JMIR Cancer

Impact Factor (2024): 3.3
Volume 12 (2026) ISSN 2369-1999 Editor in Chief: Naomi Cahill, PhD, RD

Contents

Reviews

The Impacts of Social Marketing Campaigns on Breast Cancer Awareness and Screening: Scoping Review (e88568)
Sahar Abduljawad. . . . 12

The Longevity of Mobile Apps for Cancer Recovery: Scoping Review (e82448)
Kenneth Färnqvist, Luisa Thiele, Sophie Johnsson, Pernilla Lagergren, Anna Schandl. . . . 33

Empathic AI for Patient-Centered Cancer Care: A Scoping Review of Patient Navigation, Support, and Clinical Practice (e82336)
Brianna White, Gabriela Aitken, Janet Zink, Parnian Kheirkhah Rahimabad, Fekede Kumsa, Soheil Hashtarkhani, Rezaur Rashid, Saba Kheirinejad, Tyra Girdwood, Christopher Brett, Robert Davis, David Schwartz, Arash Shaban-Nejad. . . . 42

The Application of eHealth in Symptom Management of Patients With Breast Cancer During Endocrine Therapy: Scoping Literature Review (e92030)
Chengjia Gu, Lili Shen, Shiyu Zhang, Nan Zhang. . . . 56

Multiturn Large Language Model–Based Conversational Agents for Patients With Cancer and Caregivers: Scoping Review (e96241)
Yeongrok Jeong, Hyejeon Cha, Eunyoung Suh. . . . 69

Interventions to Reduce Fear of Cancer Recurrence Among People With Cancer: Scoping Review (e81579)
Niu Li, Cynthia Wensley, Lisa Reynolds, Wal Baraza, Tatiana Leyton, Andrew Jull. . . . 84

Advanced Clinical-Based Technologies for Monitoring Physical Function in Breast Cancer Survivors: Scoping Review (e77894)
Mahtab Azhdar, Adalberto Loyola-Sanchez, Amber Wardrop, Martin Ferguson-Pell. . . . 114

Measurements and Digital Technology Solutions to Monitor Physical Activity in Patients With Pediatric Cancer: Scoping Review (e73889)
Greta Jermolenko, Guna Semjonova, Aija Klavina, Evita Dubinina, Keita Augstkalne, Klavs Balamovskis-Kalnins, Alina Cesuna, Emil Syundyukov, Dace Bertule, Madara Blumberga, Ilze Kundzina. . . . 129

AI-Enhanced Predictive Analytics to Optimize Tele-Oncology Implementation in Rural Settings: Scoping Review (e78005)
Laiba Husain, Megan Mullins, Bella Etingen, Raaed Zafar, Mediha Siddiqui, George Jackson. . . . 171

XSL•FO
RenderX

Digital Health Interventions and Psychosocial Outcomes in Women With Breast Cancer: Systematic Review (e96858)
Rocío Zúñiga-Tapia, Mercè Boixadós, Eulàlia Hernández Encuentra. . . . 202

Use of AI to Predict and Support Medication Adherence in Patients With Breast Cancer: Systematic Review (e80128)
Massimo Pezzolato, Viktorya Voskanyan, Ilaria Cutica, Chiara Marzorati, Gabriella Pravettoni. . . . 222

Viewpoints

Internet of Things–Enhanced Mathematical Oncology: Conceptual Framework for Adaptive Cancer Care Modeling (e73997)
Adeniyi Onasanya, Ramona Kyabaggu, Daryl Hepting. . . . 149

AI Agents for Multimodal Oncology Diagnosis: Toward Transparent and Traceable Clinical Decision Support (e103545)
Liuyang Yang, Liyu Shan, Xiangmei Yao, Renbin Zhao, Zengzheng Li, Shuai Feng, Yajie Wang. . . . 161

Practical Recommendations for Implementing Remote Patient Monitoring for Patients Receiving Bispecific Antibodies for Multiple Myeloma (e95949)
Cindy Varga, Eric Schaefer. . . . 1695

Original Papers

Perspectives of Adolescents and Young Adults, Caregivers, and Health Care Providers on Regional Cancer Care: Qualitative Study (e85096)
Joanne Tay, Mohammad Jarrar, Telford Yeung, Jessica Kichler. . . . 188

Improving Long-Term Adherence to Endocrine Therapy Among Breast Cancer Survivors: Development of a Multiscale Modeling and Intervention System (e68255)
Manuel Gonzales IV, Cristian Garcia-Alcaraz, Navreet Kaur, Anna Baglione, Sarah Livermon, Laura Barnes, Kristen Wells. . . . 243

Feasibility and Implementation of an eHealth Dashboard for the Remote Monitoring of Dutch Patients With Chronic Myeloid Leukemia: Multimethods Approach (e76096)
Lynn Verweij, Julie Swillens, Sanne Metsemakers, Yolba Smit, Esri Wener, Geneviève Ector, Rosella Hermens, Nicole Blijlevens. . . . 262

Cultural Adaptation of a Web-Based Ostomy Care Intervention for Hispanic Patients With Cancer and Caregivers: Mixed Methods Study (e70354)
Xiaomeng Wang, Lixin Song, Amy Vondenberger, Fei Yu, Jia Liu, Roxana Delgado, Robert Svatek, Natalia Rodriguez, Lurheinna Rivera, Mark Bonnen. . . . 280

New Technologies and Digital Health Tools in Patients With Solid Tumors and Hematological Malignancies: Cross-Sectional Multicenter Survey Study (e58823)
Alberto Lopez-Garcia, Carola Diaz-Aizpun, Beatriz Gallego-Diaz, Carolina Miranda-Castillo, Maria Yuste-Platero, Pilar Beltran-Alvarez, Cristina Carames-Sanchez, Jesus Garcia-Foncillas, Pilar Llamas-Sillero, Marta del Olmo-Rodriguez, Jorge Short-Apellaniz, Bernadette Pfang, Javier Arcos-Campillo, Raul Cordoba. . . . 298

Evaluation of GPT-5 for Esophageal Cancer Staging Using Fluorodeoxyglucose Positron Emission Tomography Maximum-Intensity Projection Images: Comparative Pilot Study (e86630)
Hiroki Maruyama, Yoshitaka Toyama, Yuya Araki, Kentaro Takanami, Masato Ito, Yumi Nakajima, Kei Takase, Takashi Kamei. . . . 307

XSL•FO
RenderX

Revealing the Physician's Task Conceptualization in Melanoma Treatment: Qualitative User Experience Study (e79294)
Eva Hartmann, Alisa Küper, Jessica Swoboda, Catharina Beckmann, Georg Lodde, Elisabeth Livingstone, Dirk Schadendorf, Sabine Sachweh. 3 2 4

Implementing UK Oncology Nursing Society–Informed Digital Symptom Triage With Episode-Based Review in Routine NHS Acute Oncology: Service Evaluation (e92586)
Mohamed Mohamed Azhar, Rishabh Ray, Sagar Sharma, Hiten Chauhan, Margaret Robinson, Yetunde Owoseni, Anupama Vijay, Sadaqat Hussain. . . . 339

Cost-Effectiveness of AI-Powered Decentralized Production for Advanced Therapies in the Hospital–Generated Chimeric Antigen Receptor T-Cell Therapy for Multiple Myeloma From a German Hospital Perspective: Early Health Technology Assessment Study (e92253)
Iris Boot, Anna Weltin, Ysanne de Graaf, Hubertus Vrijhoef. . . . 352

Detection of Cervical Nodal Metastasis in Oral Squamous Cell Carcinoma Using Elastic Scattering Spectroscopy: Ex Vivo Diagnostic Modeling Study (e103960)
Waseem Jerjes, Zaid Hamdoon, Dara Rashed, Colin Hopper. . . . 365

Exploring the Readiness for Digital Health Literacy Transformation and Intervention Preferences From the Perspectives of Patients With Cancer, Caregivers, and Health Care Professionals: Qualitative Interview Study (e77738)
Hind Mohamed, Turki Alanzi, Jon Salsberg, Mudathir Mohamed, Dervla Kelly. . . . 390

Real-World Evidence of Treatment Outcomes in Small Cell Lung Cancer: A Bayesian Mixed Effects and Competitive Risk Approach (e84042)
Luca Marzano, Asaf Dan, Adam Darwich, Luigi De Petris, Salomon Tendler, Rolf Lewensohn, Jayanth Raghothama, Sebastiaan Meijer. . . . 406

Piloting a Clinical Decision Support System for Unintended Weight Loss in Primary Care: Mixed Methods Study on Early Cancer Detection (e90885)
Javiera Martinez-Gutierrez, Sophie Chima, Barbara Hunter, Alex Lee, Lucas De Mendonca, Deborah Daly, George Fishman, Fong Lim, Benny Wang, Craig Nelson, Brian Nicholson, Jo-Anne Manski-Nankervis, Jon Emery. . . . 426

Explainable Machine Learning–Based Prediction of Progression-Free Survival in Prostate Cancer: Retrospective Cohort Study (e91510)
Hein Tun, Lin Naing, Owais Malik, Muhammad Abdullah, Thu Ta, Hanif Rahman. . . . 444

Scan, Screen, Support—A Digital Pathway for Assessing Supportive Care Needs in Oncology: Implementation Study (e82392)
Theres Fey, Nicole Thurner, Ulrike Haidn, Friederike Mumm, Birgit Haberland, Rachel Wuerstlein, Sebastian Theurich, Georg Wolfrum, Marie-Louise Troidl, Susan Müller, Claudia Bausewein, Timo Schinköthe, Volker Heinemann, Nicole Erickson. . . . 457

Patient Perspectives on AI-Powered Medical Robots in Breast and Prostate Cancer Care: Qualitative Study (e69710)
Mahiya Habib, Janet Ellis, Aaron Palachi, Melissa Korman, Tatjana Kay, Karen Barlow, Jordana DeSouza, Rosanna Macri, Abdullah Alabousi, Mehran Anvari. . . . 474

Monitoring Steps and Heart Rate Using a Withings Smartwatch in Children and Adolescents With Cancer: Validation Study (e77766)
Emma den Hartog, Marta Fiocco, Patrick van der Torre, Wim Tissing, Emma Verwaaijen. . . . 489

Online Intervention on Lung Cancer Screening Among High-Risk Individuals: Pilot Intervention Study (e89823)
Fang Lei, Feifei Huang, Yufang Guo, Sun Kim, Yang Liu. . . . 503

XSL•FO
RenderX

Large Language Models for Breast and Cervical Cancers Communication: Mixed Methods Evaluation Study Assessing Linguistic Quality, Safety, and Accessibility (e82971)
Agnik Saha, Victoria Churchill, Anny Rodriguez, Ugur Kursuncu, Muhammed Idris. . . . 516

Generative AI Chatbots as Digital Adjuncts for Sexual Health Information After Prostate Cancer in Men Who Have Sex With Men: Auto-Netnographic Study (e81745)
Mats Christiansen, Henrik Eriksson, Lisbeth Fagerström. . . . 532

Patients' Attitudes and Beliefs Toward Artificial Intelligence Use in Cancer Care: Cross-Sectional Survey Study (e81346)
Marco Santos Teles, Karolina Bryl, Susan Chimonas, Atif Khan, Andrew Epstein, Bobby Daly, Han Xiao, Jun Mao. . . . 541

Leveraging Innovative Electronic Health Record Data to Characterize Social Determinants of Health Among Survivors of Cancer in Persistent Poverty Areas: Cross-Sectional Study (e81054)
Shiori Tanaka, Laura Rogers, Lisa Zubkoff, Lori Bateman, Wendy Demark-Wahnefried, Gabrielle Rocque, Sejong Bae, Maria Pisu, James Booth, Carrie Howell. . . . 548

Digital and Paper-Based Hospital Workflows and 60-Day Mortality in Acute Leukemia: Retrospective Natural Experiment (e71306)
Christian Ramos Peñafiel, Álvaro Cabrera García, Carolina Balderas Delgado, Rocío Luna Tentle, Andrea Milán Salvatierra, Nora Godínez Cubillo, Gustavo Acosta Altamirano, Rosa Sánchez Conejo, Adán Gallardo Rodríguez. . . . 566

Assessing the Accuracy, Completeness, and Reference Quality of GPT-4 and Google for Gynecologic Cancer Information: Comparative Quantitative Study (e76471)
Alexandra Smick, Martins Ayoola, Rajeshree Rajpara, Isabel Lazo, Dana Chase. . . . 580

Experiences of Informal Caregivers Facing Cancer While Caring for Someone With Dementia: Qualitative Interview Study (e93688)
Elizabeth Travis, Mollie Price, Millie Slessor, Jacqueline Crewe, Laura Ashley. . . . 590

Identification of Early Signs of Mental Health Disorders in Older Survivors of Cancer Using Patient-Generated Health Data: Observational Study (e75050)
Georgios Petridis, Antonios Billis, Georgios Meditskos, Athina Tsanousa, Paraskevas Lagakis, Juan Naranjo, José Zenóglio de Oliveira, Vania Tavares, Panagiotis Bamidis. . . . 602

Fear of Cancer Recurrence Before Treatment in Patients Newly Diagnosed With Breast Cancer: Cross-Sectional Study (e89346)
Yun Ding, Huiting Zhang, Qiaoling Zhong, Lijuan Zhang, Huiying Qin. . . . 618

Pathways and Roadblocks in Navigating Online Cancer Communities: Qualitative Study Among Young Adult Cancer Survivors (e79893)
Qi Chen, Erin Donovan, Leyi Zhou, Lailea Noel, Barbara Jones. . . . 630

Benefits and Challenges of a Digital Exercise and Mind-Body Program During Active Cancer Treatment: Qualitative Study of Patients' Perceptions (e80075)
Karolina Bryl, Sierra Silverwood, Krupali Desai, Kelsey Schobert, Xiaotong Li, Susan Chimonas, Jun Mao, Erin Gillespie. . . . 645

Surrogate Decision-Making by Family Caregivers for Hyperthermic Intraperitoneal Chemotherapy in Gastric Cancer: Qualitative Study in a High-Volume Chinese Center (e80471)
Zheng-Ke-Ke Tan, Dan-Ni Li, Kui Jia, Wen-Zhen Tang, Xin Chen, Li Yang. . . . 660

Cancer Information Burden Among US Cancer Survivors Using Health Information National Trends Survey 2022 and 2024: Cross-Sectional Study (e101310)
Jingjing Gao, Sharon Munroe, Yue Zhang, Jason Windett. . . . 674

XSL•FO
RenderX

Predictors of Telehealth Use Among Cancer Survivors: Retrospective Study (e87403)
Jasskiran Kaur, Jeffrey Zhong, Carley Mitchell, Annie Zhang, Qian Wang, Melinda Hsu. 690

Cardiovascular Symptom Tracking Among Patients With Cancer in Cardio-Oncology Care: Qualitative Study Using the Capability, Opportunity, Motivation–Behavior (COM-B) Framework (e100279)
Roberto Benzo, Macy Tetrick, Rujul Singh, Sara St George, Vanina Pavia Aubry, Peter Washington, Anvitha Gogineni, Sanam Ghazi, Olivia Preston, Daniel Addison. 702

Factors Associated With the Use of the Lifestyle and Empowerment Techniques in Survivorship of Gynecologic Oncology (LETSGO) mHealth App in Routine Follow-Up After Gynecologic Cancer Treatment: Observational Study (e89918)
Zaklina Tarabar, Elin Børøsund, Sindre H Fosstveit, Sveinung Berntsen, Milada Hagen, Ingvild Vistad. 719

Assessment of Physical Activity Engagement and Promotion Among Cancer Survivors in Cardio-Oncology Care: Qualitative Study (e104332)
Macy Tetrick, Daniel Addison, Rujul Singh, Sara St. George, Vanina Pavia, Anvitha Gogineni, Peter Washington, Sanam Ghazi, Olivia Preston, James Fisher, Roberto Benzo. 734

Quality, Empathy, and Readability of AI Chatbot Responses to the Survivorship Needs of Adolescents and Young Adults With Melanoma: Evaluation Study (e84234)
Jordan Jafarnia, Priscilla Haff, Reece Moore, Alyssa Osheim, Katherine Riley, Sabrina Zheng, Michael Roth, Madeleine Salge, Kelly Nelson.
7 5 1

Stakeholder Perspectives on Cancer Survivors' Return to Work and Well-Being: Qualitative Interview Study (e89954)
Sara Fällman, Katarina Wijk, Anna Efverman, Linda Eklund, Per Fessé, Sven Trygged, Maria Fjell. 765

Correlates of Fitness Tracker Ownership and Use in Cancer Survivors: Cross-Sectional Survey (e92876)
Roberto Benzo, James Fisher, Macy Tetrick, Rujul Singh, Alex Osei, Jessica Krok-Schoen, Brett Nickerson, Electra Paskett, Ali Kargarandehkordi, Sachin Kumar, Sara St George, Michael Rayo, Frank Penedo, Peter Washington. 781

Symptom Management Following Radioactive Iodine Treatment in Adults With Thyroid Cancer: Usability Evaluation and Proof-of-Concept Study (e84151)
Alaina Carr, Angela Jenkins, Valeria Gómez, Janine Lorenzo, Ryan Moclair, James Pocchia, Devansh Gandhi, Kristen Miller, Kristi Graves, Jacqueline Jonklaas. 799

Clinical Feasibility of Six Inflammatory Markers for Predicting the Mortality of Patients With Cancer: Longitudinal Study (e99410)
Kangwei Wang, Ce Shi, Dingchao Xia, Ya Lin. 813

Immunotherapy Safety in Thymic Epithelial Tumors: Disproportionality Analysis of the Food and Drug Administration Adverse Event Reporting System (e76908)
Ruilian Chen, Hanrui Chen, Lingling Sun, Yang Cao, Lizhu Lin. 828

Identifying Hemophagocytic Lymphohistiocytosis and Describing Outcomes Using Computable Phenotypes: Retrospective Cohort Study (e87347)
Suheyla Ocak, Martin Yi, Agata Wolochacz, Ida Mehrdadi, Ahmed Naqvi, Sumit Gupta, Lillian Sung, Adam Yan. 843

Predicting Overall Survival in Patients With Multiple Primary Lung Cancer: Nomogram Development and Validation Study (e87275)
Wenzhi Luo, Kengliang Rao, Hongjia Chen, Cailian Hu, Shengming Liu, Jing Wang, Dongdong Zhang, Li Chen. 854

Using Latent Dirichlet Allocation Topic Modeling to Uncover Latent Research Topics and Trends in Renal Cell Carcinoma: Bibliometric Review (e78797)
Javier De La Hoz-M, Karime Montes-Escobar, Carlos Salas-Macias, Martha Fors, Santiago Ballaz. 869

XSL•FO
RenderX

Co-Ablation System for Pain Management in Bone Metastases: Retrospective Exploratory Study (e86301)
Hang Yuan, Wei-Li Xia, Ho-Young Song, Lin Zheng, Wei-Jun Fan, Hong-Tao Hu. 888

Quality of Life, Symptom Burden, and Associated Factors Among Patients With Lung Cancer in Sub-Saharan Africa: Cross-Sectional Study (e87232)
Mpho Ratshikana, Kija Malale, Abdul-Rauf Sayed, Oluwatosin Ayeni, Peedi Mathobela, Rofhiwa Mathiba, Charmaine Blanchard, Merika Tsitsi, Daniel Osei-Fofie, Anita Graham, Herry Dhudha, Nestory Masalu. 896

An Information Tool Incorporating Real-World Outcome Data for Women With Metastatic Breast Cancer Eligible for Treatment With a CDK4/6 Inhibitor: Development and Evaluation (e73156)
Ellen Engelhardt, Mariska Hackert, Anne Vogelaar, Afke de Jong, Cristina Guerrero Paez, Annette van der Velden, Christa Putker, Mariette Agterof, Ewoudt van de Garde, Cornelia van Uden-Kraan. 908

Clinician Perspectives on Integrating Mobile Sensor Data Into Cancer Care: Mixed Methods Study (e86412)
Christianna Bartel, Leeann Chen, Krina Durica, Jennifer Fedor, Akiera Palm, Tommy Selwood, Pari Lakhia, Roby Thomas, Heidi Donovan, Carissa Low. 925

Evaluation of Experiences With Ecological Momentary Assessment Among Women With Metastatic Breast Cancer: Qualitative Study (e80467)
Emily Taylor, Brenna Mossman, Mikaela Velazquez-Sosa, Zaynab Rashid, Natalie Kloster, Maureen Schwarz, Hannah Kang, Jennifer Rodriguez, Elizabeth Addington, Judith Moskowitz, Lesley Glenn, Shontè Drakeford, Roxana Guerra, Claudine Isaacs, Ami Chitalia, Christopher Gallagher, Deena Graham, Suzanne O'Neill, Claire Conley. 934

Metastasis Extraction From Non–Small Cell Lung Cancer Clinical Notes Using Large Language Model–Based Classification: Retrospective Comparative Study (e99573)
Sweta Balaji, Kiersten Campbell, Rou-Zhen Chen, Daniel Smith, Matthew Reyna, Abeed Sarker, Joshua Wallach, Ravi Parikh, Selen Bozkurt.
9 5 3

Information-Seeking Patterns and Communication Preferences Among Japanese Survivors With Cancer: Cross-Sectional Analysis (e79065)
Manae Hirayama, Daisuke Kasai, Ken-ichi Tabei. 970

Extracting Quality of Life Information of Patients Diagnosed With Breast Cancer From Health Care Online Forum Posts: Data Feasibility Study (e76044)
David Schmidt, Raoul Schubert, Brian Chen, Deborah Kuk, Valmeek Kudesia, Andreas Hinz, Philipp Cimiano. 993

Feasibility of Implementing an Electronic Patient-Reported Outcome-Based Digital Health Platform in Pediatric Oncology and Palliative Care: Insights From the MyPal4Kids Multicenter Observational Study (e80718)
Marcel Meyerheim, Lenka Radová, Anna Burns-Gebhart, Marvin Mergen, Pantelis Natsiavas, Panos Bonotis, Christina Karamanidou, Sheila Payne, Julia Downing, Julie Ling, Annette Sander, Kasra Mirzaie, Susan Dokutur, Jana Vaculíková, Petr Lokaj, Ond ej Rohleder, Norbert Graf.
1 0 1 0

Large Language Models for Supporting Clear Writing and Detecting Spin in Randomized Controlled Trials in Oncology: Comparative Analysis of GPT Models and Prompts (e78221)
Carole Koechli, Fabio Dennstädt, Christina Schröder, Daniel Aebersold, Robert Förster, Daniel Zwahlen, Paul Windisch. 1038

Physical Activity Monitoring in Adolescents and Young Adults With Cancer: Observational Feasibility Study (e87591)
Savannah Rauschendorfer, Beatrice Cantoni, Michelle Audirac, Susan Gilchrist, J Livingston, Michael Roth, Corwin Zigler, Eugenie Kleinerman.
1 0 4 8

Efficacy of Telehealth-Based Coaching to Improve Physical Activity and Overall Experience for Cancer Survivors: Secondary, Mixed Methods Analysis of a Randomized Controlled Trial (e78968)
Saif Khairat, Erin Kent, John Geracitano, Kaushalya Mendis, Zhaoqiang Zhou, Carly Bailey, William Wood. 1060

XSL•FO
RenderX

The Feasibility and Acceptability of a Remotely Delivered, Combined Exercise Intervention on Cognitive Function in Patients With Breast Cancer Following Chemotherapy: Randomized Controlled Trial (e73393)
Linda Trinh, Natalie Cuda, Catherine Sabiston, Ellen Warner, Jennifer Ryan, Arthur Kramer, Michelle Ha, Hui Xiao, Sarah O'Rourke, Edward McAuley. . . . 1075

Digital Intervention for Electronic Patient-Reported Outcomes in Advanced Cancer: Mixed Methods Study (e91416)
Christina Sauer, Simeon Sauer, Elena Doll, Stefanie Zschäbitz, Jan Raue, Maik Turni, Till Bugaj, Kiriaki Hiller, Imad Maatouk. . . . 1095

Real-World Impact of Electronic Patient-Reported Outcomes on Early Intervention Among Older Patients With Lung Cancer: Prospective Cohort Study (e97890)
Yasuhiro Nakajima, Tetsushi Maesako, Takashi Murata, Daisuke Minami, Toshiyuki Minami, Arihiko Kanehiro, Kozo Kuribayashi, Takashi Kijima. . . . 1110

“Like a Clinical Nurse Consultant in Your Pocket”—Parents’ and Health Care Providers’ Perceptions of a Co-Designed Smartphone App Following Treatment for Pediatric Cancer: Mixed Methods Evaluation Study (e90177)
Hannah Pring, Cinzia De Luca, Georgia Taylor, Chris Williams, Megan Nagarajah, David Eisenstat, Maria McCarthy. . . . 1120

Acceptance, Perceived Usefulness, and Data Sharing in Mobile Health Apps Among Patients With Breast Cancer: Cross-Sectional Survey Study (e77750)
Madeleine Flaucher, Tabea Ott, Michael Nissen, Peter Dabrock, Hanna Huebner, Matthias Beckmann, Peter Fasching, Heike Leutheuser, Bjoern Eskofier, Alexander Mocker. . . . 1137

Exploring the Adoption of Mobile Health Apps Among Patients with Head and Neck Cancer After the COVID-19 Pandemic: Secondary Analysis of Cross-Sectional Survey (e65192)
Christian Wilhelm, Matthias Scheich, Michael Winter, Johannes Allgaier, Johanna Rippstein, Sylke Zeissig, Rüdiger Pryss, Carsten Vogel, Johannes Schobel, Katrin Radeloff, Joanna Karolonek, Agmal Scherzad, Stephan Hackenberg, Rudolf Hagen. . . . 1153

Evaluation of the ALIBIRD mHealth Platform for Care of Patients With Lung Cancer: Prospective Pilot Study (e69525)
Jose Iniesta-Chamorro, Maria Sereno, Beatriz Garrido-Rubiales, Marta Villarino Sanz, Isabel Espinosa-Salinas, Sandra Falagán Martínez, Gustavo Rubio Romero, Juan Moreno-Rubio, Beatriz Tabarés, Yolanda Martín, Irene Hernández de Córdoba, Maria Morales Parga, Natalia Molinero, M Moreno-Arribas, Guillermo Reglero Rada, Ana Ramírez de Molina, Enrique Gómez, Enrique Casado. . . . 1170

Impact of a Mobile Nutrition App on Dietary Outcomes in Cancer Survivors: Pilot Feasibility Study (e79215)
Sang Park, Woo Kim, Oh Kwon, Inhyeok Yim. . . . 1195

Evaluating a Telehealth Coaching and Mobile-Based Digital Engagement Intervention for People With Cancer Using the Patient-Reported Outcomes Measurement Information System Global Health: Pilot Questionnaire Study (e72647)
Joanne Lewis, Shen Wang, Toni Jue, Tim Atkins, Scott Irwin, Puneet Nanda, Davinder Sangar, Lynette Mackenzie, Raghav Murali-Ganesh. . .
1 2 1 2

Gait Changes After a Mobile Health Exercise Intervention in Older Adults With Myeloid Neoplasms: Single-Arm Pilot Trial (e80909)
Ying Wang, Marielle Jensen-Battaglia, Po-Ju Lin, Ian Kleckner, Elisabeth Hayward, Jason Mendler, Jane Liesveld, Marissa LoCastro, Cait Fallone Sharma, Soroush Mortaz Hedjri, Richard Dunne, Karen Mustian, Kah Loh. . . . 1224

A Multilingual AI-Driven Conversational System for Cervical Cancer Awareness and Risk Assessment: Development and Evaluation Study (e103866)
Temitope Olorunfemi, Shafat Rashid, Patricia Khashayar. . . . 1235

Beta Testing an mHealth Symptom Management Intervention for People With Advanced Cancer: Single-Cohort Feasibility and Acceptability Study (e94676)
Sharla Wells-Di Gregorio, Tanya Smit, Sarah Baltimore, Don Benson, Steve Clinton, Ashley Davenport, Kari Kendra, Carolyn Presley, Donald Marks. . . . 1249

XSL•FO
RenderX

User Experiences of a Chatbot for Supporting the Self-Management of Peripherally Inserted Central Catheter for Chemotherapy: Mixed Methods Study (e81026)
Bokyung Jo, Su Kim, Min Kim, Nayeon Kim, Ayoung Lee, Hoyoung Kim, Mangyeong Lee, Juhee Cho. 1264

Supporting Patients With Breast Cancer and Providers Through Treatment and Survivorship: Multimethod Implementation Study of the MyJourney Platform (e87973)
Isabella Herrington, Julie Makarski, Krystle Amog, Leigh Hayden, Kate Mossman, Norna Abbo, Nicole Jedrzejko, Fahima Osman, Monika Kastner. 1281

ChatGPT Versus DeepSeek for Breast Cancer Information Retrieval: Quantitative Comparative Study (e72839)
Rima Hajjo, Dima Sabbah, Sanaa Bardaweel. 1305

Planning and Developing a Symptom Diary Intervention for Breast Cancer Survivors With Concerns About Medication Brands (ENABLE Study): User-Centered Design Approach (e91234)
Yolanda Eraso, Duncan Stewart, Ruth Edwards, Bal Nanray, Jacqueline Cochrane. 1322

Patient-Reported Outcomes on Quality of Life and Psychological Distress After Focal High-Intensity Focused Ultrasound for Prostate Cancer: Prospective Multicenter Study (e91984)
Toni Franz, Cornelia Hartung, Roman Ganzer, Sigrun Holze, Jens-Uwe Stolzenburg. 1342

Reinforcement Learning–Based Digital Therapeutic Intervention for Postprostatectomy Incontinence: Development and Pilot Feasibility Study (e83375)
Fan Fan, Hao Huang, Jingwen Yan, Chao-Yue Xu, Xiuhua Wu, Chunmei Zhou, Dandan Wen, Hai Huang, Ho Li, Yihong Qiu. 1365

Hashimoto Thyroiditis and Progression of Papillary Thyroid Cancer: 10-Year Retrospective Cohort Study (e80535)
Xin Li, Xiangyun Yao, Rui Shan, Fang Mei, Shibing Song, Bangkai Sun, Chunhui Yuan, Zheng Liu. 1383

Exploring the Experiences and Perspectives of Patients With Early Breast Cancer, Caregivers, and Health Care Professionals: Italian Social Media Listening Study (e73371)
Matteo Suter, Rosanna D'Antona, Manuelita Mazza, Enrica Francavilla, Alice Messi, Abhinav Maurya, Shiva Padhi, Diletta Valsecchi. 1394

Acceptability of Sharing Internet Browsing History for Cancer Research: Think-Aloud and Interview Study (e82009)
Nicola Gradwell, Mel Ramasawmy, Sanjula Arora, Manoj Mistry, Reshma Punjabi, Christina Derksen, Suzanne Scott. 1409

Health Care Providers' Perceptions of Unmet Needs Among African American Cancer Caregivers: Qualitative Investigation Among US Medical Professionals (e76266)
Brad Love, Gerold Dermid, Sean Upshaw, Amy Stark. 1423

Scalp Cooling as a Biopsychosocial Intervention in Patients Receiving Highly Alopecia-Inducing Chemotherapy: Prospective Single-Arm Study (e94635)
Dörthe Schaffrin-Nabe, Anke Josten-Nabe, Adrian Heinze, Andrea Tannapfel, Merle Schaffrin, Rudolf Voigtmann. 1443

A Multimethod Evaluation to Assess Feasibility, Acceptability, and Preliminary Efficacy of HPVVaxFacts, a Tailored Mobile Web App, for Parents With Unvaccinated Children: Pilot 2-Arm Randomized Controlled Trial (e78910)
Jennifer Cunningham-Erves, Pamela Hull, Amanda Dempsey, Tatsuki Koyama, Lili Sun, Mariya Chalise, Elizabeth Stewart, Jessica Jones, Douglas Landsittel, Meredith Smalls, Freneka Minter, Lindsay Mayberry, Lora Harnack, Michael Hook, Janet Cates, Tammy Van Maanen, Taylor Burnett, Consuelo Wilkins. 1459

XSL•FO
RenderX

The Influence of Social Media Messaging on Human Papillomavirus Vaccine Attitudes and Confidence Among Adolescent Males: Group Discussion Study (e82210)
Neha Trivedi, Rachel Van Vleet, Czesia Eid, Amelia Burke-Garcia, Amy Leader. . . . 1481

AI-Driven Patient Screening for Clinical Trials in Pancreatic Cancer: The PANCR-AI Pilot Retrospective Comparative Study (e80268)
Arthur Claessens, Alizée Simon, Agathe Manchart, Laura Desmonts, Cassandre Leguay, Jean-Christophe Faivre, Sébastien Cambier, Julia Salleron, Aurélien Lambert. . . . 1496

Navigating the Complexity of Lung Cancer Surveillance Practices: Qualitative Pilot Study on Provider Perspectives (e80659)
Jenny Woo, Sydney Conover, Caroline Gray, Shipra Arya, Julie Wu. . . . 1506

Awareness of Lung Cancer Signs and Symptoms, Risk Factors, and Screening in Jordan: Cross-Sectional Study (e86390)
Ahmad Yousef, Mohammad Saleh, Hani Shihadeh, Sanad Ahmad, Mo'tasem Alaiwah, Heba Shbailat, Karen Alsyouf, Malek Altaamreh, Eman Alshawish, Rasmieh Al amer. . . . 1518

Cancer Clinician Experiences With Telehealth Across Oncology Specialties: Cross-Sectional Survey Study (e94610)
Benjamin Baker, Sakshith Chintala, Sarah Azari, James Murphy, Hala Madanat, Maria Martinez, Richard Cripps, Aditya Bagrodia, Humberto Parada Jr, Juan Javier-DesLoges. . . . 1530

Financial Toxicity of Hematologic Malignancy Therapies, Including Cellular Therapy and Its Impact on Access to Care: Prospective Pilot Study (e68101)
Lemchukwu Amaeshi, Lauren Laufer, Kateryna Fedorov, Alex Sisto, David Levitz, Rosmi Mathew, Karen Wright, Dennis Cooper, Mendel Goldfinger, Ioannis Mantzaris, R Sica, Kira Gritsman, Marina Konopleva, Ridhi Gupta, Noah Kornblum, Chenxin Zhang, Mimi Kim, Amit Verma, Bruce Rapkin, Aditi Shastri, Nishi Shah. . . . 1548

Patient and Clinician Perspectives on Expanding Telehealth Use for Older Adults Across the Cancer Control Continuum: Mixed Methods Study (e73058)
Robin Higashi, Bella Etingen, Jessica Lee, Suzanne Cole, John Mansour, Alice Zhao, Timothy Hogan. . . . 1563

Understanding Social Media Strategies to Improve Engagement in and Equitable Dissemination of Online Cancer Nutrition Information: Cross-Sectional Study (e69886)
Ellie Harrison, Sara Buzali-Soto, Eileen Rillamas-Sun, Heather Greenlee, Echo Warner. . . . 1575

Patient and Provider Experiences With Precision Oncology: Qualitative Descriptive Study at the Department of Veterans Affairs (e84858)
Daniel Becker, Kenneth Csehak, Alexander Barbaro, Christian Miller, Antoinette Vo, Stefanie Roman, Danil Makarov, Scott Sherman, Allison Squires. . . . 1590

Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer Using Routinely Collected Clinical Data: Multicenter Retrospective Machine Learning Model Development and External Validation Study (e101162)
Gehui Xu, Xiaohe Sun, Jiaxin Jiang, Jin Sun, Liu Li, Haibo Cheng, Yuekun Zhu, Haiyi Liu. . . . 1605

Psychosocial Distress and Radiation Treatment Interruptions in Patients With Cancer Using Machine Learning and Regression: Retrospective Cohort Study (e92899)
Soheil Hashtarkhani, Christopher Brett, Fekede Kumsa, Rezaur Rashid, Parnian Kheirkhah Rahimabad, Saba Kheirinejad, Robert Davis, David Schwartz, Arash Shaban-Nejad. . . . 1623

Machine Learning–Based Survival Prediction Models for Young Patients With Gastric Cancer: Model Development and Validation Study (e86418)
Ha Kang, Wooyeong Jang, Minsam Ko, Kwang Ryu. . . . 1637

XSL•FO
RenderX

Prognostic Value of Dynamic Lactate Dehydrogenase Trends in Immunotherapy for Advanced Esophageal Squamous Cell Carcinoma: Retrospective Cohort Study (e73576)
Baidong Zhang, Zhichao Kang, Yi Ding, Wang Jing, Alei Feng, Renya Zeng, Jianan Li, Yi Zhao, Yuanliu Nie, Wentao Zhang, Lu Sun, Zhe Yang. . . . 1648

Explainable AI for Predicting Mortality Risk in Metastatic Cancer: Retrospective Cohort Study Using the Memorial Sloan Kettering-Metastatic Dataset (e74196)
Polycarp Nalela, Deepthi Rao, Praveen Rao. . . . 1658

Assessment of Predictive Factors That Shorten Duration of Treatment in Patients With Multiple Myeloma Using AI: Real-World Longitudinal Study Using Data From Medical Data Vision Claims Database (e75586)
Hiroshi Handa, Tadao Ishida, Shuji Ozaki, Shinsuke Iida, Kittima Wattanakamolkul, Chika Sakai, Kenichi Kato, David Bin-Chia Wu, DaeYoung Yu, Shota Nemoto, Yasuho Yamashita, Takuma Shibahara. . . . 1678

Patient Perspectives

Mediastinal Germ Cell Tumors: Up Close and Personal (e93678)
B Bequette. . . . 486

Short Paper

The Role of Resilience in Coping With Future Uncertainty Among People With Brain Tumors: Cross-Sectional Study (e71674)
Li-Ting Longcoy, Shu-Yuan Liang, Ardith Doorenbos. . . . 561

Letters to the Editor

Authors' Reply: Urological Content Accuracy as an Unmeasured Dimension in Generative AI Chatbot Communication About Sexual Health After Prostate Cancer (e109004)
Mats Christiansen, Henrik Eriksson, Lisbeth Fagerström. . . . 1542

Urological Content Accuracy as an Unmeasured Dimension in Generative AI Chatbot Communication About Sexual Health After Prostate Cancer (e108419)
Bilgi Isler, Ahmet Bayraktar. . . . 1544

Adapting Interventions for Home Hospice Caregivers Using Digital Health Innovation (e81589)
Qing Huang. . . . 1546

Corrigenda and Addenda

Correction: Social Media as a Platform for Cancer Care Decision-Making Among Women: Internet Survey-Based Study on Trust, Engagement, and Preferences (e87683)
Anna Johnson, Grace Longfellow, Clara Lee, Benjamin Ormseth, Gary Skolnick, Mary Politi, Yonaira Rivera, Terence Myckatyn. . . . 1603

XSL•FO
RenderX

Research Letters

Evaluation of a Retrieval-Augmented Generation–Based Large Language Model for Evidence-Based Herb and Supplement Information in Cancer Care (e86073)
Yen-Nien Hou, Jyothirmai Gubili, Pulkit Jain, Chun Lam, Avijit Chatterjee, Jun Mao. 1706

Participation in Cervical and Breast Cancer Screening Among Working Women in Japan: Panel Survey Study (e86885)
Miho Satoh, Naoko Sato. 1710

Remote Patient Monitoring Use Among Commercially Insured Adults With Cancer (e84788)
Joseph Joo, Changchuan Jiang, Jessica Billig, Nadia Lieu, Shining Yang, Yujia Jin, Arthur Hong, Joshua Liao. 1717

Tutorial

Development of a Web-Based Experiential Learning Intervention for the Public to Reduce Cancer Stigma: Tutorial on the Application of Intervention Mapping (e71166)
Miyako Tsuchiya, Keiichiro Adachi, Akiko Kimata, Kaori Kumagai. 1721

XSL•FO
RenderX

The Impacts of Social Marketing Campaigns on Breast Cancer Awareness and Screening: Scoping Review

Sahar Abduljawad, MPH, DrPH

Health Administration And Medical Information, College of Al-Leith Health Sciences, Umm Al-Qura University, Al-Abdiya Campus, Makkah Al-Mukarramah, Makkah, Saudi Arabia

Corresponding Author:
Sahar Abduljawad, MPH, DrPH
Health Administration And Medical Information, College of Al-Leith Health Sciences, Umm Al-Qura University, Al-Abdiya Campus, Makkah Al-Mukarramah, Makkah, Saudi Arabia

Abstract

Background: Promotional activities, such as social marketing approaches, are important for raising cancer awareness and increasing breast cancer screening participation. Such interventions are designed to facilitate change and maintain positive health behaviors at both the community and society level. With the growing use of these promotional activities in public health, it is pertinent to map available studies to inform future evaluations and the development of effective interventions.

Objective: This study aimed to review existing studies on how social marketing interventions have been used to promote breast cancer screening and early diagnosis. The specific objectives were to assess the characteristics of social marketing interventions aimed at promoting breast cancer screening, identify the methodological approaches used in evaluating the interventions, identify the outcomes used to measure the impact of social marketing interventions for breast cancer screening, and assess the impacts of social marketing interventions when promoting breast cancer screening.

Methods: A comprehensive literature search was performed in 6 electronic databases to identify qualitative and quantitative evaluations of social marketing interventions targeting breast cancer screening and early diagnosis. Titles and abstracts were reviewed independently by two reviewers, followed by data extraction and verification. The methodological quality was evaluated using the Cochrane Handbook for Systematic Reviews of Interventions. Outcomes were categorized using an adapted version of key performance indicators and metrics related to social media use in health promotion.

Results: Overall, 15 research articles were eligible and analyzed. The 4 types of social marketing interventions identified were health professional and educational strategies, social media or mass media campaigns, community outreach campaigns, and video-based marketing. The predominant intervention source was public health government bodies, followed by regional health services, cancer charities, and educational institutions. Most interventions were designed to influence breast cancer screening behaviors by raising awareness through advertisements, educational interventions, and fundraising activities. Most interventions reached a large audience (>500 people), but participants had mixed awareness and perceptions toward social marketing strategies, and the outcomes were shaped by demographic factors such as cultural backgrounds, age, education, marital status, and occupation. Knowledge gaps, cost concerns, and the lack of physician recommendations also constituted key barriers to screening uptake. Behavior change outcomes were sparingly assessed in most interventions, thereby contributing to limited understanding of mechanisms affecting screening uptake.

Conclusions: Available evidence suggests that social marketing interventions may enhance breast cancer screening and early diagnosis. However, robust evaluations of these interventions and assessment of behavior change outcomes could be performed in future research by using well-established evaluation frameworks and reporting guidelines. Developers and funders of these campaigns should also consider using an evaluation plan before conducting the actual campaign.

(JMIR Cancer 2026;12:e88568) doi:10.2196/88568

KEYWORDS

social marketing; social media; breast cancer; awareness; diagnosis; screening

Introduction

Social marketing campaigns (SMCs) are widely accepted approaches and tools for enhancing public awareness of diseases and altering health behaviors [1]. SMCs entail applying marketing principles to facilitate actions in the pursuit of efficient, effective, fair and sustained social transformation, both individually and collectively [1,2]. According to Thompson and Heinberg [3], social marketing applies marketing strategies to develop interventions that enhance individual and societal well-being.

XSL•FO
RenderX

Despite the similarity to conventional commercial marketing in terms of strategic approach and goal [4], SMCs emphasize voluntary behavioral change and broader benefits directed toward the community and society [1,5]. By shifting from simply providing information, SMCs assist66 with designing a comprehensive, target-centric experience that makes health behaviors highly desirable and easily adopted [1]. Although the product of commercial marketing is specific, social marketing products or services are mostly invisible and conceptual.

Social and health interventions can be enabled by using advertising techniques that promote related commercial products [1]. Typical examples of social marketing include disaster relief and recovery [6], organ donation [7], campaigns promoting recycling [8], and health promotion programs delivered via social media platforms . Social media has become extremely popular, estimated to be used by more than one-half of the global population [[9]. Hence, apart from being widely used for public engagement, it has increasingly extended into social marketing strategies for driving public health interventions [10]. Such integrations can shape behavior by targeting emotional and cognitive responses, triggering discussions, and altering social norms. In addition, the combination of social marketing and social media approaches allows rapid dissemination of health information, at a low cost, to a large audience of people across wide geographical locations [11].

In this article, the term “intervention” refers to a spectrum of activities ranging from national awareness-raising public health campaigns to targeted activities that use social marketing approaches to deliver information to specific audiences, specifically to enhance breast cancer screening decisions and uptake. The effectiveness of these social marketing strategies may differ depending on the approach used, whether official, emotive, or persuasive. The source of the message is pivotal in its persuasiveness, as sources that hold higher status and are familiar to the recipient are more likely to be recalled and internalized [12].

Accumulated evidence suggests that raising awareness and promoting the prevention of cancers has several challenges. The barriers to cancer screening uptake among susceptible individuals can be resolved via interventions that educate people about the benefits of early detection and screening [5,10]. Promotional activities, such as social marketing approaches, are critical for raising cancer awareness and likely to increase screening participation among high-risk persons, thereby reducing related morbidity and mortality rates [10]. The goal of promotional activities is to facilitate change and maintain positive health behaviors at the community and society levels.

In terms of breast cancer, the level and impact of social marketing strategies on screening decisions and uptake are underreported despite the extensive barriers to screening and adherence in vulnerable populations [13,14]. Nevertheless, empirical studies have demonstrated the effects of awareness programs on health behaviors for other diseases and disorders. Examples include the strong empirical evidence on the reduction of smoking prevalence and promotion of smoking cessation following mass media campaigns [15]. It is estimated that public health campaigns designed to address coronary heart disease have led to a 30% reduction in cholesterol in the past 3 decades [16].

Given these social marketing effects, the potential benefits of awareness raising and promotional activities on breast cancer screening should not be underestimated, thus warranting further investigation. Many scholars have attempted to explore the effectiveness of promotional campaigns in raising awareness and motivating screening for breast cancer and other cancers [17,18]. The outcomes from these studies have varied. The perceived credibility of breast cancer information is shaped by factors such as education, socioeconomic status, age, health attitudes, health status, and cultural background [19,20].

Moreover, despite the increasing advocacy for promotional campaigns, there is little evidence to support their efficacy in improving breast cancer screening and early diagnosis. Review publications describing the various approaches in which social marketing strategies are used to promote breast cancer screening and how such interventions might foster behavior change are lacking. Previous scoping and systematic reviews looking at a range of interventions for breast cancer screening focused on social media-based and mobile health (mHealth) technology [5]. However, the use of social marketing interventions for public health is rapidly evolving, reflecting the need for a broader mapping of available studies to inform future evaluations and the development of effective interventions.

This study aimed to review existing studies on how social marketing interventions have been used to promote breast cancer screening and early diagnosis. The specific objectives were to assess (1) the characteristics of social marketing interventions aimed at promoting breast cancer screening, including methodological approaches and outcomes measures, and (2) the impacts of social marketing interventions on promoting breast cancer screening. By emphasizing nonclinical initiatives, this study reviewed SMCs with the potential to be implemented at the community level to promote breast cancer screening and preventive behaviors.

Methods

Study Design

This scoping review was conducted in accordance with PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [21] and guided by the scoping review framework by Arksey and O’Malley [22]. The review focused on the use of social marketing interventions for promoting breast cancer screening and positive behavioral changes in vulnerable populations. The review was designed to address the following research questions: (1) What are the characteristics of social marketing interventions that aim to promote cancer screening? (2) What methodological approaches have been used to evaluate the interventions? (3) What are the outcomes used to measure the impact of interventions? and (4) What are the impacts of social marketing interventions in promoting breast cancer screening?

Search Strategy and Search Terms

A comprehensive literature search was conducted in 6 electronic databases: Web of Science, PubMed, EBSCOhost (via

XSL•FO
RenderX

CINAHL), MEDLINE, PsycINFO, and Scopus. The search included studies using quantitative, qualitative, and mixed methods designs to ensure a broad retrieval of relevant literature. To focus on recent advancements and growing research on social marketing strategies for breast cancer screening uptake and preventive behaviors, the search was limited to articles published between 2010 and 2025.

The literature search was performed from August 2025 to October 2025. In addition to the database search, snowballing techniques were applied, including backward reference checking (screening reference lists of selected articles) and forward reference searching using the "cited by" function in Google Scholar. These manual searches were performed to retrieve possible articles that might be missed from the electronic databases. Aligning with the research objectives, the following 6 key search components were identified: (1) breast cancer; (2) screening or testing; (3) SMC, promotional activities, or advertisement; (4) social media, mass media, or mass communication; (5) intention to screen or test; and (6) behavior change or preventive behavior.

Alternative keywords were used for each component and combined using the Boolean operators "OR" within categories and "AND" between categories. Boolean operators were not applied in Google Scholar. Additionally, a truncation symbol (*) was used where applicable to capture word variations, as detailed in Tables 1 and 2.

Table . Search strategy for the literature review in the following databases: Web of Science, PubMed, EBSCOhost (CINAHL Complete), Scopus, MEDLINE, and PsycINFO.

Main keywords	Search strings
Breast cancer	"Breast cancer" OR "Breast cancer patients" OR "Women with Breast Cancer"
Screening, testing	"Breast cancer screening" OR "Breast cancer testing" OR "Screening" OR "Testing
Social marketing campaign, promotional activities, advertisement	"Social marketing" OR "Campaign" OR "Promotion" OR "Advert"
Social media, mass media, mass communication	"Social media" OR "Mass media" OR "Mass communication" OR "Communication"
Intention to screen, test	"Intention to screen" OR "Intention to test"
Behavior change, preventive behavior	"Behavior change" OR "Preventive behavior"

XSL•FO
RenderX

Table . Specific search strings used in each database for the literature search.

Search number	Specific search query and string	Search engine(s)
1	((Breast cancer screening OR Breast cancer test OR Breast cancer testing OR intent to test)) AND (Social marketing OR social marketing campaign OR mass communication OR mass media OR mass media campaign)	Scopus
2	((Breast cancer screening [Title/Abstract] OR Breast cancer test[Title/Abstract] OR Breast cancer testing[Title/Abstract] OR intent to test[Title/Abstract])) AND (Social marketing campaign [Title/ Abstract] OR mass communication [Title/Abstract] OR mass media campaign [Title/Abstract])	PubMed
3	((Breast cancer screening [Title/Abstract] OR Breast cancer test [Title/Abstract] OR Breast cancer testing [Title/Abstract] OR intent to test [Title/Abstract])) AND (Social marketing [Title/Abstract] OR social marketing campaign [Title/Abstract] OR mass communication [Title/Abstract] OR mass media [Title/Abstract] OR mass media campaign [Title/Abstract])	Web of Science
4	((Breast cancer screening[Title/Abstract] OR Breast cancer test[Title/Abstract] OR Breast cancer testing[Title/Abstract] OR intent to test[Title/Abstract])) AND (Social marketing[Title/Abstract] OR social marketing campaign[Title/Abstract] OR mass communication[Title/Abstract] OR mass media[Title/Abstract] OR mass media campaign[Title/Abstract] OR exchange theory[Title/ Abstract] OR audience segmentation[Title/Abstract]	EBSCOhost (CINAHL Complete), MEDLINE, PsycINFO

Study Selection and Screening

The study included patients with breast cancer and those at risk who had participated in interventions involving any type of social marketing strategy, such as health promotion activities and community awareness programs. Titles, abstracts, and full-text articles were screened using PDF content to ensure alignment with the study objectives. Selected articles were subsequently stored in a Zotero database for reference management.

Eligibility Criteria

Articles were included if they met the specific criteria outlined in Table 3, which included the language and publication period, study design, intervention and outcomes, and target population. Meanwhile, studies not fulfilling the criteria listed in Table 3, not published in English, not focusing on breast cancer patients, or not involving cancer survivors or health professionals were excluded. Review articles, online materials, conference proceedings, nonoriginal articles, and articles that evaluated static internet pages, such as blog posts, were also excluded.

Table . Criteria and description considered for eligible studies.

Criterion	Description
Language and publication period	Original research articles or short communications written in English, published between 2011 and 2024
Study design	Interventional studies, including quantitative, qualitative, and mixed methods
Intervention and outcomes	Studies had to examine social marketing interventions/campaigns/programs and address at least one of the following outcomes: breast cancer screening, early diagnosis, preventive practices, or help-seeking for potential breast cancer symptoms.
Target population	Studies had to focus on women with breast cancer or high-risk populations.

Screening Process and Data Extraction

Title and abstract screening were conducted by 2 researchers, followed by full-text assessment for eligibility. Disagreements were resolved through consensus, with a third reviewer serving as a mediator. An interrater reliability analysis was performed for title and abstract screening, and an acceptable level of agreement was achieved (Cohen α=0.72). Data extraction was

XSL•FO
RenderX

conducted independently by 2 research assistants to ensure accuracy and consistency. The lead researcher predefined the key data points to be recorded. A structured data extraction form was used, capturing the following details: authors' names, research objectives, study location and design, population of interest, interventions applied, and main findings. Any discrepancies during data extraction were resolved using the same consensus approach described in the screening process, with a third reviewer providing mediation when necessary.

Assessment of Risk of Bias

Although quality appraisal is not always required in scoping reviews due to the diversity of study designs [23], this review used a risk of bias assessment given its specific focus on social marketing strategies or programs for women with breast cancer and high-risk groups. The methodological quality of each study was evaluated using the Cochrane Handbook for Systematic Reviews of Interventions [24]. The assessment considered 5 key domains: sequence generation, allocation concealment, adjustment for confounders, blinding, completeness of data, and selective reporting.

Data Synthesis

The extracted data and outcomes from the reviewed studies are shown in Table 3. We categorized the outcomes using an adapted version of key performance indicators and metrics related to social media use in health promotion [25]. The main aspects include insights, exposure and reach, and various levels of engagement. Information on the design, theoretical foundation, and delivery methods of interventions was also extracted from the reviewed studies. Furthermore, information regarding the potential mechanisms by which social marketing interventions shape behavior change was gathered by using the behavior change technique (BCT) taxonomy described by Michie et al [26].

The extracted data were analyzed using a narrative approach and thematic analysis, as recommended for scoping reviews. An inductive approach was used to identify and synthesize key themes based on the main findings from each study. The data synthesis process involved familiarization and coding, performed independently by 2 research assistants. The next step entailed developing the themes, whereby systematic coding was used to generate overarching themes. The last stage was validation and consensus, in which all the themes were reviewed by the authors to ensure consistency with the extracted data. Discrepancies were also resolved through discussion. All data management and analysis procedures were conducted using Excel (Microsoft Corp).

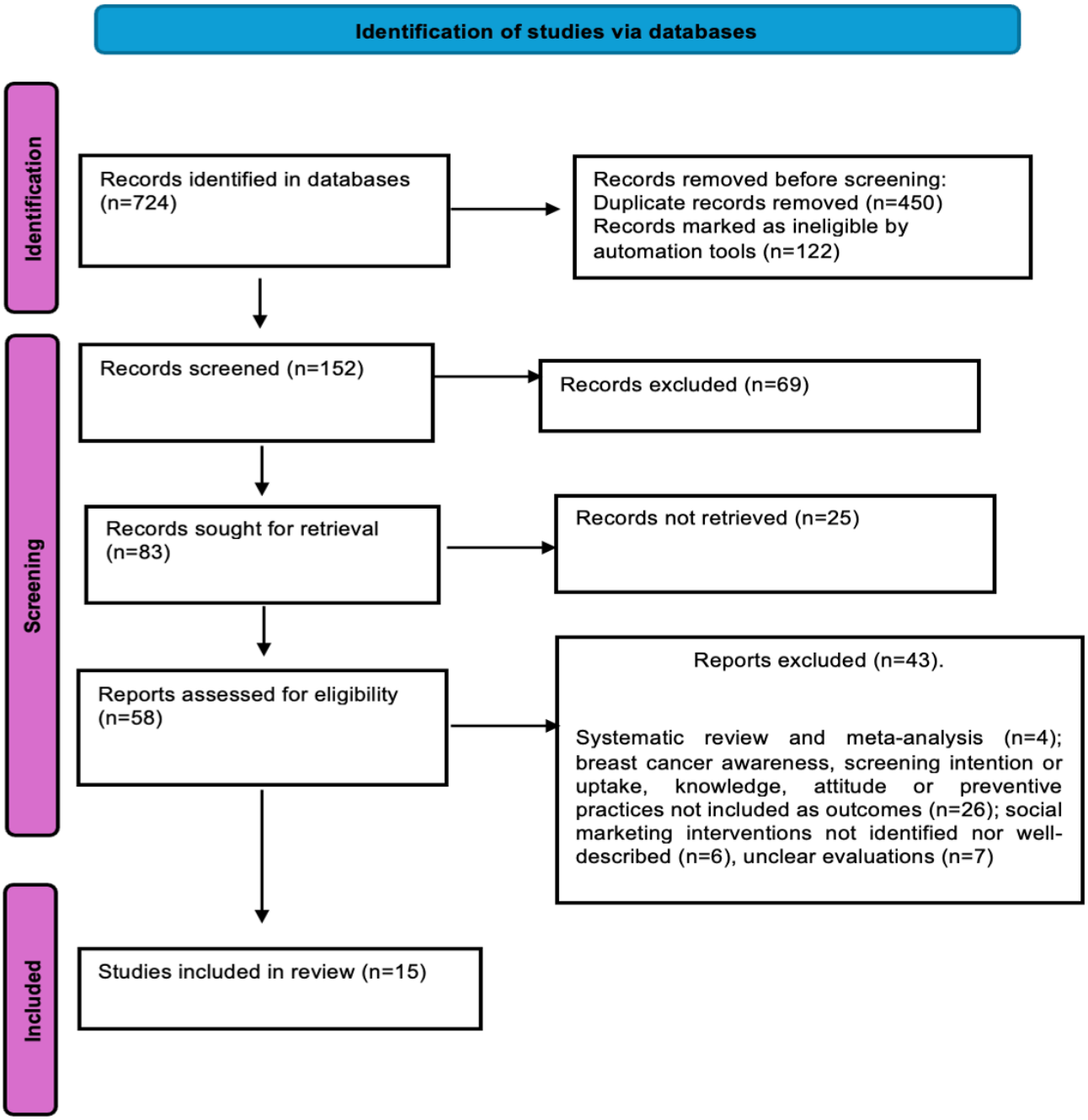
Results

Search Results

The literature search retrieved 724 articles from 6 databases. More than one-half (572/724, 79%) of the results were removed before screening, as 450 were identified as duplicates and 122 records were marked as ineligible. Upon title and abstract screening, 69 records were excluded. In the final stage, 15 of the 58 articles were eligible and included in the qualitative and thematic analysis. The reasons for excluding 43 full-text articles are outlined in Figure 1.

XSL•FO
RenderX

Figure 1. Flowchart of the scoping review literature search and screening process.



Studies and Participants' Characteristics

Table 4 outlines the characteristics of the reviewed studies. All the articles were original studies published in peer-reviewed journals. Most of the studies were conducted in the United Kingdom (n=4), and 1 each was conducted in Greece, Malaysia, Romania, Saudi Arabia, the United Arab Emirates, India, the United States, Israel, Turkey, China, and South Korea. In terms of study design, 4 studies were either single or multicenter randomized controlled trials (RCTs), 4 were quasi-experimental designs, 4 comprised a combination of cross-sectional surveys and pre- and posttest designs, 2 were case studies, and 1 article involved content analysis and in-depth interviews.

XSL•FO
RenderX

Table . Characteristics and main findings from the 15 studies on social marketing campaigns.

Search number	Reference	Objective	Study design, sample population	Sample population	Mode of delivery	Specific instruments and interventions	Main findings
1	[18]	To examine the effects of a TV-led breast cancer mass media campaign on the awareness of breast cancer–related factors and intention to engage in preventive behaviors	3 campaigns performed over 2 weeks: precampaign, followed immediately by campaign wave 1 and campaign wave 2 in the northeast of England	572 in the precampaign; 576 in wave 1; 552 in wave 2	Cancer charity	TV advertisements showing the connection between breast cancer and alcohol consumption, with a voiceover communicating the various types of cancer caused by alcohol; target online audience also reached via advertisements shown on promoted posts on Facebook	Awareness of breast cancer and readiness to engage in preventive behaviors increased significantly postcampaign (45% increment) compared with the precampaign level, especially after wave 2. Awareness of alcohol as a cancer risk factor and support for related policies increased significantly between surveys.
2	[27]	Effects of an early presentation (PEP[a]) intervention on breast cancer knowledge, skills, and motivation for early presentation into care among older women	Randomized controlled trial	867 women 67 - 70 years old attending a breast cancer screening program	Regional health services	PEP intervention using a 10-minute interaction with a radiographer	At 2-year follow-up, the intervention increased the proportion of participants who were aware of breast cancer and ready to perform early screening compared with those in usual care (21% versus 6%; OR[b] 8.1, 95% CI 2.7% - 25.0%)
3	[28]	Phase 2 assessment of the impact of a brief health professional–delivered intervention on breast cancer awareness and prevention intention among older women	Multicenter clinical trial; mammographers trained to deliver a brief health intervention for older women (n=495) in 4 screening centers, with 1 used as a control (no intervention); breast cancer awareness and prevention intention assessed at baseline and 1 month postintervention	875 older women (no intervention)	Regional health services	The intervention was based on the English National Health Service Breast Screening Program, designed to equip women to present early with breast symptoms, as a critical aspect of the concluding invited mammogram at an estimated age of 70 years.	A total of 27 mammographers were trained. Overall, breast cancer awareness increased 7-fold at 1 month in older women exposed to the intervention relative to a 2-fold increase in the control group (OR 15.2, 95% CI 10.0 to 23.2).

XSL•FO
RenderX

Search number	Reference	Objective	Study design, sample population	Sample population	Mode of delivery	Specific instruments and interventions	Main findings
4	[29]	To identify the breast cancer screening advertisements preferred by women at risk of breast cancer	A hybrid study design, comprising content analysis and repertory grid interviews; first approach: identified pertinent messages mostly promoted in breast cancer screening advertisements; interview aimed to elucidate women's perceptions regarding the messages; 13 different breast cancer screening advertisements presented to assess level of authoritativeness, credibility, emotional appeal, and persuasiveness; conducted in Greece	32 women	Charity body, celebrities	13 breast cancer screening advertisements	Some women found some of the breast cancer screening promotional messages, including award-winning advertisements, offensive and failed to connect with the visual design of these adverts. Respondents were more motivated to perform screening when exposed to positive advertisements, capturing reliable data on breast cancer data and endorsements from celebrities who genuinely support the cause.

XSL•FO
RenderX

Search number	Reference	Objective	Study design, sample population	Sample population	Mode of delivery	Specific instruments and interventions	Main findings
5	[30]	To assess the effects of a mass media campaign on breast cancer symptoms awareness and screening uptake	Pre and posttest experimental study (quasi-experimental design) in Malaysia	676 women aged >40 years from randomly selected households (intervention group)	Public health government bodies	A culturally adapted mass media campaign (TV, radio, print media, and social media); outcomes: awareness of breast cancer symptoms, self-efficacy to identify breast cancer symptoms and clinical outcomes, and campaign reach; mammogram screening and clinical breast examination data collected from hospitals and clinics	At least one campaign material was identified by most participants (65.2%); Chinese women (OR 0.25, 95% CI 0.15 to 0.40) and those aged >70 years (OR 0.47, 95% CI 0.23 to 0.94) demonstrated lower odds of being reached during the campaign relative to Malays and younger women; higher level of postcampaign awareness was observed among participants who recognized the campaign compared with nonrecognizers, especially for primary symptoms such as discharge or bleeding from nipple (79.7% vs 55.3%) and “a lump or thickening in your breast” (88.9% vs 62.1%); none of the demographic variables influenced improvement in symptom awareness scores.
6	[31]	To explore the impact of social media awareness campaigns on the early detection of breast cancer	Online survey conducted after periodic campaigns in Romania	1945 participants selected from a Facebook community (intervention) and 289 participants in a control group	Public health government bodies	Intervention comprising of periodic campaigns on the significance of breast cancer screening for prevention and early diagnosis	Intervention group had significant improvement in breast cancer screening awareness and preventive practices relative to the control group.

XSL•FO
RenderX

Search number	Reference	Objective	Study design, sample population	Sample population	Mode of delivery	Specific instruments and interventions	Main findings
7	[14]	To assess the efficacy of an institutional breast cancer health campaign on the knowledge, attitudes, and practices of female employees	Cross-sectional descriptive study. in Saudi Arabia	290 female workers from various university departments	University	A group-based breast cancer awareness campaign, supported with posters, brochures, lectures and advertisements, followed by an online self-administered questionnaire	Knowledge and positive attitudes toward breast cancer increased significantly postintervention; participants who had previously attended any prior breast cancer health campaign scored higher knowledge (mean 7.45, SD 0.98) and attitude (mean 5.30, SD1.48) scores than those who had never attended (knowledge: mean 6.10, SD 1.10; attitude: mean 4.10, SD 1.41); to facilitate screening behavior, counseling services and the promotion of existing breast cancer screening are necessary; diverse, culturally sensitive strategies tailored to younger and older women are recommended in future interventions.

XSL•FO
RenderX

Search number	Reference	Objective	Study design, sample population	Sample population	Mode of delivery	Specific instruments and interventions	Main findings
8	[32]	To ascertain the effects of breast cancer awareness programs on the use and uptake of breast screening among women in the United Arab Emirates	Cross-sectional study in the United Arab Emirates using a content-validated interviewer-administered questionnaire	959 women aged ≥20 years	University	—[c]	One-third of the participants (31.7%) had attended awareness programs, and attendance was influenced by younger age, higher education, and being single (P<.001); screening was performed by 38.3% of participants who had ever attended any breast cancer awareness programs than nonattendees (13.3%; P<.001); important barriers included knowledge gaps, cost concerns, and the lack of physician recommendations.
9	[33]	To evaluate the impact of a video-based intervention on the knowledge of breast cancer risk factors, symptoms, and attitudes toward screening practices	A quasi-experimental design in Bihar, India	267 rural-dwelling women	Regional health services	Intervention entailing serial exposure to validated YouTube Hindi videos focusing on educating respondents on breast cancer, changing attitudes toward screening methods, and encouraging healthy prevention practices	Knowledge level of risk factors, symptoms, and breast cancer screening increased significantly (100%) at 3 months post-intervention. Attitude towards breast self-examination improved significantly in 58% of women, 16.5% for clinical breast examination, and 9.7% for mammogram. Breast self-examination practices improved significantly in 60% of women. Associated factors of breast cancer awareness include education, age, and occupation.

XSL•FO
RenderX

Search number	Reference	Objective	Study design, sample population	Sample population	Mode of delivery	Specific instruments and interventions	Main findings
10	[34]	Exploring participatory social marketing principles for culturally adapted evidence-based cancer screening programs at the community level	Case studies in New York City, United States	Muslim women	Public health government bodies	The case entailed a culturally adapted program to increase breast and cervical cancer screening among a multiracial and ethnic population.	The case study revealed the benefits of integrating social marketing and CBPR[d] approaches as key strategies for the development of public health campaigns; they were effective in reaching and influencing health behaviors among vulnerable populations in the community, particularly the socially marginalized.
11	[35]	To explore the effects of peer trainer intervention on awareness of early detection and diagnosis, use of early diagnosis methods, and screening centers among women in Turkey	Quasi-experimental design in Turkey	40 women (>40 years old) participated as peer educators, with 20 selected as principal peer educators; 200 women received the education	Public health government bodies	Peer trainers educated their peers, followed by arranging a mammography appointment for women who were interested in taking the test	Participants demonstrated positive changes in health beliefs and breast self-examination knowledge; of the ~20% (n=1040) who did get mammograms, 8% were positive for breast cancer.
12	[36]	To assess the effects of a culturally based intervention program on adherence to breast cancer screening	Randomized controlled trial with a 3-month follow-up survey in Israel	589 ultra-Orthodox Jewish or Arab women	Public health government bodies	Individually tailored intervention based on each participant's degree of adherence, knowledge, and beliefs	In both groups, the performance of breast awareness practice and clinical breast examination was associated with a lower level of religious beliefs and a better understanding of the importance of screenings.

XSL•FO
RenderX

Search number	Reference	Objective	Study design, sample population	Sample population	Mode of delivery	Specific instruments and interventions	Main findings
13	[37]	Impact of mass social events on awareness of breast and cervical health among Black women	Case studies, with an average attendance time >1 hour at each event in the United Kingdom	>450 attendees at each event	Public health government bodies	Series of virtual educational and cancer screening events based on the motive to build a breast cancer resource hub for Black women (Breast/Best Health for Black Women [BHF-BW]); the program was a collaboration between the research institution and relevant community organizations; the intervention was designed to target the individual needs of Black women.	Most respondents (>85%) perceived that the event assisted them with feeling supported; about 46-48 women underwent a screening mammography during the events in 2022 and 2023; in both periods, >9% of participants affirmed they were likely to go for a mammogram when necessary.
14	[38]	To determine the influence of fear appeal messages available in media outlets on intention to perform breast self-examination among women in China	2-by-2 factorial experiment in China	488 women aged 25-50 years	Public health government bodies	Provision of messages about the risks and precautions of breast cancer across different media outlets and the intention to perform breast cancer self-examination using the EPPM[e]	Threat and efficacy had significant main and interaction effects on intention to perform breast cancer self-examination; no significant difference was detected in terms of the effectiveness of fear appeal messages between traditional and social media outlets.

XSL•FO
RenderX

Search number	Reference	Objective	Study design, sample population	Sample population	Mode of delivery	Specific instruments and interventions	Main findings
15	[39]	To promote screening mammography and correct breast cancer myths among women living in an urban community in Korea	A multicity, multicomponent community intervention trial, with surveys at baseline and 7 months later in South Korea	480 women	Public health government bodies	The intervention city was equipped with a community outreach and clinic and pharmacy-based in-reach strategies; the community campaign emphasized the efficacy of screening mammography for detecting masses that are not reachable; messages were tailored to correct myths about breast cancer and screening mammography.	A 20.4%-point decrease in the myths regarding the association between breast size and cancer, a 19%-point decrease in myths about the cost of mammography, and a 14% increase in the intention to perform screening mammography. In the intervention city, about a 23% increase in the proportion of participants at the action stage of the transtheoretical model, but such improvement was either lacking or inconsistent in the control city.

[a]PEP: Promoting Early Presentation.

[b]OR: odds ratio.

[c]Not available.

[d]CBPR: community-based participatory research.

[e]EPPM: extended parallel process model.

The sample size of participants in the studies ranged from 32 to 1945 with a total of 10,287 women enrolled in the intervention groups and 1164 in the control groups accross the 15 studies. The participants' ages were presented in 8 studies, with a minimum age of ≥20 years to a maximum age of 70 years old. Most studies used >40 years old as the eligibility criterion.

Characteristics and Methodological Approaches Used in Social Marketing Interventions

This section presents findings regarding the first objective, focusing on the characteristics of social marketing interventions in the context of breast cancer screening. Specifically, the results are described based on intervention source and types; underpinning theories; insights, exposure, and reach; and outcome measures.

Intervention Source and Types

Public health government bodies were the predominant intervention source (8/15; 53%), followed by regional health services (3/15, 20%), cancer charities (2/15, 13%) and a university (2/15, 13%). The interventions can also be broadly divided into national (n=10) and regional (n=5) breast cancer awareness campaigns, as well as targeted (n=9) and untargeted types (n=6). National breast cancer awareness campaigns refer to those performed to enhance breast cancer screening and awareness among the general public. Meanwhile, regional types focus on a specific region or community. Targeted interventions encompassed those that used well-defined activities to disseminate breast cancer screening and awareness information to specific groups, whereas untargeted types refer to interventions delivered to the general public without specifically focusing on breast cancer.

Overall, 4 types of social marketing interventions were identified: (1) video-based marketing (n=2) [18,33], (2) health professional and educational strategies (n=5) [14,15,27,28,32], (3) community outreach campaigns (n=3) [34,36,39], and (4) social media or mass media campaigns (n=5) [30,31,37,38,40]. Thus, the majority of studies involved social marketing strategies incorporating health professional and educational approaches, followed by social media and outreach campaigns.

Behavior Change Theories Used in Interventions

Five studies provided insight into the theories that informed the development and implementation of the intervention. The predominant models considered were the health belief model [30,35,36], the transtheoretical model [39], and the extended parallel process model of persuasion [38].

All interventions were designed to influence breast cancer screening behaviors, such as improving attendance at cancer

XSL•FO
RenderX

screening, performing self-examination, and adhering to recommended prevention practices. This was achieved by disseminating information on breast cancer screening and raising awareness through advertisements, educational interventions, and promotional or fundraising activities.

The interventions shaped behavior change by informing participants or vulnerable people regarding the health consequences of breast cancer, specifically by providing videos, posters, advertisements, brochures, modules, and links to cancer charity and public health websites to access information about screening, risk factors, and symptoms. The interventions delivered information from well-established and credible sources, such as cancer charities, national and regional government health organizations, and university or institutional bodies, but certain information from untrusted individuals, users, or celebrities was deemed less credible [18]. Interventions that used social comparisons, individuals' approval, and social support to motivate cancer screening were reported in 4 studies. Examples included sharing messages of support from organizations, the public, cancer survivors, charities, and celebrities [18,31,35,36,39].

Insights, Exposure, and Reach

Different exposures were used in line with the type of intervention. Some of the studies provided detailed information on the exposure and the reach achieved by the social marketing interventions. One of the interventions with the largest reach involved 2 waves of a social campaign performed over 2 weeks with 1200 participants in the northeast of England. The target online audience was also reached via advertisements shown on promoted posts in Facebook [18]. A similarly large audience was reached by studies conducted in the United Arab Emirates [32], reaching 959 women aged ≥20 years; Turkey [35] with 20 trained peer educators and each educating approximately 200 women; and more than 450 attendees in a series of campaign events organized to build a breast cancer resource hub for Black women in the United Kingdom [37]. In Malaysia, about two-thirds of the participants identified at least one of the campaign materials used for the intervention, and the probability of being reached during the campaign was lower among Chinese women (odds ratio [OR] 0.25, 95% CI 0.15 to 0.40) and those aged >70 years (OR 0.47, 95% CI 0.23 to 0.94) relative to Malays and younger women [30].

A few studies shared insights into participants' views on the acceptance of social marketing strategies as an approach to delivering breast cancer awareness and screening information [18,37]. However, 1 study revealed that some women perceived that some breast cancer screening promotional messages, including award-winning advertisements, were offensive and did not represent the benefits or messages portrayed in the advertisements [29].

Outcome Measures

Descriptively, 5 studies reported outcomes relating to breast cancer awareness [18,27,28,30,40]; 6 studies reported outcomes for early cancer screening and preventive behaviors [27,31,35,38,39,40]; 4 studies reported outcomes related to perceptions regarding social marketing strategies for prompt cancer screening [29,30,34,37]; 6 studies reported outcomes regarding knowledge, attitudes, and practices toward breast cancer screening [14,32,33,35,36,39]; and 4 studies reported on the predictors of breast cancer awareness and screening [14].

Impacts of Social Marketing Interventions

Breast Cancer Awareness

Martin et al [18] used TV advertisements to explore the relationship between breast cancer and alcohol consumption, supported with voice notes and Facebook posts on different types of cancer caused by consuming alcohol. The intervention led to a significant postcampaign increase (45% increase) in breast cancer awareness compared with the precampaign level. In Malaysia and Romania, culturally adapted and mass media campaigns (TV, radio, print media, and social media) led to significantly higher levels of awareness among participants who recognized the campaign than nonrecognizers [30]. However, the true effects of these social marketing interventions on breast cancer awareness remain uncertain given the methodological limitations, particularly the use of a 1-group pre and postdesign and the lack of underpinning theories.

Meanwhile, more reliable findings were observed in 2 RCTs [27,28]. A multicomponent intervention, comprising Promoting Early Presentation (PEP) and a 10-minute interaction with a radiographer, improved breast cancer awareness by 2-fold (OR 8.1, 95% CI 2.7% - 25.0%) compared with participants in usual care at the 2-year follow-up [27]. A similar intervention based on the English National Health Service Breast Screening Program for older women (approximately 70 years old) was associated with a 7-fold increase in breast cancer awareness (OR 15.2, 95% CI 10.0 to 23.2) relative to a 2-fold increase in the control group. These results suggest that a multicomponent intervention incorporating SMCs and health promotional activities is more effective at upscaling breast cancer awareness. However, clinical trials comparing these various approaches are required to test the hypothesis.

Breast Cancer Screening and Preventive Behaviors

The significant effects of SMCs on early breast cancer screening and preventive behaviors were revealed in 3 intervention trials [27,38,39]. The PEP intervention investigated by Forbes et al [27] among vulnerable women in the United Kingdom led to an 8-fold increase in participants' readiness to perform early screening. The intention to perform breast cancer self-examination also improved following the provision of messages about the risks and precautions of breast cancer across different media outlets [38]. Likewise, community outreach and clinic and pharmacy-based in-reach strategies led to a 14% and 23% increase, respectively, in the intention to perform screening mammography and the proportion of participants at the action stage [39].

A few studies were designed to shape participants' perceptions and positive behavioral change toward breast cancer screening via community outreach campaigns [34], social media or mass media campaigns [29,37], and health professional and educational strategies [14,32]. Yfantidou and Skandali [29] conducted a qualitative interventional study by presenting breast cancer screening advertisements to 13 participants, followed

XSL•FO
RenderX

by interviews to capture their perspectives on the promotional activities. Some of the breast cancer screening and promotional activities were considered offensive and failed to connect with the visual designs; however, the participants were motivated to perform screening when exposed to positive advertisements capturing reliable data on breast cancer and endorsements from highly respected celebrities. Williams et al [37] used a similar intervention involving a series of virtual educational and cancer screening events used to build a breast cancer resource hub for Black women. As a result, most participants (>85%) perceived the SMC to be very helpful and supportive, leading to a 9% increase in the proportion of those likely to visit health care centers for mammogram screening.

Some of the studies in this category, particularly those involving health professional and educational strategies [32,35,37] and community outreach campaigns [36,39], highlighted a positive relationship between improved knowledge and a positive perception toward breast cancer screening. The video-based social marketing used by Sinha and Sharma [33] in a regional study in India also reported significant improvements in participants' knowledge of risk factors, symptoms, and breast cancer screening, leading to positive attitudes toward breast self-examination. Direct comparisons between these interventions is challenging given the different approaches. More importantly, the evidence from these studies is weak given the qualitative or quasi-experimental design and the use of an online survey for follow-up assessment [35].

Factors Influencing the Primary Outcomes

The factors influencing breast cancer screening awareness were investigated by 4 articles. Alomairy et al [14] found that cultural backgrounds shaped younger and older women's attitudes toward breast cancer screening services. Abdulla et al [32] and Sinha and Sharma [33] reported that younger age, higher education, occupation, and being single improved the likelihood of attending breast cancer awareness programs, with knowledge gaps, cost concerns, and the lack of physician recommendations constituting the major barriers. The performance of breast awareness practices and clinical breast examinations was associated with having a lower level of religious beliefs and better understanding of the importance of screenings [36]. In contrast, none of the demographic variables influenced improvement in symptom awareness scores among Malaysian women who were exposed to a culturally adapted and mass media campaign (TV, radio, print media, and social media) [30]. These findings reflect the multifactorial nature of breast cancer screening behaviors and the need for tailored social marketing interventions.

Discussion

Principal Findings

Most articles on social marketing interventions assessed breast cancer awareness and screening intentions using a mixture of experimental and observational study designs to measure exposure, reach, and impacts on investigated outcomes. A few studies revealed that national and regional breast cancer awareness campaigns and targeted interventions might improve cancer awareness, screening intentions, and uptake. The available evidence also suggests that exposure, reach, and engagement attained by these interventions differed by age, gender, educational level, and users' ethnicity.

Although several reviews have been published on interventions to address cancer screening and prompt diagnosis, to the best of our knowledge, this scoping review is the first attempt to focus on social marketing interventions to improve breast cancer screening, awareness, early diagnosis, and presentation to care. The findings add to the existing knowledge pool by reflecting the data paucity and robust clinical trials exploring the level of engagement, mechanisms of behavior change, and attendance at screenings specifically designed for women with breast cancer and high-risk groups. It is difficult to attain high engagement levels given the challenge with accessing timely observational data on breast cancer screening attendance [40,41]. This is further complicated by the fact that social marketing interventions are usually designed without provision for evaluation and impact metrics [42]. A similar situation was observed in the scoping review conducted by Plackett et al [5], whereby the authors commented, based on expert validation of emerging themes, that campaigns are designed rapidly without room for evaluation.

Immediate changes are difficult to measure since cancer screening and early diagnosis are long-term goals, which is in sharp contrast to the primary design of SMCs. More comprehensive evaluations that reflect outcomes of behavior change and engagement could be achieved by developing an evaluation framework that integrates components of behavior change theories [43]. Nevertheless, it is difficult to equivocally demonstrate that a specific campaign is primarily responsible for a change in behavior or outcome, in addition to the organizations' limited resources and time to perform evaluations. Such assessments could, however, be conducted in campaigns involving collaboration with public health and institutional bodies.

In this review, we attempted to use the BCT taxonomy to delineate the various BCTs used in social marketing interventions to shape participants' health behaviors, including social support, incentives, and equipping vulnerable populations with health information. The health belief theory was the framework predominantly used to develop the investigated interventions. According to BCTs, providing actionable health information about breast cancer symptoms and risk factors could have stronger and direct positive effects on behavior change. The BCTs used in breast cancer campaigns, such as sharing personal stories, could help create social support and influence public opinion, thereby encouraging help-seeking behaviors, increasing screening uptake, and improving health outcomes. Future interventions could explore these assumptions, particularly how national and regional social marketing strategies might facilitate positive behavior change and the most important messages required in such interventions.

This scoping review revealed the degree to which social marketing interventions addressed inequalities in breast cancer screening, awareness, and early diagnosis. The available evidence suggests a low reach and engagement with ethnic minorities [1,30], which is mainly associated with limited access

XSL•FO
RenderX

to breast cancer screening, stigma, and low engagement with social media and SMCs [44,45]. This inequality is further exacerbated when the information is not culturally adapted [46]. The low uptake of screening and awareness may also stem from the lack of campaign role models that are highly respected and resonate with ethnic minorities [18].

Inequalities in the use and impacts of SMCs were also identified in terms of age, educational level, marital status, and occupation. Younger age, higher education, being employed, and being single improved the odds of exposure to breast cancer awareness programs and perceived benefits toward screening services [14]. These results align with recent data on social marketing and media use in the United Kingdom and high-income countries by age and gender [47]. The finding underscores the need for specific campaign strategies preferred by target users and those more likely to facilitate broad engagement.

Future interventions involving social marketing approaches should assess inequalities in exposure, reach, and engagement, as well as consider the effectiveness directly in relation to the target groups. Designing interventions toward populations with a disproportionate disease burden could address health inequalities observed in breast cancer screening and early diagnosis. This could be achieved by shifting from broad, "one-size-fits-all" approaches to audience-segmented interventions that effectively tackle systemic barriers. Social marketing strategies involving mass and social media can shift the "product" from fear-inducing to capacity-building, relatable, and manageable actions [48]. Educational-related barriers could be addressed by incorporating language and cultural nuances, partnering with community leaders, and designing messages that align with cultural beliefs and mitigating local cancer myths [49].

Meanwhile, inequities introduced by unemployment could be tackling systemic and hidden costs. Examples include campaigns that promote screenings related to tangible solutions, such as free and accessible mammography centers in local communities. These marginalized groups could also be further reached by expanding distribution channels beyond digital spaces, focusing on physical and trusted touchpoints [48]. Other possible channels to address the aforementioned inequities include leveraging trusted intermediaries to deliver face-to-face messaging and engaging the support ecosystem while considering the role of family structures in health care decisions [5].

Based on the syntheses of the 15 articles, most SMCs and interventions had significant impacts on breast cancer screening awareness and preventive behaviors. Overall, the findings are consistent with previous reviews on the use of social marketing strategies in HIV testing and related outcomes [49], whereby a consistent pattern of positive results were observed. Nevertheless, effectiveness seems to differ depending on specific elements used in the intervention, such as mass media, social media, and recognizable figures. Methodological limitations also need to be considered when interpreting the findings.

For instance, 2 RCTs [27,28] incorporating SMCs and health promotional activities led to significant improvement in upscaling breast cancer awareness, with a large effect size. These promising results may stem from the efficacy of such campaigns at enabling the transition from mere information to behavioral change. By integrating health promotional activities such as culturally tailored education, the intervention directly addresses psychosocial barriers, strengthens self-efficacy, and encourages proactive health actions [50]. Moreover, a comprehensive marketing mix is established by collaborating with stakeholders to tackle barriers such as accessibility and cost [14].

Social media or mass media and community outreach campaigns [39] were also instrumental at enhancing breast cancer awareness. These campaigns combined broad reach and localized, culturally tailored messaging as well as mitigated specific barriers to care such as stigma and fear. As a result, women were empowered to identify symptoms early and encouraged to visit screening centers [30,39]. However, the actual effect size of changes in breast cancer awareness could not be ascertained given the low-quality research designs. Promising results were also reported in a few qualitative interventional studies involving social media or mass media campaigns [29,37]. Positive advertisements capturing reliable data on breast cancer and endorsements from highly respected celebrities were perceived more positively, leading to proactive health actions. Celebrities can upscale breast cancer awareness and promote positive attitudes toward screening by leveraging their massive reach to destigmatize the cancer, break obstacles, and act as crucial relatable models for early diagnosis [51]. Involving recognizable figures bridge gaps between the general public and complex medical information, translating awareness into life-saving actions [52].

Video-based marketing was primarily used as the source of intervention in 2 studies [18,33], with both reporting greater awareness and improved knowledge and attitudes toward breast cancer screening postintervention. This approach was effective at bypassing literacy barriers, capturing attention, and directly inspiring proactive health behaviors because they rapidly translate complex health information into accessible, emotionally resonant and culturally adaptable formats [53].

Nevertheless, the evidence from the studies needs a deeper assessment and to be interpreted cautiously. Direct comparisons between the interventions was challenging given the different study designs and interventional approaches. However, the large effect sizes reported for SMCs incorporating health promotional activities suggest that these elements are effective at upscaling breast cancer awareness and positive screening decisions. This is further supported by the robust research designs used in the reviewed studies [14]. Community outreach and mass media campaigns also showed promising results in terms of breast cancer awareness and preventive behaviors, with the inclusion of recognizable figures providing additional benefits [54]. Meanwhile, evidence from video-based marketing approaches is weak, and the incorporation of other proven marketing elements may be required to optimize their effects.

Overall, the reviewed studies depict the importance of identifying SMCs and delivery methods that strongly resonate with the target audience [55]. Significant efforts are required to achieve audience segmentation and formative research with the target market for planning and branding the campaign. A

XSL•FO
RenderX

catchy slogan and name are critical components of effective SMCs, but the information was lacking in most studies. As a result, it is unclear whether these important elements contributed to the promising outcomes reported in the reviewed studies.

Quality appraisal also revealed that only 2 clinical trials (RCTs) had a robust design. Based on the weak study designs, it is difficult to conclusively infer that the SMCs led to improved breast cancer awareness, screening uptake, and early diagnosis or presentation into care. Although the use of different approaches and evaluation methods provided a clear picture of the population's exposure to the SMC and the participants' awareness levels, a standardized approach is required to facilitate robust comparisons between campaigns in terms of effectiveness and economic benefits. We acknowledge the diverse nature of target populations when conducting SMCs, but a standardized reporting of exposure and impact is crucial for significant progress to be achieved in this unique field of public health.

Limitations

Research into social marketing in public health interventions is rapidly growing, and this scoping review presents only the current and available evidence for social marketing interventions. Only articles written in English were included in our review, thereby neglecting information from relevant interventions either conducted in non-English speaking countries or not written in English. This limitation affects the generalizability of the review findings. Many of the SMCs used combined interventional approaches and sources, possibly contributing to the positive outcomes in terms of breast cancer awareness and screening participation.

Given the lack of a standardized protocol for report evaluations of social marketing interventions, the data synthesis was based on the outcomes measured and reported in the reviewed studies. In other words, some degree of reporting bias might be present in the included studies. The exclusion of gray literature may also introduce some bias, as only articles retrieved from selected databases and peer-reviewed journals were analyzed. The quality appraisal highlighted moderate to weak study designs in 13 of the 15 articles. We did not perform a meta-analysis due to the heterogeneity between the studies, particularly in terms of study designs, delivery methods, type of social marketing interventions, and diverse outcomes. Developing a reporting protocol based on available frameworks for evaluating social marketing research should be considered in future research endeavors.

Conclusion

This scoping review revealed that behavior change outcomes were sparingly assessed in most evaluations of social marketing interventions to improve breast cancer awareness, screening, and early diagnosis. The limited evidence reflects various types of social marketing interventions, ranging from community outreach to health professional and educational strategies, social media or mass media campaigns, and video-based marketing, which may improve breast cancer awareness and intended and actual screening uptake. Robust evaluations of these interventions could be performed in future research by using well-established evaluation frameworks and reporting guidelines. This will assist with capturing behavior change outcomes and elucidating their underlying mechanisms. Specific users or engaged individuals should be identified, and measures should be used to ascertain whether some health inequalities could be addressed by social marketing interventions. Developers and funders of these campaigns should also consider using an evaluation plan before conducting the actual campaign. Such evaluation plans may consider using feasible mixed methods and quasi-experimental designs. This is crucial to ensuring that campaign exposure and impact are measured adequately, as well as identifying the specific elements responsible for the observed behavioral change.

Acknowledgments

The authors extend their appreciation to Umm Al-Qura University, Saudi Arabia, for funding this research work through grant number 26UQU4350354GSSR01.

No generative artificial intelligence (AI) tool was used in this study.

Funding

This research work was funded by Umm Al-Qura University, Saudi Arabia, under grant number 26UQU4350354GSSR01.

Conflicts of Interest

None declared.

Checklist 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.
[PDF File, 180 KB - cancer_v12i1e88568_app1.pdf]

References

XSL•FO
RenderX

1. Amarasinghe H, Warnakulasuriya S, Johnson NW. Evaluation of a social marketing campaign for the early detection of oral potentially malignant disorders and oral cancer: Sri Lankan experience. J Oral Biol Craniofac Res 2021;11(2):204-208. [doi: 10.1016/j.jobcr.2021.01.013] [Medline: 33665068]
2. Evans WD. How social marketing works in health care. BMJ 2006 May 20;332(7551):1207-1210. [doi: 10.1136/bmj.332.7551.1207-a] [Medline: 16710002]
3. Thompson JK, Heinberg LJ. The media's influence on body image disturbance and eating disorders: we've reviled them, now can we rehabilitate them? Journal of Social Issues 1999 Jan;55(2):339-353 [FREE Full text] [doi: 10.1111/0022-4537.00119]
4. Anne Garretson Folse J, Guidry Moulard J, Raggio RD. Psychological ownership: a social marketing advertising message appeal? International Journal of Advertising 2012 Jan;31(2):291-315. [doi: 10.2501/IJA-31-2-291-315]
5. Plackett R, Kaushal A, Kassianos AP, et al. Use of social media to promote cancer screening and early diagnosis: scoping review. J Med Internet Res 2020 Nov 9;22(11):e21582. [doi: 10.2196/21582] [Medline: 33164907]
6. Ogie RI, James S, Moore A, Dilworth T, Amirghasemi M, Whittaker J. Social media use in disaster recovery: a systematic literature review. International Journal of Disaster Risk Reduction 2022 Feb;70:102783. [doi: 10.1016/j.ijdrr.2022.102783]
7. Harrison TR, Morgan SE, Chewning LV. The challenges of social marketing of organ donation: news and entertainment coverage of donation and transplantation. Health Mark Q 2008;25(1-2):33-65. [doi: 10.1080/07359680802126079] [Medline: 18935879]
8. Shrum LJ, Lowrey TM, McCarty JA. Recycling as a marketing problem: a framework for strategy development. Psychology and Marketing 1994 Jul;11(4):393-416 [FREE Full text] [doi: 10.1002/mar.4220110407]
9. Jafar Z, Quick JD, Larson HJ, et al. Social media for public health: reaping the benefits, mitigating the harms. Health Promot Perspect 2023;13(2):105-112. [doi: 10.34172/hpp.2023.13] [Medline: 37600540]
10. Olawepo JO, Pharr JR, Kachen A. The use of social marketing campaigns to increase HIV testing uptake: a systematic review. AIDS Care 2019 Feb;31(2):153-162. [doi: 10.1080/09540121.2018.1533631] [Medline: 30304940]
11. Roberts-Lewis SF, Baxter HA, Mein G, et al. The use of social media for dissemination of research evidence to health and social care practitioners: protocol for a systematic review. JMIR Res Protoc 2023 May 12;12:e45684. [doi: 10.2196/45684] [Medline: 37171840]
12. Evans WD, McCormack L. Applying social marketing in health care: communicating evidence to change consumer behavior. Med Decis Making 2008;28(5):781-792. [doi: 10.1177/0272989X08318464] [Medline: 18556638]
13. Ponce-Chazarri L, Ponce-Blandón JA, Immordino P, Giordano A, Morales F. Barriers to breast cancer-screening adherence in vulnerable populations. Cancers (Basel) 2023 Jan 18;15(3):604. [doi: 10.3390/cancers15030604] [Medline: 36765561]
14. Alomairy NA. The impact of breast cancer awareness campaigns on the knowledge, attitudes, and practices of breast cancer screening among Saudi female employees. PLOS ONE 2025;20(9):e0331765. [doi: 10.1371/journal.pone.0331765] [Medline: 40911590]
15. Bala MM, Strzeszynski L, Topor-Madry R. Mass media interventions for smoking cessation in adults. Cochrane Database Syst Rev ;2017(11). [doi: 10.1002/14651858.CD004704.pub4]
16. Jones DS, Greene JA. The decline and rise of coronary heart disease: understanding public health catastrophism. Am J Public Health 2013 Jul;103(7):1207-1218. [doi: 10.2105/AJPH.2013.301226] [Medline: 23678895]
17. Cohen SA, Cohen LE, Tijerina JD. The impact of monthly campaigns and other high-profile media coverage on public interest in 13 malignancies: a Google Trends analysis. Ecancermedicalscience 2020;14:1154. [doi: 10.3332/ecancer.2020.1154] [Medline: 33574899]
18. Martin N, Buykx P, Shevills C, Sullivan C, Clark L, Newbury-Birch D. Population level effects of a mass media alcohol and breast cancer campaign: a cross-sectional pre-intervention and post-intervention evaluation. Alcohol Alcohol 2018 Jan 1;53(1):31-38. [doi: 10.1093/alcalc/agx071] [Medline: 29155922]
19. Pecchioni LL, Sparks L. Health information sources of individuals with cancer and their family members. Health Commun 2007;21(2):143-151. [doi: 10.1080/10410230701307162] [Medline: 17523860]
20. Ruco A, Dossa F, Tinmouth J, Llovet D, Kishibe T, Baxter NN. Social media and mobile health technology for cancer screening: a systematic review and meta-analysis protocol. BMJ Open 2020 Feb 5;10(2):e035411. [doi: 10.1136/bmjopen-2019-035411] [Medline: 32029500]
21. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
22. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
23. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libraries J 2009 Jun;26(2):91-108 [FREE Full text] [doi: 10.1111/j.1471-1842.2009.00848.x]
24. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series: Wiley; 2008. URL: https://onlinelibrary.wiley.com/doi/book/10.1002/9780470712184 [doi: 10.1002/9780470712184]
25. Neiger BL, Thackeray R, Van Wagenen SA, et al. Use of social media in health promotion. Health Promot Pract 2012 Mar;13(2):159-164. [doi: 10.1177/1524839911433467]

XSL•FO
RenderX

26. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
27. Forbes LJL, Forster AS, Dodd RH, et al. Promoting early presentation of breast cancer in older women: implementing an evidence-based intervention in routine clinical practice. J Cancer Epidemiol 2012;2012:835167. [doi: 10.1155/2012/835167] [Medline: 23213334]
28. Forbes LJL, Linsell L, Atkins L, et al. A promoting early presentation intervention increases breast cancer awareness in older women after 2 years: a randomised controlled trial. Br J Cancer 2011 Jun 28;105(1):18-21. [doi: 10.1038/bjc.2011.205] [Medline: 21654683]
29. Yfantidou I, Skandali D. Resonating messages: a mixed-methods investigation of breast cancer screening advertisements. CCIJ 2025 Jun 17;30(4):696-723. [doi: 10.1108/CCIJ-09-2024-0172]
30. Schliemann D, Htay MNN, Dahlui M, et al. Impact of a mass media campaign on breast cancer symptoms awareness and screening uptake in Malaysia: findings from a quasi-experimental study. BMJ Open 2020 Aug 20;10(8):e036503. [doi: 10.1136/bmjopen-2019-036503] [Medline: 32819988]
31. Todor RD, Brătucu G, Candrea AN, Strempel CG, Anastasiu CV. Social media campaigns: a game changer for the prevention of breast cancer in Romania. Health Care (Don Mills) ;12(8):865. [doi: 10.3390/healthcare12080865]
32. Abdulla RA, Kareem NA, Assadi RA, et al. Impact of breast cancer awareness program on breast screening utilization among women in the United Arab Emirates: a cross-sectional study. BMC Public Health 2025 Feb 12;25(1):578. [doi: 10.1186/s12889-025-21512-1] [Medline: 39939986]
33. Sinha N, Sharma A. Digital media intervention for breast cancer awareness among rural women: a quasi-experimental study from Bihar, India. Clin Epidemiol Glob Health 2024 Jul;28:101705. [doi: 10.1016/j.cegh.2024.101705]
34. Kwon SC, Kranick JA, Islam NS, et al. Applying social marketing principles for community-based cancer screening programs: two case studies. Health Educ Behav 2025 Aug;52(4):382-391. [doi: 10.1177/10901981251322806]
35. Gözüm S, Karayurt O, Kav S, Platin N. Effectiveness of peer education for breast cancer screening and health beliefs in eastern Turkey. Cancer Nurs 2010;33(3):213-220. [doi: 10.1097/NCC.0b013e3181cb40a8] [Medline: 20357655]
36. Freund A, Cohen M, Azaiza F. A culturally tailored intervention for promoting breast cancer screening among women from faith-based communities in Israel: a randomized controlled study. Res Soc Work Pract 2019 May;29(4):375-388. [doi: 10.1177/1049731517741197]
37. Williams C, Goulbourne E, Gyansa E, et al. Initiatives to increase breast and cervical cancer-related knowledge, screening, and health behaviours among Black women. Can J Public Health 2025 Feb;116(1):100-108. [doi: 10.17269/s41997-024-00953-y] [Medline: 39436534]
38. Chen L, Yang X. Using EPPM to evaluate the effectiveness of fear appeal messages across different media outlets to increase the intention of breast self-examination among Chinese women. Health Commun 2019 Oct;34(11):1369-1376. [doi: 10.1080/10410236.2018.1493416] [Medline: 30080982]
39. Park K, Hong WH, Kye SY, Jung E, Kim MH, Park HG. Community-based intervention to promote breast cancer awareness and screening: the Korean experience. BMC Public Health 2011 Jun 14;11:468. [doi: 10.1186/1471-2458-11-468] [Medline: 21669004]
40. Hutton G, Li SJ, Quaife SL, et al. Understanding barriers to breast screening: an online survey of non-attenders as part of a service evaluation in the breast screening programme in England. BMC Public Health 2025 Jul 19;25(1):2509. [doi: 10.1186/s12889-025-23691-3] [Medline: 40684139]
41. Pagoto S, Waring ME, Xu R. A call for a public health agenda for social media research. J Med Internet Res 2019 Dec 19;21(12):e16661. [doi: 10.2196/16661] [Medline: 31855185]
42. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
43. Dunn AG, Mandl KD, Coiera E. Social media interventions for precision public health: promises and risks. NPJ Digit Med 2018;1(1):47. [doi: 10.1038/s41746-018-0054-0] [Medline: 30854472]
44. Vrinten C, Gallagher A, Waller J, Marlow LAV. Cancer stigma and cancer screening attendance: a population based survey in England. BMC Cancer 2019 Jun 11;19(1):566. [doi: 10.1186/s12885-019-5787-x] [Medline: 31185949]
45. Marlow LAV, Waller J, Wardle J. Does lung cancer attract greater stigma than other cancer types? Lung Cancer 2015 Apr;88(1):104-107. [doi: 10.1016/j.lungcan.2015.01.024] [Medline: 25704958]
46. Viswanath K, Emmons KM. Message effects and social determinants of health: its application to cancer disparities. J Commun 2006 Aug 1;56(suppl_1):S238-S264. [doi: 10.1111/j.1460-2466.2006.00292.x]
47. Adults: media use and attitudes report 2019. Ofcom. 2019 May 30. URL: https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/adults/media-use-and-attitudes-2019/adults-media-use-and-attitudes-report.pdf?v=323935 [accessed 2026-08-16]
48. Kilibarda T, Sinanović Š, Bubanj S, Trgovčević S, Ivanović S, Milutinović S. Health promotion and the components of social marketing in the function of early diagnosis of breast cancer. Journal of Regional Section of Serbian Medical Association in Zajecar 2021;46:4 [FREE Full text]

XSL•FO
RenderX

49. Wei C, Herrick A, Raymond HF, Anglemyer A, Gerbase A, Noar SM. Social marketing interventions to increase HIV/STI testing uptake among men who have sex with men and male-to-female transgender women. Cochrane Database Syst Rev 2011 Sep 7;2011(9):CD009337. [doi: 10.1002/14651858.CD009337] [Medline: 21901734]
50. Larsen L. "I think it is a powerful campaign and does a great job of raising awareness in young women": findings from breast cancer awareness campaigns targeting young women in Canada. Can Oncol Nurs J 2022;32(1):61-67. [doi: 10.5737/236880763216167] [Medline: 35280071]
51. Evans DGR, Barwell J, Eccles DM, et al. The Angelina Jolie effect: how high celebrity profile can have a major impact on provision of cancer related services. Breast Cancer Res ;16(5). [doi: 10.1186/s13058-014-0442-6]
52. Majali TAO, Omar A, Alsoud M, Barakat S. Use of social media in promoting breast cancer awareness among Malaysian women of generation Y: a conceptual framework. Multicultural Education 2021 [FREE Full text] [doi: 10.5281/zenodo.4495161]
53. Muthineni RR, Suhani S, Kaur R, et al. Effectiveness of video health education on breast cancer awareness and self-examination in the new age of digitalisation: community-based evidence from a developing nation. Eur J Breast Health 2026 Mar 24;22(2):147-155. [doi: 10.4274/ejbh.galenos.2025.2025-9-4] [Medline: 41873842]
54. Xu S, Markson C, Costello KL, Xing CY, Demissie K, Llanos AA. Leveraging social media to promote public health knowledge: example of cancer awareness via Twitter. JMIR Public Health Surveill 2016;2(1):e17. [doi: 10.2196/publichealth.5205] [Medline: 27227152]
55. Noar SM, Palmgreen P, Chabot M, Dobransky N, Zimmerman RS. A 10-year systematic review of HIV/AIDS mass communication campaigns: have we made progress? J Health Commun 2009;14(1):15-42. [doi: 10.1080/10810730802592239] [Medline: 19180369]

Abbreviations

BCT: behavior change technique
mHealth: mobile health
OR: odds ratio
PEP: Promoting Early Presentation
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
RCT: randomized controlled trial
SMC: social marketing campaign

Edited by M Balcarras; submitted 27.Nov.2025; peer-reviewed by J Brabin, S Beasley; revised version received 12.Jul.2026; accepted 22.Jul.2026; published 05.Oct.2026.

Please cite as:
Abduljawad S
The Impacts of Social Marketing Campaigns on Breast Cancer Awareness and Screening: Scoping Review
JMIR Cancer 2026;12:e88568
URL: https://cancer.jmir.org/2026/1/e88568
doi:10.2196/88568

© Sahar Abduljawad. Originally published in JMIR Cancer (https://cancer.jmir.org), 5.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Longevity of Mobile Apps for Cancer Recovery: Scoping Review

Kenneth Färnqvist[1*], MSc; Luisa Thiele[1,2*], MSc; Sophie Johnsson[1], MSc; Pernilla Lagergren[1,3], PhD; Anna Schandl[1,4], PhD

[1]Department of Molecular Medicine and Surgery, Karolinska Institute, Blombäcksväg 23, Stockholm, Sweden

[2]Department of Social Sciences, Technology and Arts, Luleå University of Technology, Luleå, Sweden

[3]Department of Surgery & Cancer, Imperial College London, London, United Kingdom

[4]Division of Perioperative and Intensive Care, Department of Clinical Sciences and Education, Södersjukhuset | Karolinska Institutet, Stockholm, Sweden

[*]these authors contributed equally

Corresponding Author:
Anna Schandl, PhD
Department of Molecular Medicine and Surgery, Karolinska Institute, Blombäcksväg 23, Stockholm, Sweden

Abstract

Background: The number of cancer survivors is steadily increasing worldwide, leading to an increased demand for long-term follow-up and supportive care. Many survivors face ongoing physical and psychosocial issues that highlight the need for innovative management approaches. Mobile health apps offer potential benefits by facilitating patient-led follow-up, self-management, and more efficient use of health care resources. Although the market for cancer-related mobile apps has grown rapidly, their sustainability and scientific basis remain unclear. In the European Union, the Medical Device Regulation (MDR), in effect since May 2021, introduced stricter criteria for classifying medical devices, including certain software apps. While aiming to improve patient safety, the MDR could pose challenges for small companies and academic developers, potentially limiting the availability of such apps. No scoping review has delineated changes in active apps before and after implementation of the legislation regulating medical devices.

Objective: This scoping review aimed to evaluate the current availability and longevity of English-language mobile apps supporting cancer recovery, with a specific focus on changes before and after the implementation of the European Union MDR, and to assess the extent to which these apps are supported by clinical evidence.

Methods: Searches were conducted in mobile app stores (Apple's App Store and Google Play) and literature databases (MEDLINE, Embase, Cochrane Library, and Web of Science), using predefined terms. Mobile apps targeting cancer recovery and published articles on their effectiveness were included. Two reviewers independently extracted data. A descriptive analysis was conducted to report trends in mobile device app availability and updates over time.

Results: A total of 151 mobile apps were identified in 2018. However, by 2024, only 45 of 151 (30%) were still available. Among these, 25 of 151 (17%) were updated within the past 2 years. During the search in December 2024, 1 new mobile app supported by scientific evidence was discovered. This mobile app was developed to assist cancer survivors in managing insomnia through cognitive behavioral therapy. Rapid turnover and a potential lack of sustainability in the mobile health app market for cancer survivors were evident, with most mobile apps identified in 2018 no longer available by 2024.

Conclusions: This review revealed a limited number of publicly available mobile apps that support cancer recovery. The longevity of existing mobile apps is limited, potentially because of regulatory and financial barriers. Prioritizing rigorous effectiveness trials, addressing implementation barriers, and developing sustainable business models are essential to ensure the long-term availability and success of mobile health apps in cancer survivorship care.

(JMIR Cancer 2026;12:e82448) doi:10.2196/82448

KEYWORDS

cancer; symptoms; neoplasm; mobile health; cancer survivorship

Introduction

Globally, the total number of new instances of cancer is steadily rising, but treatment and survival rates are also consistently improving. Thus, the prevalence of people having undergone cancer treatments is rapidly increasing [1]. Currently, there are approximately 12 million cancer survivors in Europe [1]. Survivors often have regular health care follow-up appointments

XSL•FO
RenderX

for many years, or even lifelong [2]. Hence, providing adequate cancer care is a significant challenge because of increasing demand [3]. The role of regular outpatient appointments after treatment completion has been criticized because of the lack of scientific evidence supporting their benefits and the anxiety and distress they could bring patients [4]. In contrast, some data indicate that patient-initiated follow-up appointments may be more valuable for patients and the health care system [5]. Structured clinical follow-up remains an essential component of survivorship care, with digital tools increasingly regarded as complementary rather than substitutive approaches. This complementary role highlights the need for scalable, accessible, and resource-efficient follow-up models, including mobile health apps that can enhance access to and reduce the burden on health care systems and foster patient empowerment [5,6].

While structured follow-up remains important, many patients who have completed curative-intent treatment for cancer experience incomplete recovery, with persistent physical and psychosocial problems [7]. These facts highlight the need to consider novel management strategies for cancer survivors. Use of modern communication tools, such as smartphones and engaging in patient-driven follow-up, may offer a successful way forward to improve cancer survivorship and promote more cost-effective use of health care resources [5,6]. In recent years, the market for mobile apps has grown rapidly [8], and many initiatives focusing on cancer patients have become available [9]. Nevertheless, the annual number of mobile apps created to assist individuals living with and beyond cancer does not correspond to the existing evidence of their usefulness and effectiveness [10,11]. The rapid pace of app development has outstripped the evidence base, creating a gap that researchers are now trying to address and highlight as an evaluation crisis in the field [12].

The content of mobile apps for cancer management available on the market can be divided into five main categories: (1) communicating cancer information, (2) planning and organizing cancer care, (3) interacting with others (including health care professionals and others affected by cancer), (4) enacting management strategies and adjusting to life with or beyond cancer, and (5) receiving feedback about cancer management, for example, by sharing self-monitoring reports with health care professionals [13]. For instance, almost 40 apps have been proposed to aid skin cancer detection [14]. Some apps aim to help cancer patients keep track of their treatment and therefore include functions such as schedules and treatment-related information, as well as symptom tracking. The majority of these apps have not been scientifically developed or classified as medical devices [11].

Numerous cancer mobile apps are available on the market; however, the current availability of these apps remains uncertain. The survival of an app is contingent on financial, political, and regulatory factors [15]. Apps solely used as research tools can often be sustained through research funding. Conversely, apps intended for commercial use or integration into sectors such as health care usually require ongoing external financing, which may lead to discontinuation due to struggles to secure such support. Furthermore, political decisions (eg, national reimbursement policies, procurement rules, or regulatory priorities) also influence the adoption of apps or similar software in health care, which can pose challenges for researchers [15].

The Medical Device Regulation (MDR) is a specific European Union regulation affecting the classification and approval of medical devices, including software apps [16]. The introduction of the MDR, which became applicable in May 2021, has significant implications [16]. For instance, software that provides a responsive functionality, such as offering self-care advice based on self-reported values, such as health-related quality of life was, from that date, classified as a medical device (MDR class IIa) [16]. This classification necessitates a costly and time-consuming validation process that forces these apps to comply with stringent rules for classification, technical documentation, clinical evaluation, and postmarket surveillance to remain on the market. Although the MDR has been established to enhance patient safety and create a more robust regulatory framework, it presents challenges, particularly for smaller companies and academic contexts with limited resources. The increased regulatory burden and associated costs may disproportionately affect these entities, potentially reducing the number of mHealth apps available on the European Union market [17].

No scoping review has delineated the changes in active apps before and after the implementation of the MDR. This study aimed to describe the longevity (ie, whether apps identified in 2018 remained accessible in 2024) of English-language mobile apps before and after implementation of legislation regulating medical devices by examining (1) the availability of these apps in mobile app stores and (2) the extent to which they are supported by scientific evidence.

Methods

Our review was reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines (Checklist 1) [18]. We used the InSynQ list for items relevant to scoping reviews [19]. Quality assessment and risk of bias were not evaluated because the primary aim was to identify mobile device app measures, consistent with scoping review recommendations [18]. An internal study protocol was established before conducting the study. Because the protocol was not publicly registered, it is available from the corresponding author on reasonable request.

Information Sources and Search for Apps

To investigate the longevity of these apps, we examined cancer-related mobile apps available since September 2018, as identified in the study by Adam et al [13], and compared them with our search conducted in December 2024. Accordingly, all mobile apps from Adam et al's [13] study were included in separate analyses. However, different inclusion and exclusion criteria were applied to address the 2 additional research questions, and multiple approaches were used to identify available apps with supporting literature.

Search Strategy and Screening

First, on December 2, 2024, we searched Apple's App Store and Google Play to identify currently available cancer recovery

XSL•FO
RenderX

mobile apps. We explored both the apps still available from Adam et al's [13] study and additional apps identified through our search. Two researchers (KF and SJ) performed the searches, resolving discrepancies through discussion or consulting with a third author (AS). The search strategy described by Adam et al [13] was used. Specifically, we used the exact same keywords from the original review ("cancer," "cancer survivor," and "cancer survivorship") and conducted 2 searches using an iPhone and a Samsung phone to search Apple's App Store and Google Play, respectively.

Second, a literature search was conducted to understand which cancer recovery apps were supported by scientific evidence. Searches were conducted in MEDLINE, Embase, the Cochrane Library, and Web of Science databases. The literature search was conducted on October 4, 2022, and updated on August 15, 2024. Medical Subject Headings (MeSH) and free-text terms were identified for each search concept. The full search strategies, including all MeSH and free-text terms, are provided in Multimedia Appendix 1. For mobile apps beyond those identified in the study by Adam et al [13] but found in our December 2024 search, we explored the literature in PubMed using free-text searches for the app names. We also conducted Google searches using the names of apps identified by Adam et al [13].

We developed and pilot-tested the abstract and full-text review criteria that reflected our inclusion and exclusion criteria. For pilot testing, the reviewers assessed 10% of abstracts and full-text articles. Any discrepancies were discussed within the research group and, if necessary, the process was updated accordingly. Two researchers (KF and SJ) independently evaluated the literature search results according to the eligibility criteria using Rayyan web-based software [20]. Discrepancies were resolved through discussion or consultation with a third author (AS). Initially, titles and abstracts were screened for eligibility, followed by a full-text review of articles deemed to align with the eligibility criteria. The search strategy was developed in MEDLINE (Ovid) in collaboration with librarians at the Karolinska Institute University Library. MeSH and free-text terms were identified for each search concept. The search strategy was then translated into the other databases using the Polyglot web-based tool [21].

No language restrictions were applied to the database searches. However, studies not written in English were excluded during screening due to limited translation resources. Articles published before 2010 were excluded. The strategies were peer-reviewed by another librarian prior to execution. Deduplication was performed as described by Bramer et al [22]. A final step was added to compare the DOIs.

Third, a snowball search was conducted to check references and citations of eligible studies identified through the database searches using the Citationchaser web-based tool [23].

Eligibility Criteria

Study Inclusion Criteria

As this was a scoping review, we aimed to include all relevant study designs that provided quantitative measures of impact. We considered experimental studies of any size that reported at least 1 measure of efficacy or effectiveness of a mobile app. Only studies published from 2010 onwards were eligible.

Study Exclusion Criteria

Observational studies, qualitative studies, editorials, reviews, opinion articles, case studies, research protocols, trial registrations, conference summaries, and papers published as abstracts were excluded. Studies that reported only qualitative measures were excluded. Due to limited resources, studies not written in English were excluded.

Population Inclusion Criteria

This study included primary studies of adult cancer survivors (18 years or older) who had completed treatment with curative intent before participating in the study.

Population Exclusion Criteria

Individuals with incurable diseases or those at the end of life were excluded. Studies that included patients younger than 18 years were excluded, even if the mean age was greater than 18 years. Studies that included patients with childhood cancer, who were subsequently included as adults, were also excluded.

Intervention Inclusion Criteria

We included all studies that examined smartphone apps designed for cancer survivors, with content presented in English that reported at least 1 measure of efficacy or effectiveness. The duration of the intervention had to be at least 6 weeks to ensure the intervention period was long enough to assess meaningful change in quantitative outcomes. Eligible studies required 1 treatment arm (or the treatment arm in single-arm studies) to consist solely of the mobile app or the mobile app in addition to usual care (as defined by the study authors).

Intervention Exclusion Criteria

We did not consider mobile apps used solely to identify or diagnose specific cancer types, provide general cancer information, manage hospital appointments or medication schedules, supply tools for health care professionals or caregivers, or support patients before or during treatment. Web-based apps were not included. Apps exclusively designed for tablets were also excluded because of their lower accessibility and use rates compared with mobile phones among the general population. Lastly, we excluded apps involving person-to-person interactions, such as cognitive behavioral therapy (CBT) or general counseling, as it would increase costs and resource demands.

Mobile App Characteristics Inclusion Criteria

Beyond apps identified by Adam et al [13], additional apps were eligible if they were available in one of the mobile app stores (Apple's App Store or Google Play) and supported by evidence that met the study eligibility criteria. The eligibility of apps was evaluated using the cancer survivorship care quality framework [24], which required that the app address at least 1 of the following domains: surveillance and management of physical effects, psychosocial effects, or chronic medical conditions.

XSL•FO
RenderX

Mobile App Characteristics Exclusion Criteria

Apps were excluded if they did not address any survivorship domain defined in the cancer survivorship care quality framework or did not meet the availability and evidence requirements described above.

Data Charting Process

During the study planning phase, a checklist for data extraction was developed (see Multimedia Appendix 1). We developed data extraction forms to collect relevant information from each article. The forms were piloted on 5 apps, and any discrepancies were discussed within the research group. Two researchers (KF and SJ) independently extracted relevant data (from mobile app stores and papers), except for the information on the app intervention (as described in mobile app stores), which was extracted by one reviewer (KF) and verified by another reviewer (SJ). This approach was deemed most acceptable, as the information on mobile app stores was transcribed in a free-text format and could not be extracted in precisely the same manner between reviewers. Any disagreements were resolved through discussion or by seeking input from a third reviewer (AS) when needed. The shortest follow-up was set at 6 weeks. We included all studies that investigated smartphone apps designed for cancer survivors that reported at least 1 measure of efficacy or effectiveness. We took this approach because we wanted to investigate not only existing mobile apps on the market but also apps that could potentially affect survivorship. The following information was collected from each study or the mobile app:

- Trial details (sample size, first author's name, publication year, country, funding)
- Participant information (type of cancer)
- Mobile app specifics (name of the app, availability in app stores, latest update of the app, intervention characteristics [as described in mobile app stores], cancer survivorship care quality framework domain covered)

Synthesis of Results

This scoping review used a narrative approach aligned with the research objectives to summarize the results. The data are presented in tabular and narrative formats, offering a comprehensive overview of the findings.

Patient and Public Involvement

The extensive involvement of patients and their caregivers in our Surgical Care Science research partnership group [25] was integral to the study design. Their contributions encompassed proposing research questions and informing study methodologies. Feedback from these stakeholders on the study's findings was actively solicited, and insights derived from group discussions were incorporated into the Discussion section.

Results

Mobile Apps Found in App Stores

In the study by Adam et al [13], 151 mobile apps were identified. Of these, 45 of 151 (30%) were still available on either Apple's App Store or Google Play in December 2024. A total of 16 of 151 mobile apps (11%) were updated in 2024, and 25 of 151 apps (17%) were updated between 2023 and 2024. In our search in Apple's App Store and Google Play in December 2024, we found 1 new app from 2024 that fulfilled our eligibility criteria [26]. This mobile app was developed to assist cancer survivors in managing insomnia through CBT. Detailed descriptions of all identified apps are provided in Multimedia Appendices 2-4.

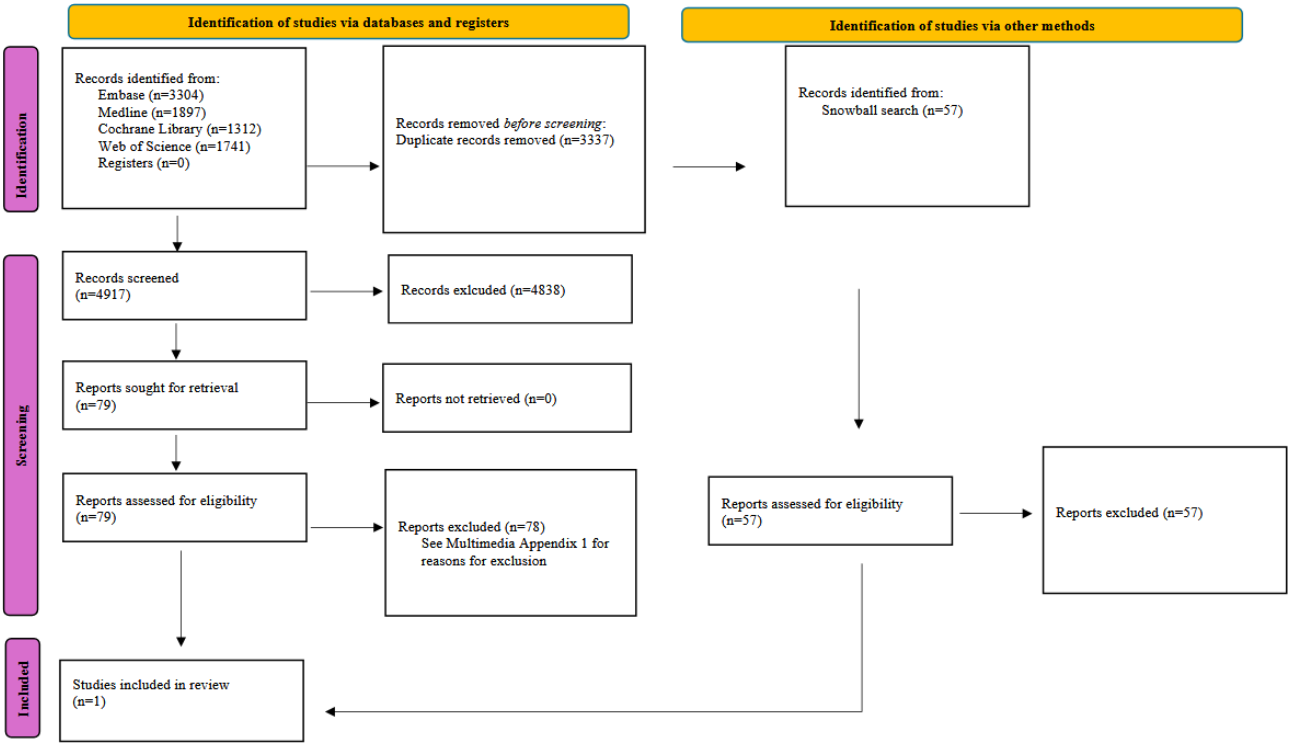
In the December 2024 Google Play search, we identified 225 search results for "cancer," 52 for "cancer survivor," and 3 for "cancer survivorship." In Apple's App Store, we obtained 175 results for "cancer," 20 for "cancer survivor," and 10 for "cancer survivorship." The searches for "cancer survivor" and "cancer survivorship" did not render any unique results but were covered by the initial search for "cancer" both in Google Play and in Apple's App Store. The search did not yield any additional mobile apps. However, 1 additional mobile app was found in the literature search (Multimedia Appendix 4) [26]. The reasons for excluding apps included that they were exclusively for clinicians (n=51); were fundraising apps (n=6); were not exclusively for the treatment period (n=11); did not fulfill any of the cancer survivorship care quality framework domains (n=66); were conference apps (n=6); were for children's cancer only (n=6); were games (n=4); were covered by Adam et al [13] (n=24); were apps for obtaining a diagnosis (n=23); were not supported by evidence (n=1); had no content related to cancer (n=164); and were not English (n=38).

Literature Supporting Apps

The initial search conducted in 2022 yielded 6238 papers, and the updated search conducted in 2024 yielded 2016 new papers. A total of 3337 duplicate papers were removed. After reviewing the literature, only 1 article was included in this study. That study evaluated a mobile app developed to help cancer survivors manage insomnia through CBT. It offers a self-guided, 7-week program featuring sleep hygiene education, personalized sleep plans, relaxation techniques, and motivational interviewing. Users can track their sleep metrics and receive tailored advice to address their negative beliefs about sleep. A single-arm interventional study involving 30 participants significantly improved insomnia severity among cancer patients [26]. Figure 1 shows a PRISMA flowchart of the search strategy. See Multimedia Appendix 2 for reasons for exclusion.

XSL•FO
RenderX

Figure 1. Flow diagram of search strategy (adapted from Page et al [27]).



Discussion

Principal Findings

This scoping review aimed to evaluate the durability of English-language mobile apps intended to aid cancer patients' recovery after treatment and to determine the level of scientific evidence supporting their effectiveness. In 2018, 151 mobile apps were available, of which 45 of 151 (30%) were still available as of December 2024. Remarkably, only 16 of 151 mobile apps (11%) were updated by 2024. In addition, we found 1 new app by December 2024.

Strengths and Limitations

To ensure the scientific rigor of this scoping review, we conducted a comprehensive literature search using 4 electronic databases, an additional snowball search to check the references and citations of eligible studies, and a broad search strategy. The review also used Apple's App Store and Google Play, which have almost 100% coverage of the mobile app market [21]. The review followed the PRISMA guidelines for scoping reviews [18] and was methodologically robust, with an independent blinded selection of studies and data extraction (except for extraction for interventions).

However, our scoping review had some limitations. To make the review more pragmatic, we focused on mobile apps available in app stores that followed at least 1 of the cancer survivorship care quality framework domains and had supporting evidence for effectiveness without cointerventions (excluding usual care). We also excluded face-to-face interactions within the app because of the cost and implementation challenges in real-world health care settings, which have been shown to hinder mobile app survival [28]. This meant that several trials and apps were excluded. Consequently, the data used in this review do not represent the entire market. This methodology does, however, draw attention to the scarcity of cancer recovery–focused mobile apps that are available for download from mobile app stores and have some scientific backing for their efficacy.

We also conducted searches using computers and tablets to determine whether the results differed owing to the algorithms provided by the manufacturers. We noticed that the search results differed depending on who conducted the searches, when they were conducted, and the devices used for the search (eg, tablets, computers, or phones). For example, when using a web browser, the maximum number of search results for Apple's App Store was 50, whereas for an iPhone, it was 250.

However, we consider the search terms to be accurate and representative of how a typical user searches for a mobile app related to the topic examined in this review. Finally, a limitation of searching for apps with associated studies is that the nomenclature must likely be included to identify them. Consequently, scientific support may exist for some apps, even though we did not find it.

Implementation Challenges

Proponents of mHealth assert that its primary advantages include improved efficacy, reduced expenditure, and expeditious delivery of health care with minimal risk [29]. Nevertheless, it is not evident how digital health technologies can complement the need for in-person care interactions and health care expenses [30]. Some research findings suggest that digital interventions may contribute to increased physical activity levels. However, the outcomes regarding psychological effects remain inconclusive. Notably, the evidence supporting these findings is limited [9].

Adherence is a well-known and complex problem that can affect the outcomes of mobile app interventions [31]. Although

XSL•FO
RenderX

developers have created a purportedly optimal app, its efficacy is contingent upon the health care sector's willingness or reluctance to integrate it into existing practices [32]. Adopting new technologies in health care settings often faces challenges owing to resistance to change and concerns about workflow disruption. Health care providers may be hesitant to integrate new mobile apps into their practice if they perceive them to be time-consuming or potentially compromise patient care [9,28]. The involvement of health care professionals in the development and implementation processes is crucial for overcoming barriers. This approach ensures that the technology addresses specific requirements and integrates effectively with existing workflows [33].

The implementation of mobile apps in health care settings presents several challenges. Strict regulations, such as the MDR in Europe, have increased development costs and complexity, potentially limiting the availability of apps, particularly from smaller enterprises [16]. Financial limitations also pose a major hurdle, as sustaining high-quality health care apps requires substantial resources [17]. Additionally, technical obstacles, adoption resistance, and insufficient evidence hinder the effective integration of mobile apps into health care environments [28]. These challenges underscore the necessity for careful consideration and strategic planning to overcome these barriers and ensure successful implementation.

Most mobile apps identified in the study by Adam et al [13] are no longer available, highlighting their rapid turnover and potential lack of sustainability in the mobile health app market for cancer survivors. This high attrition rate raises concerns regarding the continuity of the research process using these digital tools, as they frequently depend on research funding [34]. Research grants may be derived from taxpayer funds, and given the research inefficiencies observed in this area, action must be taken to enhance the sustainability of these apps and avoid research waste [35]. Implementing mobile apps in health care settings would optimize the allocation of public funds instead of merely conducting pilot studies and subsequently discarding apps. Despite the emergence of new mobile apps, overall scientific evidence supporting their effectiveness remains limited [11]. Several studies suggest that apps may reduce health care use related to disease management (eg, fewer clinic visits, hospital admissions, or unplanned contacts) and the overall financial burden of the disease; however, the cost-benefit relationship remains contentious due to high development and implementation costs, as well as limited user adoption [36]. These 2 critical factors must be established when proposing a change in the care-setting environment [37,38].

Conclusions

Despite the well-documented rapid growth of mobile health apps in the health care market, this scoping review found that there appear to be few publicly accessible, frequently updated, evidence-based apps for cancer recovery that are supported by clinical evidence. Even though many apps are available, the longevity of care recovery apps is very low. In addition, the rapid pace of mobile device app development has outstripped the evidence base, creating a gap that researchers are now trying to address while highlighting an evaluation crisis in the field. Ensuring sustained availability of mobile health apps necessitates heightened emphasis on more robust effectiveness trials, addressing barriers to implementation, and developing sustainable business models. Although some cost-effectiveness evaluations of mobile health apps exist, there is a clear need for more comprehensive and standardized assessments within care settings.

Acknowledgments

We thank Narcisa Hannerz and Sabina Gillsund from the Karolinska Institute University Library for their assistance in developing and peer reviewing the literature search strategies. We are also grateful to the members of the Surgical Care Science patient and caregiver partnership group for their valuable input on the study design and interpretation of the findings.

Funding

This work was supported by Region Stockholm (ALF project RS2021-0822), Sjöberg Foundation contract number 2024-01-04:06; The Swedish Research Council contract number 2023-02178, and the Swedish Cancer Society (230665 FE). The funders had no involvement in the study design, data collection, analysis, interpretation of results, or writing of the manuscript.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files in Multimedia Appendices 1-4. No additional datasets were generated. The full search strategies and extracted application-level data are provided in the supplementary files in Multimedia Appendices 1-4. The study protocol is available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: KF, LT, PL, AS
Methodology: KF, LT, SJ, AS
Data curation: KF, SJ
Formal analysis: KF, SJ
Investigation: KF, SJ, LT

XSL•FO
RenderX

Writing – original draft: KF, LT
Writing – review and editing: KF, LT, SJ, PL, AS
Supervision: AS, PL
Project administration: KF

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full search strategy, checklist template, data collection process, data items, study risk of bias assessment, effect measures, mobile apps organized by last update year, and papers that were excluded after reading their full texts.
[DOCX File, 27 KB - cancer_v12i1e82448_app1.docx]

Multimedia Appendix 2
Description of available apps, according to Adam et al [13].
[DOCX File, 51 KB - cancer_v12i1e82448_app2.docx]

Multimedia Appendix 3
Longevity of cancer recovery apps, 2018-2024.
[DOCX File, 35 KB - cancer_v12i1e82448_app3.docx]

Multimedia Appendix 4
Additional cancer recovery apps found in December 2024.
[DOCX File, 19 KB - cancer_v12i1e82448_app4.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 29 KB - cancer_v12i1e82448_app5.docx]

References

1. Lagergren P, Schandl A, Aaronson NK, et al. Cancer survivorship: an integral part of Europe's research agenda. Mol Oncol 2019 Mar;13(3):624-635. [doi: 10.1002/1878-0261.12428] [Medline: 30552794]
2. Bashore L, Hobbie W. Emerging and ongoing survivorship challenges among childhood cancer survivors and providing risk-based focused follow-up care. Semin Oncol Nurs 2021 Jun;37(3):151163. [doi: 10.1016/j.soncn.2021.151163] [Medline: 34144848]
3. Gatta G, Rossi S, Capocaccia R. Cancer burden estimates and forecasts: uses and cautions. Tumori 2013;99(3):439-443. [doi: 10.1177/030089161309900320] [Medline: 24158074]
4. Høeg BL, Bidstrup PE, Karlsen RV, et al. Follow-up strategies following completion of primary cancer treatment in adult cancer survivors. Cochrane Database Syst Rev 2019 Nov 21;2019(11):CD012425. [doi: 10.1002/14651858.CD012425.pub2] [Medline: 31750936]
5. Holme Ø, Bretthauer M, Løberg M, Kalager M, Adami HO. Revisiting surveillance schemes for cancer survivors. Ann Oncol 2019 Jun;30(6):881-883. [doi: 10.1093/annonc/mdz107]
6. Murray E, Hekler EB, Andersson G, et al. Evaluating digital health interventions. Am J Prev Med 2016 Nov;51(5):843-851. [doi: 10.1016/j.amepre.2016.06.008]
7. Stein KD, Syrjala KL, Andrykowski MA. Physical and psychological long-term and late effects of cancer. Cancer 2008 Jun 1;112(11 Suppl):2577-2592. [doi: 10.1002/cncr.23448] [Medline: 18428205]
8. Lau N, O'Daffer A, Yi-Frazier J, Rosenberg AR. Goldilocks and the three bears: a just-right hybrid model to synthesize the growing landscape of publicly available health-related mobile apps. J Med Internet Res 2021 Jun 7;23(6):e27105. [doi: 10.2196/27105] [Medline: 34096868]
9. Fournier V, Duprez C, Grynberg D, Antoine P, Lamore K. Are digital health interventions valuable to support patients with cancer and caregivers? An umbrella review of web-based and app-based supportive care interventions. Cancer Med 2023 Dec;12(23):21436-21451. [doi: 10.1002/cam4.6695] [Medline: 37937812]
10. Dickinson R, Hall S, Sinclair JE, Bond C, Murchie P. Using technology to deliver cancer follow-up: a systematic review. BMC Cancer 2014 May 3;14(1):311. [doi: 10.1186/1471-2407-14-311] [Medline: 24885758]
11. Changrani K, Chima S, Sharma A, et al. A systematic review of smartphone applications for cancer survivors. J Cancer Surviv 2024 Dec;18(6):1951-1973. [doi: 10.1007/s11764-023-01435-9] [Medline: 37700151]

XSL•FO
RenderX

12. Prentice C, Peven K, Zhaunova L, et al. Methods for evaluating the efficacy and effectiveness of direct-to-consumer mobile health apps: a scoping review. BMC Digit Health 2024;2(1). [doi: 10.1186/s44247-024-00092-x]
13. Adam R, McMichael D, Powell D, Murchie P. Publicly available apps for cancer survivors: a scoping review. BMJ Open 2019 Sep 30;9(9):e032510. [doi: 10.1136/bmjopen-2019-032510] [Medline: 31575584]
14. Kassianos AP, Emery JD, Murchie P, Walter FM. Smartphone applications for melanoma detection by community, patient and generalist clinician users: a review. Br J Dermatol 2015 Jun;172(6):1507-1518. [doi: 10.1111/bjd.13665] [Medline: 25600815]
15. Lee SM, Kim NR, Hong SG. Key success factors for mobile app platform activation. Serv Bus 2017 Mar;11(1):207-227. [doi: 10.1007/s11628-016-0329-y]
16. Bretthauer M, Gerke S, Hassan C, Ahmad OF, Mori Y. The new European medical device regulation: balancing innovation and patient safety. Ann Intern Med 2023 Jun;176(6):844-848. [doi: 10.7326/M23-0454] [Medline: 37068279]
17. Mohn D, Zehnder M. Medical device regulation (MDR) from a dental perspective. Front Dent Med 2023;4:1155820. [doi: 10.3389/fdmed.2023.1155820] [Medline: 39916893]
18. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-SCR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
19. Cumpston MS, McKenzie JE, Ryan R, Flemyng E, Thomas J, Brennan SE. Development of the InSynQ checklist: a tool for planning and reporting the synthesis questions in systematic reviews of interventions. Cochrane Evid Synth Methods 2023 Dec;1(10):e12036. [doi: 10.1002/cesm.12036] [Medline: 40476007]
20. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
21. Clark JM, Sanders S, Carter M, et al. Improving the translation of search strategies using the Polyglot Search Translator: a randomized controlled trial. J Med Libr Assoc 2020 Apr;108(2):195-207. [doi: 10.5195/jmla.2020.834] [Medline: 32256231]
22. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc 2016 Jul;104(3):240-243. [doi: 10.3163/1536-5050.104.3.014] [Medline: 27366130]
23. Haddaway NR, Grainger MJ, Gray CT. Citationchaser: a tool for transparent and efficient forward and backward citation chasing in systematic searching. Res Synth Methods 2022 Jul;13(4):533-545. [doi: 10.1002/jrsm.1563] [Medline: 35472127]
24. Nekhlyudov L, Mollica MA, Jacobsen PB, Mayer DK, Shulman LN, Geiger AM. Developing a quality of cancer survivorship care framework: implications for clinical care, research, and policy. JNCI 2019 Nov 1;111(11):1120-1130. [doi: 10.1093/jnci/djz089]
25. Schandl A, Mälberg K, Haglund L, Arnberg L, Lagergren P. Patient and public involvement in oesophageal cancer survivorship research. Acta Oncol 2022 Mar 4;61(3):371-377. [doi: 10.1080/0284186X.2021.2016950]
26. Arring N, Barton DL, Lafferty C, Cox B, Conroy DA, An L. Mi Sleep coach mobile app to address insomnia symptoms among cancer survivors: single-arm feasibility study. JMIR Form Res 2024 Apr 26;8:e55402. [doi: 10.2196/55402] [Medline: 38669678]
27. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
28. Zakerabasali S, Ayyoubzadeh SM, Baniasadi T, Yazdani A, Abhari S. Mobile health technology and healthcare providers: systemic barriers to adoption. Healthc Inform Res 2021 Oct;27(4):267-278. [doi: 10.4258/hir.2021.27.4.267] [Medline: 34788907]
29. Wattanapisit A, Teo CH, Wattanapisit S, Teoh E, Woo WJ, Ng CJ. Can mobile health apps replace GPs? A scoping review of comparisons between mobile apps and GP tasks. BMC Med Inform Decis Mak 2020 Jan 6;20(1):5. [doi: 10.1186/s12911-019-1016-4] [Medline: 31906985]
30. Peters L, Greenfield G, Majeed A, Hayhoe B. The impact of private online video consulting in primary care. J R Soc Med 2018 May;111(5):162-166. [doi: 10.1177/0141076818761383] [Medline: 29490152]
31. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 5;13(3):e52. [doi: 10.2196/jmir.1772] [Medline: 21821503]
32. Garavand A, Samadbeik M, Kafashi M, Abhari S. Acceptance of health information technologies, acceptance of mobile health: a review article. J Biomed Phys Eng 2017 Dec;7(4):403-408. [Medline: 29445717]
33. Moorthy K, Halliday LJ, Noor N, Peters CJ, Wynter-Blyth V, Urch CE. Feasibility of implementation and the impact of a digital prehabilitation service in patients undergoing treatment for oesophago-gastric cancer. Curr Oncol 2023 Jan 30;30(2):1673-1682. [doi: 10.3390/curroncol30020128] [Medline: 36826089]
34. Grimpe C. Extramural research grants and scientists' funding strategies: beggars cannot be choosers? Res Policy 2012 Oct;41(8):1448-1460. [doi: 10.1016/j.respol.2012.03.004]
35. Thornton JP. Flypaper nonprofits: the impact of federal grant structure on nonprofit expenditure decisions. Public Finance Rev 2012;42(2):176-198. [doi: 10.1177/1091142112446845]
36. An Q, Kelley MM, Hanners A, Yen PY. Sustainable development for mobile health apps using the human-centered design process. JMIR Form Res 2023 Aug 25;7:e45694. [doi: 10.2196/45694] [Medline: 37624639]

XSL•FO
RenderX

37. Herledan C, Falandry C, Huot L, et al. Clinical impact and cost-saving analysis of a comprehensive pharmaceutical care intervention in older patients with cancer. J Am Geriatr Soc 2024 Feb;72(2):567-578. [doi: 10.1111/jgs.18585] [Medline: 37818698]
38. Higgins AM, Harris AH. Health economic methods: cost-minimization, cost-effectiveness, cost-utility, and cost-benefit evaluations. Crit Care Clin 2012 Jan;28(1):11-24. [doi: 10.1016/j.ccc.2011.10.002] [Medline: 22123096]

Abbreviations

CBT: cognitive behavioral therapy
MDR: Medical Device Regulation
MeSH: Medical Subject Headings

Edited by M Balcarras; submitted 15.Aug.2025; peer-reviewed by AM Yacob, A Donate-Martinez; accepted 06.Jan.2026; published 11.Feb.2026.

Please cite as:
Färnqvist K, Thiele L, Johnsson S, Lagergren P, Schandl A
The Longevity of Mobile Apps for Cancer Recovery: Scoping Review
JMIR Cancer 2026;12:e82448
URL: https://cancer.jmir.org/2026/1/e82448
doi:10.2196/82448

© Kenneth Färnqvist, Luisa Thiele, Sophie Johnsson, Pernilla Lagergren, Anna Schandl. Originally published in JMIR Cancer (https://cancer.jmir.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Empathic AI for Patient-Centered Cancer Care: A Scoping Review of Patient Navigation, Support, and Clinical Practice

Brianna M White[1*], MPH, DrPHc; Gabriela L Aitken[2*], MD; Janet A Zink[1], PhD; Parnian Kheirkhah Rahimabad[1], PhD; Fekede Asefa Kumsa[1], PhD; Soheil Hashtarkhani[1], PhD; Rezaur Rashid[1], PhD; Saba Kheirinejad[1], PhD; Tyra Girdwood[3], PhD; Christopher L Brett[4], MBA, MD; Robert L Davis[1], MPH, MD; David L Schwartz[2], MD; Arash Shaban-Nejad[1], MSc, MPH, PhD

[1]Department of Pediatrics, College of Medicine, The University of Tennessee Health Science Center—Oak Ridge National Laboratory Center for Biomedical Informatics, 50 N Dunlap, 492R, Memphis, TN, United States

[2]Departments of Radiation Oncology & Preventive Medicine, The University of Tennessee Health Science Center, Memphis, TN, United States

[3]Department of Community and Population Health, College of Nursing, The University of Tennessee Health Science Center, Memphis, TN, United States

[4]Department of Radiation Oncology, University of Tennessee Graduate School of Medicine, Knoxville, TN, United States

[*]these authors contributed equally

Corresponding Author:
Arash Shaban-Nejad, MSc, MPH, PhD
Department of Pediatrics, College of Medicine, The University of Tennessee Health Science Center—Oak Ridge National Laboratory Center for Biomedical Informatics, 50 N Dunlap, 492R, Memphis, TN, United States

Abstract

Background: Artificial intelligence (AI) is rapidly reshaping oncology, offering advancements in clinical care and patient support. A growing area of interest is the integration of empathic AI: systems integrating clinical precision with emotional intelligence to support medical decision-making and the emotional and psychosocial well-being of patients and caregivers.

Objective: This review aimed to explore the role of empathic AI in cancer care, with a focus on its applications in patient education, clinician support, and emotional care. It also evaluated the ethical, cultural, and implementation challenges associated with its integration into clinical practice in oncology.

Methods: A systematic search of literature was conducted in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines. Peer-reviewed papers published between January 2018 and January 2025 were identified through a search of PubMed, Scopus, and IEEE Xplore, and citation tracking. Eligible studies focused on applications of empathic AI in oncology. A total of 44 studies were included and analyzed thematically across 3 core clinical domains: tailored communication and education, diagnostics and care plan optimization, and emotional and psychosocial support.

Results: Empathic AI demonstrates the potential to improve cancer care by enhancing patient education, clinical decision-making, and emotional support. Common applications include personalized education for patients and providers, support for diagnostic and therapeutic decisions, and tools designed to recognize and respond to patient distress. Several studies noted improved patient engagement and reduced clinician burden. However, concerns were identified regarding overreliance on AI systems, cultural insensitivity, and patient privacy.

Conclusions: Empathic AI represents a promising advancement in patient-centered oncology, integrating emotional intelligence into clinical care. Its successful implementation will require careful attention to ethical, cultural, and clinical considerations to ensure health equity, trust, and safety in AI-assisted cancer care.

(JMIR Cancer 2026;12:e82336) doi:10.2196/82336

KEYWORDS

empathic AI; oncology; cancer care; patient-centered care; AI in health care; digital health; health equity; artificial intelligence

Introduction

Artificial intelligence (AI) has become increasingly embedded in medical practice over the past decade, enhancing care capabilities from the development of personalized health libraries and innovative patient education strategies to disease screening, diagnostics, and treatment planning [1-4]. These technologies are reshaping how care is accessed, delivered, and experienced. In oncology, AI models are being developed to standardize and improve the consistency and quality of care, offering new ways to support both clinicians and patients in navigating complex diagnostic and treatment pathways [5].

XSL•FO
RenderX

These technologies provide timely, accurate, and accessible information that can empower patients and caregivers, supporting engagement, comprehension, and autonomy throughout the cancer care continuum [6-8].

For example, AI has rapidly become a critical adjunct for medical image analytics, demonstrating that computer models can achieve state-of-the-art tumor detection, segmentation, extraction, and classification across a wide range of cancer types [9]. These advancements in AI in health care have improved diagnostic accuracy and enabled earlier cancer detection, allowing clinicians to initiate timely, evidence-based interventions. Beyond imaging, AI supports personalized oncology through predictive models that generate tailored treatment recommendations based on individual patient demographics, genomics, and clinical profiles [10]. These technologies have ushered in a new era of precision health and medicine [11], specifically impacting the field of oncology, informing care strategies that range from recurrence prediction and surgical planning to emotional assessment. By integrating psychosocial considerations into clinical decision-making, AI holds promise not only for improving treatment efficacy but also for enhancing overall quality of life [10]. This progress may help address systemic gaps in access to emotional support and culturally responsive communication, which have been long-standing public health challenges in cancer care.

However, despite these advances, the integration of AI into emotionally sensitive domains such as oncology introduces a new set of unresolved ethical, relational, and epistemological challenges [12]. The emotional toll of a cancer diagnosis underscores the critical need for empathetic care. Empathy, long considered a distinctly human element of care, has historically been difficult to replicate through technology. Yet recent advancements suggest that AI systems may approximate emotional attunement, recognize distress, and respond in ways perceived as empathic and compassionate [13]. In this context, empathic AI refers to technologies designed to detect and respond to users' emotional states to foster interactions that feel emotionally responsive and humanlike [13].

This review examines the role of empathic AI in cancer care, with a focus on its potential to enhance patient-centered, holistic treatment. We synthesize current advancements, challenges, and ethical considerations to inform the development of empathic AI systems that support both clinical effectiveness and emotionally responsive care. Given the distinct emotional, ethical, and communicative demands of oncology, this review underscores the need for a field-specific framework to guide the safe, equitable, and meaningful integration of empathic AI into cancer care delivery. From a public health perspective, thoughtful integration can inform scalable interventions, guide equitable policy design, and improve outcomes for historically underserved populations.

Methods

Search Strategy

The initial review of literature incorporated peer-reviewed studies that were identified on January 6, 2025, from PubMed, published between January 2018 and January 2025. This timeline was adopted to ensure that the review focuses on the most recent advancements and innovations in the use of empathic AI in cancer care as technological developments are rapidly evolving. Our search terms were used based on the thesaurus and keywords including AI, empathic, empathy, emotion, emotional intelligence, cancer, tumor, and oncology. Terms and keywords were combined to comprise search phrases, allowing for a more sensitive search. Search phrases included "artificial intelligence" OR "AI" AND "empathic" OR "empathy" OR "emotion" OR "emotional intelligence" AND "cancer" OR "oncology".

Following peer review, the search strategy was expanded to strengthen database scope. An updated search was conducted in Scopus and IEEE Xplore on January 1, 2026, using the same search terms and inclusion criteria, and newly identified studies were screened and extracted using the same procedures as the original search. Studies identified through the updated search were fully integrated into the descriptive and thematic analyses, and totals reflect the combined results of both search phases.

Eligibility Criteria

Studies were included for review if they met the following eligibility criteria: (1) published from January 2018 to January 2025, (2) explored the relationship between empathic AI and cancer care, and (3) available in the English language. Studies were excluded if they were not original research or were not peer-reviewed.

Study Selection, Data Extraction, and Analysis

The initial electronic database search yielded 78 results, which were collectively screened by 2 reviewers (GA and JZ) using the eligibility criteria. Following the abstract and full-text screening, 64 studies were eliminated because they did not meet the inclusion criteria. To expand the scope of relevant literature, a snowballing technique was used, where key studies identified during the initial search were used to trace additional sources through references and citations. This process led to the identification of 15 additional studies, enhancing the comprehensiveness of the review. A third reviewer (BW) verified the full set of included and excluded studies, cross-referencing decisions with the original criteria and extraction tables. Any uncertainties or inconsistencies were discussed collaboratively among the 3 reviewers to ensure a consistent and transparent review process. Study quality indicators (ie, peer-review status, study design, and sample size) were collected during data extraction to contextualize findings and mitigate potential sources of bias.

Terminology

Throughout this review, we use the term "empathic AI" as an umbrella descriptor for AI systems designed to recognize, respond to, or adapt to users' emotional states in health care contexts. Closely related terms such as empathetic AI or emotionally intelligent AI appear sparingly and only when reflecting terminology used in the source literature. These terms are treated as conceptually overlapping within this review, and their use does not imply substantive distinctions in system function or intent.

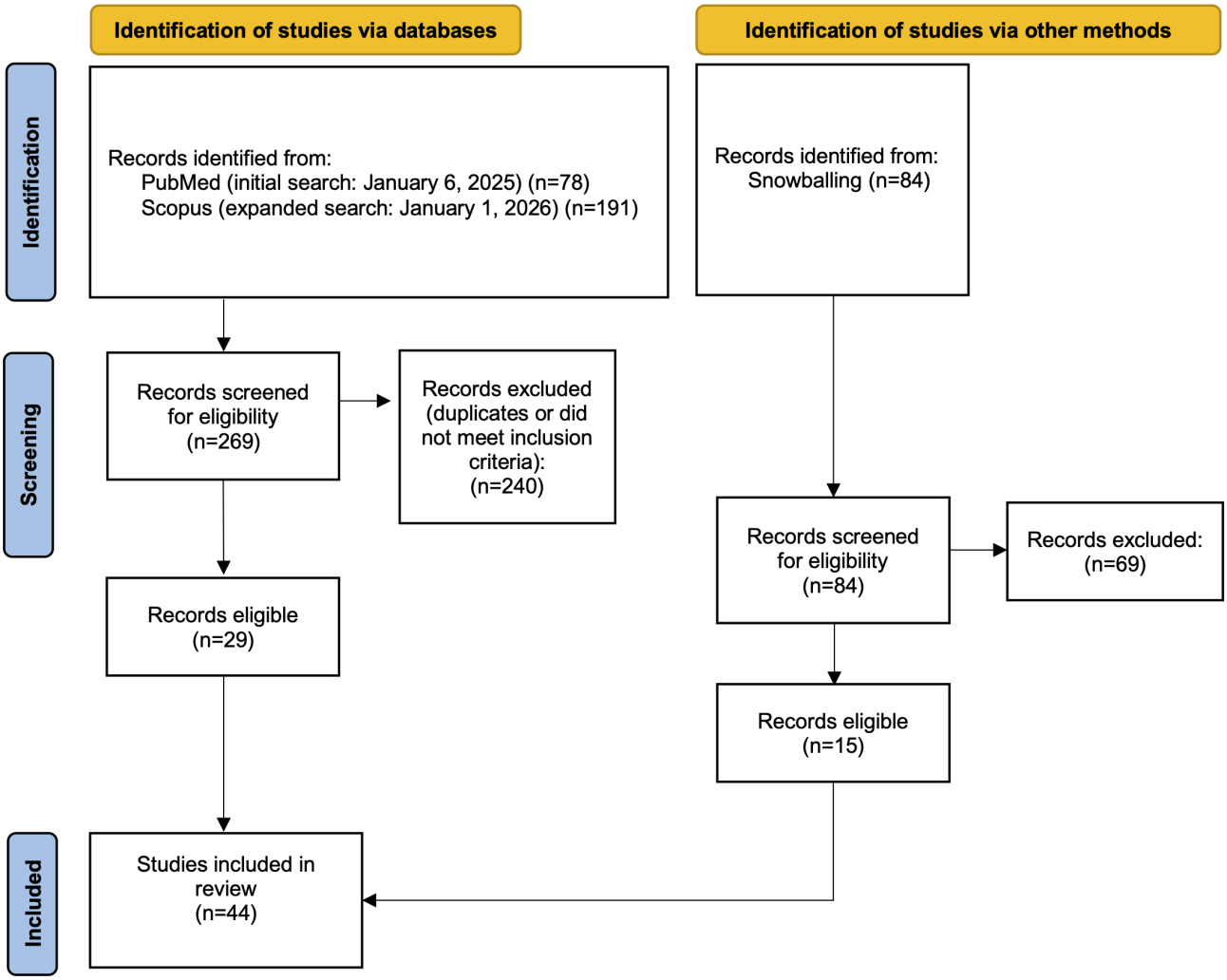
XSL•FO
RenderX

Results

Overview

As shown in Figure 1, a total of 44 studies met the inclusion criteria for review. Study characteristics are summarized in Table S1 in Multimedia Appendix 1 [14-57]. The included literature spans from June 2018 to January 2025, reflecting the evolution of empathic AI applications in cancer care over time. A synthesis of key applications, thematic domains, and representative findings across the included studies is shown in Table 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram of study selection process.



XSL•FO
RenderX

Table . Summary of key findings by clinical domain of empathic artificial intelligence in cancer care.

Domain	Key applications	Representative findings	Included references
Tailored communication and education	• Chatbots delivering emotionally calibrated, context-aware patient education • LLMs[a] and NLP[b] tools supporting empathic provider communication • Real-time summary tools and simulation-based training	• Improved patient comprehension, reduced distress, and greater engagement in care • Enhanced clinician ability to convey complex or difficult information empathetically • Reduced provider burnout and communication fatigue	[14-51]
Enhanced diagnostics and treatment planning	• Integration of clinical and emotional data for diagnostics and care planning • AI[c]-assisted screening, documentation, and decision-making • Tailored treatment strategies based on emotional state, literacy, preferences, and trust	• Increased diagnostic accuracy and efficiency • More personalized and emotionally attuned treatment plans • Improved adherence, patient satisfaction, and provider-patient trust	[17,18,20,22,23,25-27,31,34,37-43,45-47,50-54]
Emotional and psychosocial support	• NLP and affective computing to detect distress and offer emotional responses • Chatbots and support tools applying positive psychology principles • LLMs and cofacilitators offering caregiver guidance • VR[d] interventions and NLP tools for end-of-life discussions	• Reduced anxiety, isolation, and emotional distress in patients • Enhanced emotional support and resilience for caregivers • Improved mood, stress reduction, and support in palliative and end-of-life contexts	[14-17,19-36,38,39,41-51,53-56]

[a]LLMs: large language models.

[b]NLP: natural language processing.

[c]AI: artificial intelligence.

[d]VR: virtual reality.

Thematic analysis was conducted through a qualitative synthesis of the studies shown in Table S1 in Multimedia Appendix 1. Each paper was reviewed to extract key concepts, use cases, and reported outcomes relating to the application of empathic AI in cancer care. To organize these findings meaningfully, we used an inductive approach, identifying patterns that emerged directly from review of the data. These patterns were then organized into three overarching clinical domains: (1) tailored communication and education, (2) enhanced diagnostics and treatment planning, and (3) emotional and psychosocial support. The domains were refined iteratively as themes reoccurred across studies, allowing grouping of findings into categories that reflect the diverse roles of empathic AI in cancer care.

Tailored Communication and Education

Tailored communication and educational interventions emerged as a central theme across the majority of included studies, underscoring the potential of empathic AI to deliver information that is both clinically relevant and dynamically attuned to the emotional and cognitive needs of both patients and providers.

Tailored Patient Education

Empathic AI systems supporting personalized patient education were reported in 86% (38/44) of studies, highlighting growing adoption of these technologies to enhance informational access and emotional resonance in patient care. These tools tailor educational content to individual patient contexts, including diagnosis, treatment phase, emotional state, and cognitive load. Chatbots and AI cofacilitators were most often used to simplify complex medical concepts, offer condition-specific coping strategies, and reduce distress through emotionally calibrated messaging.

For example, Greer's *Vivibot* chatbot leveraged a positive psychology framework to support emotional regulation in young adult cancer survivors, resulting in reduced anxiety and high user satisfaction [14]. Studies by Leung et al [15,16] evaluated AI cofacilitators that adapted group-based educational content in real time based on participant mood and interaction dynamics, enhancing cohesion and engagement. Similarly, Chaix et al [17] and Siglen et al [18] reported that chatbot-delivered education improved adherence to treatment regimens and facilitated greater engagement in shared decision-making. By modifying tone,

XSL•FO
RenderX

language, and complexity in real time, AI tools improved patient comprehension and engagement.

Chen et al [19] evaluated the use of AI chatbots such as GPT-3.5, GPT-4, and Claude AI to generate high-quality, empathetic, and readable responses to cancer-related questions posed by patients on social media. The study found that AI-generated responses, particularly those from Claude AI, outperformed those from oncologists in terms of empathy, clarity, and emotional resonance, suggesting that AI chatbots can serve as effective tools for enhancing patient education and communication. Other studies specifically highlighted the role of empathic AI in expanding access to reliable, emotionally attuned health information among historically underserved or low-literacy populations [20,21]. These findings position AI as a tool to enhance patient education equity, particularly in contexts where human resources are limited, or in settings where patients may be hesitant to ask questions directly.

Recent studies further extend this work by demonstrating how generative AI tools tailor educational content based on patient demographics, clinical context, and informational needs, including evaluations of ChatGPT and other large language models (LLMs) responding to cancer-related questions and decision-making scenarios [22-25].

However, concerns about misinformation, cultural insensitivity, and overreliance on AI-generated outputs were raised across multiple studies [26-28]. These findings emphasize the importance of clinical oversight and algorithmic guardrails. Scalability and safety remain central to the integration of AI-mediated education in oncology as these tools move closer to frontline deployment.

Tailored Provider Education

Empathic AI was also used to support tailored provider education in 52% (23/44) of studies, reflecting a growing interest in leveraging AI to enhance clinical decision-making and support emotionally intelligent communication. These tools were positioned as cognitive extenders, offering context-specific recommendations on treatment protocols, synthesizing complex medical data, and facilitating more empathic responses during patient interactions. As 1 study highlighted, this may be particularly important in high-stress oncology settings for reducing clinician burnout and compassion fatigue [29].

Several studies examined the use of LLMs to help clinicians better understand caregiver perspectives, enabling the cocreation of more responsive group interventions [30,31]. Providers received AI-generated conversation summaries or context-sensitive prompts, which improved their ability to convey difficult information with empathy and precision. Others found that AI-generated patient-facing responses were rated significantly higher in quality, clarity, and empathy than those written by clinicians under time pressure, pointing to the potential of AI to support consistency in emotionally attuned communication [21,28,32]. One study comparing physician and AI chatbot responses with patient inquiries on social media found that AI replies were preferred 78.6% of the time [32].

Furthermore, De Silva et al [33] explored the use of machine learning and natural language processing (NLP) to analyze patient interactions in online cancer support groups. By identifying patient-reported behaviors, emotions, and treatment decisions, AI tools provided valuable insights into patient decision-making, emotional dynamics, and care preferences. The findings highlighted how AI-driven insights could help clinicians better understand behaviors, decisions, side effects, and emotional cues within support groups to derive actionable insights for patient-centered care. Recent work has also explored hybrid and simulation-based approaches, including live-chat models emphasizing digital clinical empathy [34] and virtual patient systems designed to train clinicians in empathic and end-of-life communication [35,36]. Other studies explored the use of empathic AI for real-time feedback and simulation-based training, with AI tools positioned not just as repositories of information but as companions in ethical and emotional reasoning [26,27].

These findings underscore the potential of AI to cultivate emotionally intelligent clinician communication. Nevertheless, integration challenges remain. Several studies raised concerns about how increasing reliance on AI may shift the balance between algorithmic input and clinical judgment, prompting debate about the potential erosion or recalibration of clinical intuition. While some view this shift as a loss, others argue that AI could enhance intuition by providing real-time feedback and mitigating cognitive biases [21,37]. Although AI may reduce cognitive load and offer scalable training solutions, its ability to cultivate true empathic capacity remains contested, which raises important questions for future research and professional development strategies.

Enhanced Diagnostics and Treatment Planning

The decision-making use of empathic AI in cancer care was explored across several areas, from diagnostics to care plan optimization. In total, 70% (31/44) of the included studies addressed the use of empathic AI in the clinical space, with approximately half examining both diagnostics and treatment planning. Studies in this theme revealed how AI systems are increasingly integrated into clinical decision-making, optimizing care goals and improving patient outcomes.

Diagnostics and Decision-Making Support

Empathic AI systems have been shown to support clinicians in making more accurate diagnoses and treatment decisions. Among the studies reviewed, 43% (19/44) discussed how AI is being used to aid in diagnostic reasoning and clinical interpretation. By integrating emotional intelligence with clinical data, these tools help enhance the accuracy of patient assessments, considering both clinical factors and the emotional context of the patient's situation. For example, ChatGPT has been used to assist in the interpretation of personalized breast cancer screening studies with 80% accuracy while reducing clinician workload by 70% [38]. Similarly, these algorithms have supported individualized radiation therapy dosage planning [39], enhancing diagnostic efficiency and precision in a patient-centered manner. While these applications are technically oriented, the authors emphasize that "using AI and combining it with human wisdom, empathy, and affection will be the method of choice for further fruitful development of tomorrow's

XSL•FO
RenderX

senology" [38], highlighting the envisioned role of empathy even within data-driven clinical tools [40,52].

Furthermore, NLP models have been used to analyze patients' decision-making behaviors and emotions [33,41], allowing clinicians to understand patients better and tailor their recommendations accordingly. When it comes to clinical documentation, AI-powered notes written by OpenAI's ChatGPT 3.5 and GPT 4.0 have been found to be more personalized and empathetic [42]. Together, these functions of AI platforms create not only a precise and tailored patient experience but also one that is more patient-specific. While AI integration into the clinical realm is in its infancy, it has a definite role in enhancing patient care delivery and optimization.

Care Plan Optimization

Several studies examined how AI-driven systems can optimize care plans by considering both medical and psychosocial factors. In total, 61% (27/44) of studies addressed how AI enhances patient-centered decisions and care. Specifically, these AI tools assist in tailoring treatment regimens that account for individual patient preferences and emotional well-being, ensuring a more holistic approach to cancer care. For example, AI models such as the Vik chatbot, designed to support patients with breast cancer[17], can personalize conversations and treatment recommendations based on a patient's literacy level and personal context. This helps patients make informed decisions and improves treatment adherence [57]. Others emphasize the influence of AI-generated guidance on treatment decision-making in advanced or complex cancer contexts, including variability in tone, content, and recommendations that may shape patient choices [22,52]. Furthermore, multiple studies examining chatbots showed that patients are capable of trusting machines [27], including the Rosa chatbot focused on breast cancer gene testing embedded into an app for patients with breast and ovarian cancer [18]. AI can, therefore, be used to detect distress and other emotions as shown by the results of using an AI-based therapy group cofacilitator [16], which allows physicians to tailor their patient-to-provider interactions accordingly, resulting in more patient-centered care. Through these functions, AI platforms foster stronger patient-provider relationships in the clinical landscape [43].

Studies included in this theme demonstrate that AI enhances patient-provider communication and real-time emotional support, leading to optimized care plans. Although concerns regarding machine privacy remain, these studies support the use of AI systems as copilots or cofacilitators in clinical decision-making to enhance the holistic work of physicians.

Emotional and Psychosocial Support

Another prominent theme in the studies was the emotional and psychosocial support that empathic AI can offer to both patients and caregivers. To underscore the importance and timeliness of this theme, the vast majority of studies (40/44, 91%) alluded to applications of the technology involving patient or caregiver emotional support. These AI systems are increasingly seen as tools for alleviating stress, providing reassurance, and offering emotional guidance during challenging cancer journeys.

Patient Emotional Support

Several studies focused on the ways empathic AI can provide emotional comfort to patients, offering them a sense of companionship and reducing feelings of isolation. A large majority of studies reviewed (40/44, 91%) highlighted the use of empathic AI in the service of patient emotional support. By using NLP and affective computing, AI systems were able to recognize and respond to emotional cues, providing timely encouragement or referrals to professional counseling services [53,54]. Likewise, Mårell-Olsson et al [44] demonstrated that socially intelligent agents can address nonmedical needs in pediatric populations by reducing feelings of isolation, supporting coping, and facilitating participation in daily activities.

The evaluation of the ability of various recently released AI chatbots to produce easily understandable, medically relevant, culturally appropriate, and empathetic responses to oncology patient queries was a key focus of many of the included studies [19,32,45,46]. While chatbots may be more easily applied to diagnostics, treatment, and monitoring [26], progress on interpersonal-type communications is being made, despite challenges noted on the algorithmic development side [45]. As a sign of how far these chatbots have come, one social media–based study found that nonhuman responses were rated higher on every measure by expert health care evaluators, including empathy and quality, and greatly preferred to the clinician responses [32]. Others further illustrate these applications through immersive and affect-aware interventions, including AI-driven virtual reality (VR) therapy for anxiety and depression [47,48] and sentiment-aware conversational agents designed to detect and respond to emotional distress [49].

In another notable example, the unique emotional needs of teen and young adult patients with cancer were addressed by the delivery of positive psychology skills via the Vivibot chatbot [14]. As this age group typically experiences heightened anxiety and depression posttreatment completion, this kind of intervention fills a previously unmet need [55].

Caregiver Support

Empathic AI also plays a critical role in supporting caregivers, who often experience high levels of stress and burnout. Unsurprisingly, 52% (23/44) of studies were also concerned with addressing aspects of caregiver emotional support. AI applications were found to offer tailored resources, emotional support, and coping strategies, helping caregivers manage their own well-being while supporting their loved ones through cancer care. For example, AI-based cofacilitators for online cancer support groups were shown to provide targeted emotional support for both patients and their caregivers in response to the detection of emotional distress [16]. Also, as patient care is a complex endeavor involving an interconnected team of clinicians, patients, and caregivers, the use of empathic AI chatbots can be at the center of an integrated and holistic approach to support [26,43,50]. Additional evidence highlights AI-supported environments that indirectly or explicitly assist caregivers through shared decision-making and emotional scaffolding, particularly in advanced cancer contexts [22,34].

XSL•FO
RenderX

Of particular interest is the potential of LLMs as providers of real-time, individually customized, and up-to-date medical information, with their concomitant risks and benefits, to allow for informed, complex, and empathic decision-making among caregivers for oncology patients [30]. A more speculative take on AI technologies in health care argues that the next step may well be social robots imbued with artificial empathy as the caregivers of the future [20]. Ultimately, any of these configurations may be thought of as hybrid intelligent human-AI care systems.

Palliative and End-of-Life Care

In palliative and end-of-life care settings, empathic AI has been particularly effective in providing compassionate communication and emotional support, with 30% (13/44) of studies referencing the use of AI in these contexts. These systems help both patients and health care providers navigate sensitive conversations, ensuring that care remains patient-centered and aligned with individual values during the final stages of life. For instance, the VR psychological support tool used in the SafeSpace study [51] sought to integrate guided relaxation exercises and compassionate mind training for oncology patients afflicted with a range of cancer types, including those under palliative care. Statistically significant improvements were found in the psychological measures of mood and emotional well-being, with overall reductions in stress levels demonstrated due to the safe and soothing environment simulated by the intervention. Simulated VR travel to memorable places was another unique application of empathic AI that boosted physiological and mental health outcomes, particularly depression and anxiety, for patients with terminal cancer[51]. Other researchers have used NLP approaches, such as analyzing emotion intensity or identifying psychosocial challenges, to address end-of-life discussions with patients with cancer and provide them with targeted resources specific to their stage of the cancer journey [22,24,35,36,41,56].

Discussion

Principal Findings

This review demonstrates that empathic AI is emerging as a powerful and transformative force within oncology, with broad clinical applications from education to diagnostics and treatment planning to emotional support. As such, empathic AI has the potential to enhance not only the efficacy and accuracy of care but also its psychosocial dimensions. We found that the majority of studies focused on patient support, with 85% (38/44) addressing tailored patient education and 91% (40/44) addressing patient emotional and psychosocial support. While less frequently represented, a subset of studies also explored applications in diagnostics and decision-making support (19/44, 43%) and palliative and end-of-life care (13/44, 30%), highlighting additional areas where empathic AI is beginning to gain traction. Across domains, evidence illustrates that AI systems are increasingly being used to address long-standing gaps in cancer care delivery. These technologies are reconfiguring traditional models of care into more adaptive, responsive, and patient-centered frameworks.

It is important to distinguish between human empathy and algorithmic or “simulated” empathy. While human empathy arises from lived experience, relational context, and moral judgment, empathic AI operates through pattern recognition, affective cue detection, and response optimization [58]. These systems do not “feel” empathy but rather approximate empathic behaviors through language, timing, and contextual sensitivity. This positions empathic AI to be understood not as a substitute for human connection but as a decision support and interaction support that may enhance communication, accessibility, and emotional scaffolding within clinician-led care.

A consistent theme throughout the examined literature is the redefinition of what it means to deliver holistic care. Empathic AI enables care that integrates clinical precision with personalized emotional support, bridging a gap that has long challenged providers operating under time and resource constraints. This convergence is particularly critical in oncology, where patients face complex treatment pathways intertwined with profound emotional and existential burdens. Across studies, LLMs and chatbot technologies were used to adapt educational content to user-specific contexts, including cancer type, treatment stage, emotional state, and cultural or linguistic preferences.

To bring these capabilities into everyday practice, clinicians can adopt empathic AI in targeted, workflow-compatible ways. For example, LLMs can be embedded into electronic health record systems to generate on-demand phrasing suggestions for emotionally sensitive conversations, such as treatment failure or end-of-life planning. Moreover, AI cofacilitators can personalize group therapy or education sessions in real time by responding to emotional cues. Additionally, sentiment analyses or AI-generated chat summaries can be used to brief providers ahead of visits, helping them tailor communication strategies to patient needs. AI’s ability to contextualize medical information, detect distress, and respond with calibrated empathy suggests that these systems can complement and bolster human care, particularly in emotionally charged settings such as diagnosis disclosure, palliative care, and end-of-life decision-making.

Evidence from this review further suggests that empathic performance in digital oncology settings is not solely a function of model sophistication but of interactional design and workflow integration. Studies measuring cancer-related live chat environments and simulation-based systems highlight that digital clinical empathy requires intentional strategies (ie, timely acknowledgment, authenticity, appropriate boundary setting, and contextual responsiveness) and may be best sustained through hybrid models that combine AI-enabled support with human oversight [34-36].

Moreover, the high levels of user satisfaction reported in interventions such as Vivibot [14], Vik [17], Rosa [18], and social media–based studies [19,32,57] underscore that patients are receptive to receiving education, treatment planning, and emotional support from AI systems. These findings raise important implications for the clinical role of AI in extending the reach and impact of human oncology care, rather than simply imitating it. Rather than merely mimicking human interaction,

XSL•FO
RenderX

empathic AI can help scale patient-centered care by offering timely, emotionally calibrated support, especially in scenarios where time constraints or emotional fatigue may limit provider capacity. Notably, findings from this review suggest that patients may sometimes prefer AI-generated responses over those from physicians when it comes to clarity, empathy, and accessibility [19,29,32]. Such findings challenge traditional assumptions about the limitations of machine-delivered care and suggest a potential reimagining of the therapeutic relationship. In oncology, clarity, emotional tone, and trust are particularly sensitive. For clinicians, this emerging preference signals an opportunity to align care delivery models with patient communication expectations. AI can serve as a buffer, or even a bridge, between emotionally overwhelmed patients and time-constrained providers. Empathic AI-mediated tools may help close gaps in emotional availability, ensure message consistency, and reduce the interpersonal strain that often accompanies sensitive care discussions.

Empathic AI also holds potential to reshape the patient-to-provider relationship. For patients, AI expands access to personalized education and emotional support, democratizing information and empowering users to take an active role in their treatment. This is especially significant for historically underserved populations or individuals with low health literacy, who may benefit most from scalable, emotionally attuned technologies. However, emerging evidence indicates that patient engagement with empathic AI tools is strongly shaped by trust, digital literacy, perceived usefulness, privacy concerns, and the extent to which AI systems are meaningfully connected to clinicians and care pathways. This highlights implementation success dependent as much on sociotechnical context as on model performance [25].

From a provider perspective, AI serves as both a diagnostic and treatment decision support system and a partner in empathic communication. AI-generated conversation summaries, real-time feedback, and emotion-sensitive prompts can enhance provider-patient dialogues and reduce burnout by offloading cognitive and emotional labor, such as the analysis of patient decision-making and emotions provided by the PRIME-2 (Patient Reported Information Multidimensional Exploration version 2) framework [46]. However, these benefits come with caution. The risk of overreliance on AI-generated outputs and the potential erosion of clinicians' empathic instincts raise critical concerns about long-term implications for professional development and patient trust.

Despite the promise of empathic AI, several challenges must be addressed before these systems can be safely and effectively deployed in clinical settings. The question of algorithmic empathy remains unresolved. Evidence suggests that generative AI responses to cancer-related decision-making scenarios may vary by patient demographics and clinical framing, with inconsistencies in tone, content, and guidance that could differentially influence patient perceptions and treatment decisions, underscoring the need for careful validation and health equity–focused oversight [22,24]. While AI can simulate or mimic affective responses and detect emotional cues, true empathic understanding rooted in shared experience remains uniquely human. As such, AI should be positioned as a decision support tool, not as a replacement for traditional, physician-led care. Overreliance on AI in emotionally sensitive encounters could lead to depersonalization or erosion of therapeutic rapport between patient and provider, particularly if patients perceive machine-generated responses as inauthentic or misaligned with their values.

Models trained on nonrepresentative data may also risk reinforcing racial, cultural, and linguistic biases, leading to emotionally inappropriate or even harmful interactions. In oncology, where communication must be attuned to cultural norms, family dynamics, and existential concerns, these missteps could carry particularly high stakes. Specifically, loss of trust in providers and institutions is detrimental to cancer care, as trust remains the foundation of health care in sensitive and difficult situations and among vulnerable populations. Safety is further complicated by the risk of misinformation or emotionally manipulative outputs. Moreover, sociocultural dynamics shape both the reception and use of empathic AI. Concepts of empathy, emotional expression, and caregiving may vary considerably across cultures; what is perceived as compassionate in one context may be viewed as intrusive or inappropriate in another. Developers must engage patients, caregivers, and clinicians from diverse backgrounds in early design to ensure cultural alignment and application for all populations.

Ensuring that AI systems provide reliable, culturally sensitive, and context-appropriate support will require robust algorithmic guardrails, clinician oversight, and rigorous validation. Without these safeguards, empathic AI may inadvertently exacerbate existing health disparities among marginalized populations already facing barriers to emotionally supportive care.

The findings of this review underscore the value of hybrid clinical care models, wherein humans and machines complement each other, termed human-AI symbiosis [39]. While empathic AI systems show promise in extending the reach of clinicians, personalizing education, and offering emotional scaffolding for patients and caregivers, their successful integration into clinical workflows remains dependent on several critical factors. Effective implementation will depend on careful design, robust training data, and a commitment to maintaining the human essence of medicine. Given the rapid pace of model development and evolving user expectations, adoption should proceed cautiously, with a focus on adaptability and long-term clinical usefulness. Moving forward, research must explore the longitudinal impacts of empathic AI interventions, develop integration frameworks grounded in ethical and cultural humility, and assess performance across diverse populations and care settings. Interdisciplinary collaboration will be essential to ensure that future systems are not only intelligent and accurate but also trustworthy, equitable, compassionate, and responsive to real-world clinical needs. To support practical interpretation, Table 2 summarizes key challenges associated with empathic AI in oncology alongside potential mitigation strategies identified across the included literature.

XSL•FO
RenderX

Table . Challenges, risks, and mitigation strategies for empathic artificial intelligence in cancer care.

Challenge	Description	Potential mitigation strategies
Algorithmic empathy versus human empathy	Empathic AI[a] systems simulate empathic behaviors through pattern recognition and affective cue detection but lack lived experience, moral reasoning, and relational accountability. This raises concerns about authenticity, depth, and appropriateness of emotional responses in sensitive oncology contexts.	Position empathic AI as decision support and interaction support rather than replacement for human care; embed clinician oversight and escalation pathways; and explicitly communicate AI's role and limitations to patients to preserve trust and relational integrity.
Reliability, trust, and interpretive risk	Empathic AI-generated responses may vary in tone, emotional sensitivity, or guidance depending on prompt structure, clinical framing, or patient demographics, potentially shaping patient perceptions, fostering over trust or skepticism, and influencing clinical decision-making.	Implement algorithmic guardrails and standardized response calibration; conduct bias and consistency audits across subgroups; intentionally design user interfaces to reinforce clinician accountability; use scenario-based validation prior to deployment in emotionally sensitive use cases; and provide clinician training on appropriate AI use.
Misinformation risk	Generative AI systems may produce fluent, confident, but clinically incorrect or incomplete responses, which can be particularly harmful in oncology decision-making and end-of-life contexts.	Maintain human-in-the-loop validation for high-stakes outputs; restrict autonomous response generation in diagnostic and treatment contexts; and integrate citation requirements, uncertainty signaling, and clinician review mechanisms.
Equity, bias, and cultural alignment	Training data may encode structural inequities or fail to account for cultural variation in emotional expression, caregiving norms, and communication styles, resulting in emotionally inappropriate or inequitable interactions.	Use participatory and co-design approaches with diverse patients, caregivers, and clinicians; prioritize inclusive datasets; localize language, tone, and interaction styles; and conduct ongoing health equity and subgroup performance audits.
Privacy, data governance, and emotional data sensitivity	Empathic AI systems often process highly sensitive emotional, psychosocial, and behavioral data, increasing risks related to privacy, consent, and secondary data use.	Establish robust data governance frameworks; apply privacy-by-design principles; ensure secure infrastructure and role-based access controls; and clarify consent processes for emotional and conversational data use.
Workflow integration and clinician burden	Poorly integrated AI tools may increase cognitive load, disrupt clinical workflows, or generate alert fatigue, undermining adoption and trust.	Design workflow-compatible tools aligned with clinical routines; pilot implementations in real-world settings; and iteratively refine systems based on clinician feedback and usability testing.

[a]AI: artificial intelligence.

Strengths and Limitations

This review offers a comprehensive analysis of empathic AI in cancer care, yet several limitations should be noted. First, this review included studies available only in English language, which may limit the generalizability of findings to global health care contexts. Cultural variations in emotional expression, communication styles, and attitudes toward AI are likely to influence both the development and reception of empathic AI systems. Furthermore, although a snowballing technique was used to enhance inclusivity, the reliance on peer-reviewed literature indexed in PubMed may have excluded relevant studies published in interdisciplinary journals, conference proceedings, or gray literature, particularly those from fields such as human-computer interaction or behavioral science.

Another limitation relates to the rapidly evolving nature of AI technologies. Given the pace of advancement in the field, some of the models and interventions included in this review may already be outdated or replaced by more sophisticated systems. While this pace of innovation reflects momentum in the field, it also complicates assessments of clinical readiness. Current tools may not yet meet the usability, safety, or integration standards required for large-scale adoption, underscoring the need for ongoing evaluation in real-world settings.

However, this fast-paced innovation can also be viewed as a strength, as the continual refinement of AI models suggests that future iterations will likely be more accurate, more empathetic, and better integrated into clinical workflows. By capturing a broad timeline of development from 2018 to early 2025, this review not only reflects the field's foundational growth but also establishes a baseline against which future progress can be measured.

It is also important to note that included studies were methodologically heterogeneous, encompassing diverse designs, populations, interventions, and outcome measures. The absence of standardized metrics for evaluating empathic performance limits direct comparison across studies, although this heterogeneity reflects the early and exploratory nature of the field.

Additionally, there was no standardized metric for evaluating "empathy" in AI-generated interactions across studies. This inconsistency complicates efforts to compare effectiveness across interventions. Some studies used patient satisfaction or

XSL•FO
RenderX

emotional resonance as proxies, while others relied on subjective expert evaluation. At the same time, emerging work in sentiment modeling and empathic conversational agent development suggests early methodological movement toward quantifying affective signals, although approaches remain heterogeneous and not yet clinically standardized [49,53,54]. Future research should focus on a shared operational definition of empathic AI performance to establish clear benchmarks for clinical implementation.

Public Health Implications

With careful design and oversight, empathic AI has the potential to advance public health by expanding access to emotionally responsive care and addressing structural gaps in patient-provider communication, trust, and psychosocial support. These technologies could help bridge long-standing inequities by reaching historically underserved populations facing linguistic, geographic, or cultural barriers to care. In resource-limited settings, empathic AI may also reduce clinician burden, promote continuity of care, and strengthen system resilience by extending the reach of essential health services.

In oncology, empathic AI platforms could deliver timely, culturally sensitive, and emotionally attuned support to help mitigate disparities in diagnosis, treatment adherence, and survivorship. By enhancing provider communication and tailoring care to diverse patient needs, these systems can counteract the effects of systemic bias and institutional mistrust. Integration into care settings may also enable early detection of patient distress and improve referral pathways to behavioral health and palliative services, offering a scalable means of bridging mental and physical care.

Implementation of empathic AI in low-resource or global health contexts may offer affordable, scalable strategies to support patient-centered cancer care where human capital is limited. As digital health infrastructure continues to evolve, empathic AI should be recognized not only as a clinical innovation but as a system-level tool for strengthening communication-centered models of care that foster patient engagement, empowerment, and trust. By capturing data on emotional needs, communication gaps, and engagement patterns, these tools may also inform more responsive, equity-driven public health surveillance and planning.

Conclusions

Empathic AI holds immense promise to revolutionize cancer care by enabling a more patient-centered, holistic, and emotionally sensitive approach. Oncology presents distinct emotional and ethical challenges, making it a uniquely high-stakes setting for the application of empathic AI. With thoughtful development and implementation, these tools can help close communication gaps, improve diagnostics and treatment decision-making, and support the emotional well-being of both patients and caregivers.

The creation of empathic AI–driven portals for patients and providers, for example, can improve engagement and satisfaction, expand access to culturally tailored education, and personalize treatment planning. These tools can and should be applied across the oncologic care continuum from identifying at-risk patients, to supporting treatment navigation and recovery, to helping patients and caregivers cope with the emotional burden of a cancer diagnosis. While this review highlights early promise and high user receptivity, clinical integration must be guided by rigorous validation, ethical oversight, and a deep understanding of real-world workflows and health system readiness.

Empathic AI should be viewed as an adjunct to human-centered care, not as a replacement for current practices and procedures. Clinicians must lead its implementation to ensure that empathy, trust, and clinical excellence remain central to oncologic practice. To support responsible adoption, health care systems should proactively invest in the infrastructure required to evaluate and scale these tools. Modular pilot programs such as leveraging LLMs to coauthor patient education materials, deploying chatbot cofacilitators in support groups, or using NLP tools to identify patients in emotional distress can help pave the way to widespread implementation. Other practical applications include automated phrasing assistance in clinical documentation, simulation-based training for difficult conversations, and tools to detect and mitigate caregiver burnout.

This review highlights the distinct value and complexity of applying empathic AI in oncology, where sensitivities differ markedly from other clinical contexts. The findings underscore the need for an oncology-specific framework that identifies key emotional touchpoints, proposes clinical use cases, and integrates safeguards to mitigate ethical and psychosocial risks. With continued advancement, empathic AI can help foster a health care environment that is both emotionally attuned and operationally efficient—particularly in cancer care, where emotional resonance and communication quality are as vital as clinical accuracy. The path forward requires rigorous validation, interdisciplinary collaboration, and an unwavering commitment to equity and patient voice. As we move toward the next generation of cancer care, the thoughtful integration of empathic AI offers an unprecedented opportunity to reimagine how we deliver care.

Acknowledgments

Generative artificial intelligence tools were used under direct author supervision to assist with structural editing and alignment with journal formatting guidelines. All substantive content, data interpretation, analytic decisions, and final wording are the original work of the authors.

Funding

This study was supported by a grant from the State of Tennessee Department of Health.

XSL•FO
RenderX

Data Availability

Data sharing is not applicable to this paper as no data sets were generated or analyzed during this study.

Authors' Contributions

BMW conceptualized the study and drafted, reviewed, and edited the manuscript. GLA conceptualized the study and drafted, reviewed, and edited the manuscript. JAZ conceptualized the study and drafted, reviewed, and edited the manuscript. PKR, FAK, SH, SK, TG, RR, CLB, and RLD critically reviewed and edited the manuscript. DLS reviewed and edited the manuscript, supervised the study, and acquired funding. AS-N drafted, reviewed, and edited the manuscript; supervised the study; and acquired funding.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Multimedia Appendix 1
Summary table representing the characteristics of selected studies (revision).
[DOCX File, 80 KB - cancer_v12i1e82336_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 100 KB - cancer_v12i1e82336_app2.pdf]

References

1. Khosravi M, Zare Z, Mojtabaeian SM, Izadi R. Artificial intelligence and decision-making in healthcare: a thematic analysis of a systematic review of reviews. Health Serv Res Manag Epidemiol 2024;11:23333928241234863. [doi: 10.1177/23333928241234863] [Medline: 38449840]
2. Shaban-Nejad A, Michalowski M, Buckeridge DL. Health intelligence: how artificial intelligence transforms population and personalized health. NPJ Digit Med 2018;1:53. [doi: 10.1038/s41746-018-0058-9] [Medline: 31304332]
3. Ammar N, Bailey JE, Davis RL, Shaban-Nejad A. Using a personal health library-enabled mHealth recommender system for self-management of diabetes among underserved populations: use case for knowledge graphs and linked data. JMIR Form Res 2021 Mar 16;5(3):e24738. [doi: 10.2196/24738] [Medline: 33724197]
4. Shaban-Nejad A, Michalowski M, Bianco S. Multimodal AI in Healthcare: A Paradigm Shift in Health Intelligence: Springer; 2023. [doi: 10.1007/978-3-031-14771-5]
5. Kolla L, Parikh RB. Uses and limitations of artificial intelligence for oncology. Cancer 2024 Jun 15;130(12):2101-2107. [doi: 10.1002/cncr.35307] [Medline: 38554271]
6. Cè M, Chiarpenello V, Bubba A, et al. Exploring the role of ChatGPT in oncology: providing information and support for cancer patients. BioMedInformatics 2024;4(2):877-888. [doi: 10.3390/biomedinformatics4020049]
7. Li Y, Gao W, Luan Z, Zhou Z, Li J. The impact of Chat Generative Pre-trained Transformer (ChatGPT) on oncology: application, expectations, and future prospects. Cureus 2023;15(11):e48670. [doi: 10.7759/cureus.48670]
8. Goel I, Bhaskar Y, Kumar N, et al. Role of AI in empowering and redefining the oncology care landscape: perspective from a developing nation. Front Digit Health 2025;7:1550407. [doi: 10.3389/fdgth.2025.1550407] [Medline: 40103737]
9. Debelee TG, Kebede SR, Schwenker F, Shewarega ZM. Deep learning in selected cancers' image analysis—a survey. J Imaging 2020;6(11):121. [doi: 10.3390/jimaging6110121]
10. Auf H, Svedberg P, Nygren J, Nair M, Lundgren LE. The use of AI in mental health services to support decision-making: scoping review. J Med Internet Res 2025 Jan 24;27:e63548. [doi: 10.2196/63548] [Medline: 39854710]
11. Shaban-Nejad A, Michalowski M. Precision Health and Medicine: A Digital Revolution in Healthcare: Springer; 2020. [doi: 10.1007/978-3-030-24409-5]
12. Farhud DD, Zokaei S. Ethical issues of artificial intelligence in medicine and healthcare. Iran J Public Health 2021;50(11). [doi: 10.18502/ijph.v50i11.7600]
13. Rubin M, Arnon H, Huppert JD, Perry A. Considering the role of human empathy in AI-driven therapy. JMIR Ment Health 2024 Jun 11;11:e56529. [doi: 10.2196/56529] [Medline: 38861302]
14. Greer S, Ramo D, Chang YJ, Fu M, Moskowitz J, Haritatos J. Use of the chatbot "Vivibot" to deliver positive psychology skills and promote well-being among young people after cancer treatment: randomized controlled feasibility trial. JMIR Mhealth Uhealth 2019 Oct 31;7(10):e15018. [doi: 10.2196/15018] [Medline: 31674920]
15. Leung YW, Wouterloot E, Adikari A, et al. Artificial intelligence-based co-facilitator (AICF) for detecting and monitoring group cohesion outcomes in web-based cancer support groups: single-arm trial study. JMIR Cancer 2024 Jul 22;10:e43070. [doi: 10.2196/43070] [Medline: 39037754]

XSL•FO
RenderX

16. Leung YW, Ng S, Duan L, et al. Therapist feedback and implications on adoption of an artificial intelligence-based co-facilitator for online cancer support groups: mixed methods single-arm usability study. JMIR Cancer 2023 Jun 9;9:e40113. [doi: 10.2196/40113] [Medline: 37294610]
17. Chaix B, Bibault JE, Pienkowski A, et al. When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot. JMIR Cancer 2019 May 2;5(1):e12856. [doi: 10.2196/12856] [Medline: 31045505]
18. Siglen E, Vetti HH, Lunde ABF, et al. Ask Rosa—the making of a digital genetic conversation tool, a chatbot, about hereditary breast and ovarian cancer. Patient Educ Couns 2022 Jun;105(6):1488-1494. [doi: 10.1016/j.pec.2021.09.027] [Medline: 34649750]
19. Chen D, Parsa R, Hope A, et al. Physician and artificial intelligence chatbot responses to cancer questions from social media. JAMA Oncol 2024 Jul 1;10(7):956-960. [doi: 10.1001/jamaoncol.2024.0836] [Medline: 38753317]
20. Morrow E, Zidaru T, Ross F, et al. Artificial intelligence technologies and compassion in healthcare: a systematic scoping review. Front Psychol 2023;13:971044. [doi: 10.3389/fpsyg.2022.971044] [Medline: 36733854]
21. McDarby M, Mroz EL, Hahne J, Malling CD, Carpenter BD, Parker PA. “Hospice care could be a compassionate choice”: ChatGPT responses to questions about decision making in advanced cancer. J Palliat Med 2024 Dec;27(12):1618-1624. [doi: 10.1089/jpm.2024.0256] [Medline: 39263979]
22. Jo MH, Kim MJ, Oh HK, et al. Communicative competence of generative artificial intelligence in responding to patient queries about colorectal cancer surgery. Int J Colorectal Dis 2024 Jun 20;39(1):94. [doi: 10.1007/s00384-024-04670-3] [Medline: 38902500]
23. Clerici CA, Bernasconi A, Lasalvia P, et al. Being diagnosed with a rhabdomyosarcoma in the era of artificial intelligence: whom can we trust? Pediatr Blood Cancer 2024 Nov;71(11):e31256. [doi: 10.1002/pbc.31256] [Medline: 39129151]
24. Alanezi F. Factors influencing patients’ engagement with ChatGPT for accessing health-related information. Crit Public Health 2024 Dec 31;34(1):1-20. [doi: 10.1080/09581596.2024.2348164]
25. Xu L, Sanders L, Li K, Chow JCL. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR Cancer 2021 Nov 29;7(4):e27850. [doi: 10.2196/27850] [Medline: 34847056]
26. Goumas G, Dardavesis TI, Syrigos K, Syrigos N, Simou E. Chatbots in cancer applications, advantages and disadvantages: all that glitters is not gold. J Pers Med 2024 Aug 19;14(8):877. [doi: 10.3390/jpm14080877] [Medline: 39202068]
27. Turner JH. Triangle of Trust in Cancer Care? The physician, the patient, and artificial intelligence chatbot. Cancer Biother Radiopharm 2023 Nov;38(9):581-584. [doi: 10.1089/cbr.2023.0112] [Medline: 37707991]
28. Alanzi TM, Alzahrani W, Albalawi NS, et al. Public awareness of obesity as a risk factor for cancer in central Saudi Arabia: feasibility of ChatGPT as an educational intervention. Cureus 2023 Dec;15(12):e50781. [doi: 10.7759/cureus.50781] [Medline: 38239542]
29. Koranteng E, Lewis ACF, Abel GA. Compassionate machines: the ethics of “artificial empathy” in cancer care. JAMA Oncol 2024 Jul 1;10(7):857-858. [doi: 10.1001/jamaoncol.2024.0824] [Medline: 38696207]
30. Sezgin E, Jackson DI, Kocaballi AB, et al. Can large language models aid caregivers of pediatric cancer patients in information seeking? A cross-sectional investigation. Cancer Med 2025 Jan;14(1):e70554. [doi: 10.1002/cam4.70554] [Medline: 39776222]
31. Leung YW, Wouterloot E, Adikari A, et al. Natural language processing-based virtual cofacilitator for online cancer support groups: protocol for an algorithm development and validation study. JMIR Res Protoc 2021 Jan 7;10(1):e21453. [doi: 10.2196/21453] [Medline: 33410754]
32. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023 Jun 1;183(6):589-596. [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
33. De Silva D, Ranasinghe W, Bandaragoda T, et al. Machine learning to support social media empowered patients in cancer care and cancer treatment decisions. PLoS One 2018;13(10):e0205855. [doi: 10.1371/journal.pone.0205855] [Medline: 30335805]
34. Luetke Lanfer H, Reifegerste D, Weber W, et al. Digital clinical empathy in a live chat: multiple findings from a formative qualitative study and usability tests. BMC Health Serv Res 2024 Mar 8;24(1):314. [doi: 10.1186/s12913-024-10785-8] [Medline: 38459522]
35. Kane B, Giugno C, Schubert L, Haut K, Wohn C, Hoque E. Managing emotional dialogue for a virtual cancer patient: a schema-guided approach. IEEE Trans Affective Comput 2024;15(3):1041-1052. [doi: 10.1109/TAFFC.2023.3317805]
36. Ali MR, Sen T, Kane B, et al. Novel computational linguistic measures, dialogue system and the development of SOPHIE: standardized online patient for healthcare interaction education. IEEE Trans Affective Comput 2023;14(1):223-235. [doi: 10.1109/TAFFC.2021.3054717]
37. Goisauf M, Cano Abadía M. Ethics of AI in radiology: a review of ethical and societal implications. Front Big Data 2022;5:850383. [doi: 10.3389/fdata.2022.850383] [Medline: 35910490]
38. Mundinger A, Mundinger C. Artificial intelligence in senology—where do we stand and what are the future horizons? Eur J Breast Health 2024 Apr;20(2):73-80. [doi: 10.4274/ejbh.galenos.2024.2023-12-13] [Medline: 38571686]
39. Turner JH. Human-artificial intelligence symbiotic reporting for theranostic cancer care. Cancer Biother Radiopharm 2025 Mar;40(2):89-95. [doi: 10.1089/cbr.2024.0216] [Medline: 39501808]

XSL•FO
RenderX

40. Rogasch JMM, Metzger G, Preisler M, et al. ChatGPT: can you prepare my patients for [18F]FDG PET/CT and explain my reports? J Nucl Med 2023 Dec 1;64(12):1876-1879. [doi: 10.2967/jnumed.123.266114] [Medline: 37709536]
41. Gliwska E, Barańska K, Maćkowska S, Różańska A, Sobol A, Spinczyk D. The use of natural language processing for computer-aided diagnostics and monitoring of body image perception in patients with cancers. Cancers (Basel) 2023 Nov 16;15(22):5437. [doi: 10.3390/cancers15225437] [Medline: 38001696]
42. Kharko A, McMillan B, Hagström J, et al. Generative artificial intelligence writing open notes: a mixed methods assessment of the functionality of GPT 3.5 and GPT 4.0. Digit Health 2024;10:20552076241291384. [doi: 10.1177/20552076241291384] [Medline: 39493632]
43. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ 2023 Sep 22;23(1):689. [doi: 10.1186/s12909-023-04698-z] [Medline: 37740191]
44. Mårell-Olsson E, Mejtoft T, Tovedal S, Söderström U. Opportunities and challenges of using socially intelligent agents: increasing interaction and school participation for children suffering from a long-term illness. IJILT 2021 Aug 12;38(4):393-411. [doi: 10.1108/IJILT-11-2020-0199]
45. Chow JCL, Li K. Ethical considerations in human-centered AI: advancing oncology chatbots through large language models. JMIR Bioinform Biotechnol 2024 Nov 6;5:e64406. [doi: 10.2196/64406] [Medline: 39321336]
46. Ranasinghe W, de Silva D, Bandaragoda T, et al. Robotic-assisted vs. open radical prostatectomy: a machine learning framework for intelligent analysis of patient-reported outcomes from online cancer support groups. Urol Oncol 2018 Dec;36(12):529. [doi: 10.1016/j.urolonc.2018.08.012] [Medline: 30236854]
47. Spiegel BMR, Liran O, Clark A, et al. Feasibility of combining spatial computing and AI for mental health support in anxiety and depression. NPJ Digit Med 2024;7(1):22. [doi: 10.1038/s41746-024-01011-0]
48. Mao W, Chen W, Wang Y. Effect of virtual reality-based mindfulness training model on anxiety, depression, and cancer-related fatigue in ovarian cancer patients during chemotherapy. Technol Health Care 2024;32(2):1135-1148. [doi: 10.3233/THC-230735] [Medline: 37781832]
49. Hamza-Cherif S, Fatiha Kazi Tani L, Settouti N. Improving healthcare communication: AI-driven emotion classification in imbalanced patient text data with explainable models. Adv Technol Innov 2024;9(2):129-142. [doi: 10.46604/aiti.2024.13523]
50. Leung YW, Park B, Heo R, et al. Providing care beyond therapy sessions with a natural language processing-based recommender system that identifies cancer patients who experience psychosocial challenges and provides self-care support: pilot study. JMIR Cancer 2022 Jul 29;8(3):e35893. [doi: 10.2196/35893] [Medline: 35904877]
51. O'Gara G, Murray L, Georgopoulou S, et al. SafeSpace: what is the feasibility and acceptability of a codesigned virtual reality intervention, incorporating compassionate mind training, to support people undergoing cancer treatment in a clinical setting? BMJ Open 2022 Feb 10;12(2):e047626. [doi: 10.1136/bmjopen-2020-047626] [Medline: 35144943]
52. Wei GX, Zhou YW, Li ZP, Qiu M. Application of artificial intelligence in the diagnosis, treatment, and recurrence prediction of peritoneal carcinomatosis. Heliyon 2024 Apr 15;10(7):e29249. [doi: 10.1016/j.heliyon.2024.e29249] [Medline: 38601686]
53. Adikari A, de Silva D, Moraliyage H, et al. Empathic conversational agents for real-time monitoring and co-facilitation of patient-centered healthcare. Future Generation Comput Syst 2022 Jan;126:318-329. [doi: 10.1016/j.future.2021.08.015]
54. Adikari A, de Silva D, Ranasinghe WKB, et al. Can online support groups address psychological morbidity of cancer patients? An artificial intelligence based investigation of prostate cancer trajectories. PLoS One 2020;15(3):e0229361. [doi: 10.1371/journal.pone.0229361] [Medline: 32130256]
55. Bandaragoda T, Ranasinghe W, Adikari A, et al. The Patient-Reported Information Multidimensional Exploration (PRIME) framework for investigating emotions and other factors of prostate cancer patients with low intermediate risk based on online cancer support group discussions. Ann Surg Oncol 2018 Jun;25(6):1737-1745. [doi: 10.1245/s10434-018-6372-2]
56. Niki K, Okamoto Y, Maeda I, et al. A novel palliative care approach using virtual reality for improving various symptoms of terminal cancer patients: a preliminary prospective, multicenter study. J Palliat Med 2019 Jun;22(6):702-707. [doi: 10.1089/jpm.2018.0527]
57. Klotz R, Pausch TM, Kaiser J, et al. ChatGPT vs. surgeons on pancreatic cancer queries: accuracy & empathy evaluated by patients and experts. HPB (Oxford) 2025 Mar;27(3):311-317. [doi: 10.1016/j.hpb.2024.11.012] [Medline: 39672696]
58. Ajeesh KG, Joseph J. The compassion illusion: can artificial empathy ever be emotionally authentic? Front Psychol 2025;16:1723149. [doi: 10.3389/fpsyg.2025.1723149]

Abbreviations

AI: artificial intelligence
LLM: large language model
ML: machine learning
NLP: natural language processing
PRIME-2: Patient Reported Information Multidimensional Exploration version 2
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
VR: virtual reality

XSL•FO
RenderX

Edited by M Balcarras; submitted 13.Aug.2025; peer-reviewed by AG Abejas, A Kapila; revised version received 01.Feb.2026; accepted 06.Feb.2026; published 09.Apr.2026.

Please cite as:
White BM, Aitken GL, Zink JA, Kheirkhah Rahimabad P, Kumsa FA, Hashtarkhani S, Rashid R, Kheirinejad S, Girdwood T, Brett CL, Davis RL, Schwartz DL, Shaban-Nejad A
Empathic AI for Patient-Centered Cancer Care: A Scoping Review of Patient Navigation, Support, and Clinical Practice
JMIR Cancer 2026;12:e82336
URL: https://cancer.jmir.org/2026/1/e82336
doi:10.2196/82336

© Brianna M White, Gabriela L Aitken, Janet A Zink, Parnian Kheirkhah Rahimabad, Fekede A Kumsa, Soheil Hashtarkhani, Rezaur Rashid, Saba Kheirinejad, Tyra Girdwood, Christopher L Brett, Robert L Davis, David L Schwartz, Arash Shaban-Nejad. Originally published in JMIR Cancer (https://cancer.jmir.org), 9.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Application of eHealth in Symptom Management of Patients With Breast Cancer During Endocrine Therapy: Scoping Literature Review

Chengjia Gu[1], BSN; Lili Shen[1], BSN; Shiyu Zhang[2], BSN; Nan Zhang[1], MSN

[1]Comprehensive Breast Health Centre, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Rui Jin Er Road, Shanghai, China
[2]Department of Nursing, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Corresponding Author:
Nan Zhang, MSN
Comprehensive Breast Health Centre, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Rui Jin Er Road, Shanghai, China

Abstract

Background: Endocrine therapy is a fundamental treatment for hormone receptor–positive breast cancer. However, treatment-related symptoms may lead to poor adherence and premature discontinuation, adversely affecting clinical outcomes. eHealth technology interventions have been widely applied as potential tools to support patient education, symptom management, and adherence.

Objective: This review aimed to evaluate the application, effectiveness, and limitations of eHealth technologies in symptom management among patients with breast cancer undergoing endocrine therapy.

Methods: A systematic literature search was conducted across 9 databases, including MEDLINE Complete, Academic Search Complete, PubMed, and CINAHL Ultimate, covering studies published between February 2015 and February 2025. Keywords included “breast cancer,” “endocrine therapy,” and “ehealth,” along with synonyms and related terms. Eligible studies examined the use of eHealth interventions in patients with breast cancer undergoing endocrine therapy. From 1352 records, 30 (2.2%) studies were included after screening and full-text review.

Results: The 30 included studies comprised 13 (43.3%) randomized controlled trials, 7 (23.3%) single-arm studies, 3 (10%) retrospective analyses, 2 (6.7%) observational studies, 3 (10%) mixed methods studies, and 2 (6.7%) qualitative studies. eHealth interventions, primarily mobile apps, telemonitoring systems, and SMS text messaging, demonstrated benefits in short-term symptom management, such as reductions in hot flashes, joint pain, anxiety, and sexual distress and improved medication adherence. However, evidence regarding long-term adherence and quality of life outcomes remains inconsistent. Considerable heterogeneity existed in study designs, adherence measures, and intervention components. Nurses were considered key contributors in implementing and monitoring eHealth interventions.

Conclusions: eHealth interventions show promise in improving symptom management and short-term adherence among patients with breast cancer receiving endocrine therapy. Methodological limitations and variability across studies restrict the strength of conclusions. Future research should prioritize long-term outcomes, standardized measures, and integration into routine clinical practice while incorporating perspectives from multiple stakeholders.

(JMIR Cancer 2026;12:e92030) doi:10.2196/92030

KEYWORDS

eHealth; endocrine therapy; symptom management; Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA; patient with breast cancer

Introduction

Breast cancer is heterogeneous, with approximately 80% of cases classified as hormone receptor positive [1,2]. Adjuvant endocrine therapy (AET), primarily selective estrogenic receptor modulators and aromatase inhibitors, is a standard treatment for patients that significantly reduces the risk of recurrence and mortality [3-5]. Clinical guidelines recommend 5 to 10 years of AET [6-8]. Despite its clinical benefits, adherence to endocrine therapy remains suboptimal. Studies indicate that up to one-third of patients discontinue treatment within 5 years, often due to treatment-related adverse effects, inadequate information, comorbidities, or patient preferences [9,10]. Common adverse effects include hot flashes, joint pain, osteoporosis, sexual dysfunction, weight gain, sleep disturbances, depression, anxiety, and cognitive impairment [6,11]. These symptoms can impair quality of life (QOL) and contribute to nonadherence, ultimately affecting prognosis [12].

XSL•FO
RenderX

Therefore, effective symptom management is critical for improving adherence and clinical outcomes.

The World Health Organization has introduced the Global Strategy on Digital Health, emphasizing the integration of information and communications technologies to improve health outcomes [13]. In this review, eHealth refers to the use of digital tools such as mobile apps, SMS text messaging, telemedicine platforms, web-based portals, wearable devices, and electronic patient-reported outcome systems for real-time symptom monitoring, medication reminders, patient education, automated alerts, and remote self-management support [14]. Over the last 2 decades, eHealth has been increasingly used to support symptom monitoring, patient education, and remote care delivery [15]. Previous research has demonstrated the value of eHealth in chronic disease and cancer management, including symptom tracking, adherence support, and self-management [16,17].

Patients with breast cancer have used eHealth technologies for issues such as pain, sleep disorders, lymphedema, psychological distress, and chemotherapy side effects [18-22]. However, no comprehensive scoping review has synthesized evidence specifically on eHealth applications for symptom management during endocrine therapy. This review addresses that gap by evaluating the current evidence on effectiveness, strengths, limitations, and implementation considerations to inform clinical practice.

Methods

Search Strategy

This review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Checklist 1) to ensure thorough search quality and standardized processes [23,24]. Nine databases were searched: MEDLINE Complete, Academic Search Complete, CINAHL Ultimate, Health Source: Nursing/Academic Edition, APA PsycInfo, PubMed, Scopus, Embase, and Web of Science. The search used combinations of terms for “breast cancer,” “endocrine therapy,” and “eHealth” supplemented with synonyms, academic terms, abbreviations, and related terminology. Boolean operators (“AND” and “OR”) were applied to combine terms for search strategy construction systematically. Limits included English-language, peer-reviewed articles focusing on the application of eHealth in breast cancer and published between February 2015 and February 2025. The full search strategy is detailed in Table 1.

Table . Summary of the search strategy for the literature review.

Domain	Description
Search keywords	
Breast cancer	“Breast Neoplasms” OR “Breast Tumor,” OR “Breast Cancer” OR “Mammary Cancer” OR “Malignant Neoplasm of Breast” OR “Breast Malignant Neoplasm” OR “Malignant Tumor of Breast” OR “Breast Malignant Tumor” OR “Cancer of Breast OR Mammary Carcinoma, Human” OR “Human Mammary Neoplasms”
Endocrine therapy	“endocrine therapy” OR “endocrine treatment” OR “aet” OR “hormone therapy”
eHealth	ehealth OR e-health OR telecare OR telemedicine OR telehealth OR “digital health” OR mHealth OR teletherapy OR “smartphone App” OR “smartphone application” OR “health apps” OR “remote monitoring” OR “mobile app” OR “mobile application” OR “wearable devices” OR telemedicine OR tele-care OR telecare OR “Medicine, Virtua” OR “Tele Referral” OR tele-referrals OR tele-referral OR “Virtual Medicine” OR “Health, Mobile” OR “Mobile Health” OR telehealth OR “Electronic App” OR “Electronic Application” OR “Application, Mobile” OR “mobile phone” OR “smartphone” OR “text-, electronic-“ OR “multimedia-message” OR “electronic mail” OR “phone applications” OR “computer” OR “podcast” OR “videos” OR “internet” OR “website” OR “chatroom” OR “message board” OR “activity tracker” OR “electronic health” OR “mobile health” OR “telemedicine”
Other search parameters	
Search limits	2015-2025
Type of material searched	English-language peer-reviewed scholarly articles
Databases searched	EBSCOhost databases (Academic Search Complete, Health Source: Nursing/Academic Edition, CINAHL Ultimate, MEDLINE Complete, and APA PsycInfo), PubMed, Scopus, Web of Science, and Embase

Eligibility Criteria

Studies were eligible for inclusion if they were empirical research articles that examined the application of eHealth interventions specifically for symptom management or related outcomes, including medication adherence and QOL, among patients with breast cancer undergoing endocrine therapy in either the early or advanced disease stages. Exclusion criteria

XSL•FO
RenderX

included animal studies, research protocols, review articles, conference abstracts, non–English-language publications, and studies that focused exclusively on other cancer treatments rather than endocrine therapy.

Study Selection and Data Extraction

A total of 1352 records were retrieved from the 9 databases, and 148 (10.9%) were removed based on the predefined inclusion and exclusion criteria. The remaining 1204 records were imported into the Zotero software (version 7; Corporation for Digital Scholarship), where 616 (51.2%) duplicates were removed. Among the remaining 588 records, title and abstract screening resulted in the exclusion of 504 (85.7%) studies, primarily due to irrelevant populations (non–breast cancer or non–endocrine therapy; n=295, 58.5%), irrelevant focus (chemotherapy only, screening, or survivorship care without endocrine therapy; n=110, 21.8%), nonempirical content (reviews, protocols, and abstracts; n=59, 11.7%), or focus on other diseases (n=40, 7.9%). Subsequently, of the remaining 84 records, 30 (35.7%) were excluded because the full texts could not be retrieved. Full-text assessment was then performed on the remaining 54 articles, of which 24 (44.4%) were excluded because they described only study protocols or were conducted on survivors of breast cancer without a specific focus on endocrine therapy. Ultimately, 30 articles met all eligibility criteria and were included in this review (see Table S1 in Multimedia Appendix 1 [25-54] for details). Data extracted included study design, participant characteristics, intervention details, outcomes (symptom management, adherence, QOL, and patient experiences), and key findings.

Risk-of-Bias Assessment

Two oncology nurse researchers assessed the risk of bias in randomized controlled trials (RCTs) using the Cochrane risk-of-bias tool independently [55]. Discrepancies, if any, were resolved by a third reviewer. Due to heterogeneity in designs, interventions, and outcome measures, no meta-analysis was performed. The Critical Appraisal Skills Programme checklists were used to evaluate methodological quality across study types [56].

Data Synthesis

Findings were synthesized thematically into three main aspects: (1) effectiveness and limitations of eHealth for symptom management and adherence, (2) the role of nurses, and (3) patient experiences and impact on QOL.

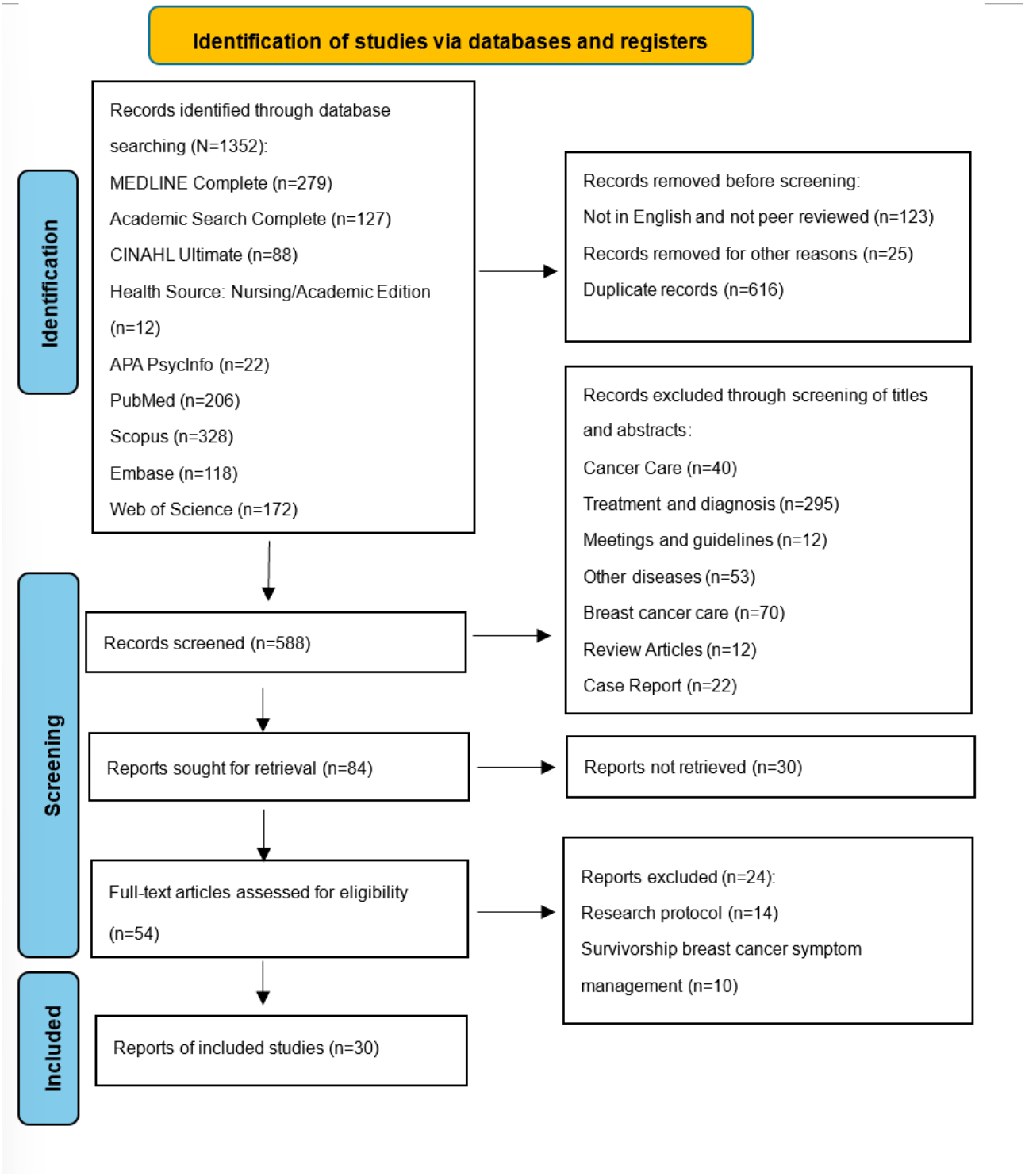
Results

Overview

The 30 selected studies included 13 (43.3%) RCTs, 7 (23.3%) single-arm studies, 3 (10%) retrospective analyses, 2 (6.7%) observational studies, 3 (10%) mixed methods studies, and 2 (6.7%) qualitative studies. Most studies (n=17, 56.7%) were conducted in the United States, followed by Sweden (n=3, 10%); Japan and the United Kingdom (n=2, 6.7% each); and Germany, Turkey, Singapore, South Korea, and China (n=1, 3.3% each). Three key themes emerged from the review: (1) effectiveness and limitations regarding symptom management and adherence to endocrine therapy, (2) nurses' role in eHealth interventions, and (3) experiences of patients with breast cancer with eHealth tools and the impact on QOL. The PRISMA flow diagram is presented in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 Fflow Ddiagram of the literature review.



Effectiveness and Limitations of eHealth in Symptom Management

eHealth interventions, including mobile apps, telemonitoring, and SMS text messaging, facilitated timely self-reporting of symptoms and triggered alerts for health care providers. In total, 10% (3/10) of the studies allowed patients to self-report and record symptoms using mobile apps [29-31], whereas another 13.3% (4/30) of the studies used questionnaires to record patients' symptoms [32-35], allowing for real-time data capture and proactive management.

A total of 20% (6/30) of the studies reported positive effects on symptom control using eHealth [36-41]. One RCT combining an educational website with telephone counseling significantly reduced sexual difficulties, distress, and anxiety compared with usual care [36]. Another RCT using a mobile education program with nurse reminders improved physical well-being, pain, emotional functioning, and endocrine-related symptoms at 3 months [57]. Teleconferencing interventions in 2 RCTs enhanced self-efficacy and reduced hot flashes and depression [41,42]. A longitudinal study of remote tai chi sessions via an app demonstrated reductions in joint pain, anxiety, and hot

XSL•FO
RenderX

flashes after 12 weeks [38]. An app-based intervention in advanced endocrine treatment improved arm symptoms, cognitive function, anxiety, and depression [39]. The mobile app integrated daily exercise logging, medication adherence tracking, Fitbit activity data synchronization, personalized goal setting, educational modules on hormonal therapy side effects, and weekly progress visualizations, with automated feedback and nurse coach messaging. These eHealth tools enabled more accurate and timely symptom tracking, supporting personalized care plans and reducing symptom burden in the short term (typically 1 to 6 months) [29,31,35,42]. However, most studies focused on short-term outcomes and specific subgroups, such as patients with poor AET adherence [41], recent sexual activity [36], or aromatase inhibitor–induced arthralgia [38], limiting the generalizability of the findings. Patients with breast cancer on endocrine therapy often experience complex, interacting symptom clusters [22]; few interventions (1/30, 3.3%) addressed this holistically.

Medication Adherence

Medication adherence is commonly defined as taking medication as prescribed; nonadherence includes delays, dose deviations, or premature discontinuation [58]. Previous studies have shown that eHealth interventions have a significant impact on medication adherence in patients with chronic diseases and those taking medication for mental health conditions [59-61]. Medication adherence is addressed solely in relation to its direct correlation with effective symptom management; premature discontinuation is a subsequent result of inadequately controlled symptoms. Therefore, studies showing improved adherence were interpreted strictly as evidence of effective symptom control rather than as a separate end point.

Retrospective studies indicated 5-year adherence rates of 48% to 53%, primarily due to adverse events and symptoms [43-45]. Access to practical information, active monitoring of side effects, and effective medication reminders have been suggested to improve adherence [62]. A total of 16.7% (5/30) of the studies found that digital health tools, including SMS text messages, mobile apps, and emails, that incorporated symptom education, self-monitoring, and alerts, improved medication adherence [31,32,34,46,47]. Another 16.7% (5/30) of the studies reported benefits from reminder-based educational interventions [30,39,48,49]. However, other studies reported inconsistent results. In total, 15.4% (2/13) of the RCTs with 1- and 3-year follow-ups on SMS text messaging found no significant differences in adherence [33,50]. One retrospective study with telephone follow-up showed no long-term benefit [45], and one RCT on a mobile program reported no improvement despite educational modules [40]. Medication adherence was measured across several aspects: prescription fill rates of more than 80% [32,39], self-reports [31], estrogen level monitoring, and discontinuation rates (Table 2). The heterogeneity in measures and data collection and analysis methods contributed to variability in the findings. Improvements were mainly evident in short-term studies; long-term data remain limited, highlighting the need for sustained interventions.

XSL•FO
RenderX

Table . Summary of medication adherence measures in the literature review.

	Mougalian et al [32]	Hershman et al [33]	Tan et al [50]	Park et al [48]	Takada et al [31]	Smith et al [34]	Wheeler et al [49]	Chan et al [39]	Krukowski et al [47]	Jacobs et al [51]
Prescription fill rate exceeding 80%	✓							✓		
Integrating urine analysis with self-reporting		✓								
Combining self-reporting with innovative bottle technology							✓			
Using estrogenic level measurements in conjunction with a simplified medication adherence report			✓							
Patient self-records					✓					
Discontinuation rates						✓				
Integrating prescription fill data with smart bottle opening data				✓					✓	
Medication adherence reporting scale+smart pill bottle										✓

Nurses' Role in eHealth for Symptom Management

Nurses were the primary implementers in most interventions, leveraging eHealth to enable real-time symptom assessment, 2-way communication, prioritization of high-risk cases, and referrals [63-66]. A total of 13.3% (4/30) of the studies found that nurses enabled 2-way communication and referral by prioritizing high-risk symptoms [32,34,40,47]. Nurses sharing information through eHealth tools could improve both patient satisfaction and the quality of care [52,53]. Nurses delivered education via SMS text messages, app modules, and teleconferences, enhancing patients' self-management skills and symptom control [29,33,37,41,67]. Moreover, the included interventions used dynamic, standardized, and structured assessments of both cognitive and somatic symptoms, which new data indicate are superior to conventional static questionnaires for earlier detection of clinically relevant changes and timely intervention [67].

Nurse involvement improved care efficiency, multidisciplinary collaboration, and resource use [68-70]. In total, 10% (3/30) of the studies, which were nurse led, showed that nurse-led approaches improved high-quality health care and reduced high-cost contacts and outpatient visits [46,47,53]. However, challenges included increased workload, role ambiguity, and the need for process adaptation [69,71]. Most evidence reflects patient perspectives; fewer studies (2/30, 6.7%) examined nurses' or physicians' acceptance and barriers. Cost-effectiveness evaluations were largely absent, underscoring the need for health economic analyses.

XSL•FO
RenderX

Patients' Experiences With eHealth Tools and the Impact on QOL

In this review, patients generally expressed high satisfaction and acceptance of SMS text message reminders [32,50], teleconferencing [37], and mobile apps [29,54], citing ease of use, helpful information, and improved self-management (satisfaction often of >90%). Combined tools such as videoconferencing applications and wearables further enhanced satisfaction, and many patients were willing to continue using them [38,47,62]. A total of 20% (6/30) of the studies examined QOL impacts. Four of these studies reported significant improvements in various domains such as physical, emotional, or endocrine-specific typically after 1 to 6 months of using eHealth tools such as mobile apps with reminders or tele-education [31,39,51,52]. One study providing tele-education showed significant improvements in patients' QOL after a 3-month intervention [52]. Another study used a teleconference education format and demonstrated significant improvement in patients' QOL after 24 weeks of intervention [51]. However, 2 RCTs found no significant differences in QOL: one study implemented a mobile app with SMS text message reminders and follow-up assessments at 6 and 12 months [40], whereas the other integrated the app with a patient e-reporting intervention and assessed QOL at 3 months after the intervention [35].

Heterogeneity and variations in QOL study results likely stem from differences in QOL instruments [31,35,37,39,40] (Table 3). Various QOL measurement scales may yield different results when evaluating due to differences in study design, cultural appropriateness, and assessment dimensions [72-74]. The assessment of QOL in survivors of breast cancer requires a multidimensional approach encompassing physical, psychological, social, and spiritual well-being, indicating that the use of a single assessment tool may not fully capture the heterogeneity in their QOL trajectories [75]. This review found that eHealth interventions effectively improved patients' QOL in the short term. However, the long-term effects and trends for patients with breast cancer undergoing endocrine treatment remain to be thoroughly researched and explored.

Table . Summary of questionnaires assessing quality of life (QOL) in the literature review.

Study	FACT-ES QLS[a]	EORTC QLQ-C30[b]	EORTC QLQ-BR23[c]	FACT-B[d]	12-Item Short Form Health Survey	Validity and reliability	Measurement results
Çınar et al [52], 2021	✓					✓	There was a statistically significant difference in the total FACT-ES QLS scores (P<.05).
Jacobs et al [51], 2022				✓		✓	There was a significant improvement in QOL in the intervention group at 24 wk (P<.01).
Takada et al [31], 2022				✓		✓	The intervention did not negatively affect QOL.
Chan et al [39], 2024		✓	✓			✓	Subscale scores were significant: P=.02 for cognitive functioning, P=.04 for future prospective functioning, and P=.04 for arm symptom score.
Graetz et al [40], 2024					✓		Not statistically significant.
Okuyama et al [35], 2024				✓		✓	Not statistically significant.

[a]FACT-ES QLS: Functional Assessment of Cancer Therapy–Endocrine Subscale Quality of Life Scale.

[b]EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30.

[c]EORTC QLQ-BR23: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Breast Cancer Module.

[d]FACT-B: Functional Assessment of Cancer Therapy–Breast.

However, several limitations remain in sample selection when examining patients' experience with eHealth interventions. In terms of sample inclusion, the mixed methods study by Jacobs et al [37] of patients with high levels of anxiety involved

XSL•FO
RenderX

predominantly White patients with high symptom burden. Additionally, the single-arm study by Gomaa et al [38] only included patients who reported joint pain. Some studies (4/30, 13.3%) predominantly selected patients who were younger, White individuals, more educated, and economically privileged [32,38,40,47]. In terms of sample size, 2 mixed methods studies and 1 single-arm study lacked a control group. They had small sample sizes of 15 [29], 10 [54], and 39 [47]. Furthermore, several studies with low response rates (3/30, 10%) were noted [46,52,53]. The main reasons for this problem included lack of personalized interventions, poor user interface design, time constraints, technical difficulties, and privacy issues [32,47,49]. Therefore, the characteristics of nonresponding patients and patients who declined to be enrolled may introduce selection bias.

Discussion

Principal Findings

This is the first scoping review to synthesize evidence on eHealth applications specifically for symptom management in patients with breast cancer undergoing endocrine therapy, incorporating effectiveness, adherence, nurse roles, patient experience, and QOL. Interventions varied (mobile apps: 8/30, 26.7%; SMS text messaging: 5/30, 16.7%; telemedicine: 6/30, 20%; online portals: 3/30, 10%; hybrid electronic patient-reported outcomes+wearables: 2/30, 6.7%; remote tai chi: 1/30, 3.3%) but shared fundamental features: real-time data capture, threshold-triggered alerts, and remote feedback. Short-term benefits for symptom burden were consistent across app- and telemedicine-based approaches, although direct comparisons were limited. This comprehensive perspective not only facilitates an in-depth understanding of the mechanisms and actual effects of eHealth interventions but also provides valuable reference information for future clinical practice and research.

This review demonstrated eHealth's efficacy in symptom monitoring and short-term management, with some positive effects on psychological symptoms and potential indirect benefits for adherence. However, many studies of symptom monitoring and management (4/30, 13.3%) applied single-arm designs with short follow-up periods, and adherence outcomes relied heavily on self-report, introducing bias [30-32,39]. The RCTs included in this review often lacked blinding, and interventions frequently targeted single symptoms or specific subgroups, overlooking complex symptom clusters and socioeconomic factors [33,40,47,48,50]. Nevertheless, this review found that eHealth interventions could improve patients' medication adherence and short-term symptom management. However, current research relies heavily on one-sided quantitative data from patients while neglecting to assess the early warning systems of physicians and nurses.

This review revealed that nurses played a pivotal role in implementation, education, and coordination, contributing to efficiency and patient engagement. However, challenges such as workload, role clarity, and limited exploration of health care providers' perspectives persist. Patient experiences were largely positive with high satisfaction, but generalizability is constrained by limited studies and equipment requirement issues [76,77]. QOL improvements were evident in the short term but inconsistent in the long term, partly due to measurement heterogeneity. In addition, most existing studies have examined the roles and functions of physicians and nurses in eHealth interventions primarily from the patient's perspective. However, as the primary implementers of these technologies, health care providers' own acceptance of eHealth tools and the barriers they encounter in daily practice remain underexplored.

This review indicates that eHealth interventions are generally associated with favorable patient experiences, high uptake rates, and positive short-term effects on QOL among patients with breast cancer receiving endocrine therapy. Prior literature similarly underscores this population's expressed need for digital support. Nevertheless, several limitations constrain the generalizability of these findings. Many quantitative studies (4/30, 16.6%) featured small, relatively homogeneous samples, often restricted to participants with the necessary devices and digital access [31,38,39,49]. Low response rates to eHealth applications in some studies (6/30, 20%) further restricted applicability across diverse populations. These issues can hinder timely symptom detection and intervention for certain patient groups. Moreover, few studies have thoroughly explored patients' specific needs in terms of tool design, particularly with respect to cultural adaptability.

The strengths of this review lie in its comprehensive literature search and the inclusion of diverse study designs, including RCTs, single-arm studies, retrospective analyses, observational studies, mixed methods studies, and qualitative research. Quality appraisal was conducted using the Critical Appraisal Skills Programme checklists, and the complementary nature of these methodologies, where RCTs help establish causality and qualitative studies provide rich insights into patient experiences, enhances the overall credibility of the synthesis. However, notable limitations were evident in the primary studies. Most (23/30, 76.6%) lacked adequate sample diversity and representativeness. In addition, the evidence base was heavily skewed toward high-income countries, with minimal data from low- and middle-income settings or rural areas, thereby limiting global applicability and scalability. Although the use of theories, models, and frameworks has been shown to strengthen the scientific rigor and explanatory depth of digital health research [78], only 13.3% (4/30) of the studies in this review were explicitly guided by such frameworks. Furthermore, most of the included RCTs (10/13, 76.9%) lacked blinding of participants or researchers, and 26.7% (8/30) of the studies used single-arm designs without control groups, both of which increase the risk of bias and reduce the robustness of the findings. In addition, the search period ended in February 2025, which may have omitted more recent high-impact publications. The restriction to English-language peer-reviewed articles may also have excluded relevant studies published in other languages. Despite these constraints, eHealth interventions demonstrate considerable potential to reduce symptom burden and improve adherence during endocrine therapy. Future research should prioritize long-term efficacy, greater equity in access, multi-stakeholder perspectives, and integration into routine

XSL•FO
RenderX

oncology workflows to maximize real-world impact and sustainability.

Conclusions

eHealth interventions offer promising support for symptom management and short-term adherence among patients with breast cancer receiving endocrine therapy, with generally high patient acceptance. Nurses are central to successful implementation. However, evidence is limited by methodological heterogeneity, short follow-ups, inconsistent long-term outcomes, and a lack of diversity. By addressing individualized patient needs and incorporating experiences from multiple stakeholders (patients, physicians, and nurses), eHealth tools can be optimized to reduce symptom burden, improve adherence, and enhance clinical outcomes. Future studies should prioritize rigorous designs, standardized measures, long-term follow-up, and equitable implementation to realize the full potential of digital health in oncology care.

Acknowledgments

The authors declare that none of them used generative artificial intelligence in the research and writing process.

Funding

The authors declared no financial support was received for this work.

Authors' Contributions

Conceptualization: CG
Data curation: CG, LS, SZ
Investigation: CG, LS, SZ
Methodology: CG
Formal analysis: CG, LS, SZ
Supervision: NZ
Writing – original draft: CG
Writing – review & editing: LS, SZ, NZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of main characteristics of identified research, including research aim, outcome measures, and results.
[DOCX File, 95 KB - cancer_v12i1e92030_app1.docx]

Checklist 1
PRISMA-ScR Checklist.
[PDF File, 185 KB - cancer_v12i1e92030_app2.pdf]

References

1. Zhang Z, Wang J, Skinner KA, et al. Pathological features and clinical outcomes of breast cancer according to levels of oestrogen receptor expression. Histopathology 2014 Oct;65(4):508-516. [doi: 10.1111/his.12412] [Medline: 24620991]
2. Rosso R, D'Alonzo M, Bounous VE, et al. Adherence to adjuvant endocrine therapy in breast cancer patients. Curr Oncol 2023 Jan 21;30(2):1461-1472. [doi: 10.3390/curroncol30020112] [Medline: 36826073]
3. Record H, Clelland E, Rothschild HT, et al. Tamoxifen or aromatase inhibitors with ovarian function suppression in pre-menopausal stage I-III lobular breast cancer. NPJ Breast Cancer 2023 Oct 26;9(1):88. [doi: 10.1038/s41523-023-00594-3] [Medline: 37884561]
4. Schiavon G, Smith IE. Status of adjuvant endocrine therapy for breast cancer. Breast Cancer Res 2014;16(2):206. [doi: 10.1186/bcr3636] [Medline: 25032258]
5. Sasada S, Kondo N, Hashimoto H, et al. Prognostic impact of adjuvant endocrine therapy for estrogen receptor-positive and HER2-negative T1a/bN0M0 breast cancer. Breast Cancer Res Treat 2023 Dec;202(3):473-483. [doi: 10.1007/s10549-023-07097-6] [Medline: 37688665]
6. Burstein HJ, Lacchetti C, Anderson H, et al. Adjuvant endocrine therapy for women with hormone receptor-positive breast cancer: ASCO clinical practice guideline focused update. J Clin Oncol 2019 Feb 10;37(5):423-438. [doi: 10.1200/JCO.18.01160] [Medline: 30452337]
7. Cardoso F, Kyriakides S, Ohno S, et al. Early breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2019 Aug;30(8):1194-1220. [doi: 10.1093/annonc/mdz173]

XSL•FO
RenderX

8. Gradishar WJ, Moran MS, Abraham J, et al. Breast cancer, version 3.2022, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw 2022 Jun;20(6):691-722. [doi: 10.6004/jnccn.2022.0030] [Medline: 35714673]
9. Yussof I, Mohd Tahir NA, Hatah E, Mohamed Shah N. Factors influencing five-year adherence to adjuvant endocrine therapy in breast cancer patients: a systematic review. Breast 2022 Apr;62:22-35. [doi: 10.1016/j.breast.2022.01.012] [Medline: 35121501]
10. Cavazza M, Banks H, Ercolanoni M, et al. Factors influencing adherence to adjuvant endocrine therapy in breast cancer-treated women: using real-world data to inform a switch from acute to chronic disease management. Breast Cancer Res Treat 2020 Aug;183(1):189-199. [doi: 10.1007/s10549-020-05748-6] [Medline: 32591986]
11. Franzoi MA, Agostinetto E, Perachino M, et al. Evidence-based approaches for the management of side-effects of adjuvant endocrine therapy in patients with breast cancer. Lancet Oncol 2021 Jul;22(7):e303-e313. [doi: 10.1016/S1470-2045(20)30666-5] [Medline: 33891888]
12. Wang L, Wei T, Liu J, et al. Experiences and management needs of endocrine therapy-related symptoms in patients with breast cancer: a systematic review and qualitative evidence synthesis protocol. BMJ Open 2023 Dec;13(12):e073915. [doi: 10.1136/bmjopen-2023-073915]
13. Global strategy on digital health 2020-2027. World Health Organization. 2025. URL: https://www.who.int/publications/i/item/9789240116870 [accessed 2026-04-26]
14. Ahern DK, Kreslake JM, Phalen JM. What is eHealth (6): perspectives on the evolution of eHealth research. J Med Internet Res 2006 Mar 31;8(1):e4. [doi: 10.2196/jmir.8.1.e4] [Medline: 16585029]
15. Boogerd EA, Arts T, Engelen LJ, van de Belt TH. "What is eHealth": time for an update? JMIR Res Protoc 2015 Mar 12;4(1):e29. [doi: 10.2196/resprot.4065] [Medline: 25768939]
16. Vaffis S, Whaley S, Axon DR, et al. Features of cancer mHealth apps and evidence for patient preferences: scoping literature review. JMIR Cancer 2023 Apr 28;9:e37330. [doi: 10.2196/37330] [Medline: 37115587]
17. Li X, Zhang N, Yang J, Geng Z, Zhou J, Zhang J. Weight management personas of breast cancer patients undergoing chemotherapy in China: a multi-method study. BMC Med Inform Decis Mak 2024 Apr 25;24(1):108. [doi: 10.1186/s12911-024-02515-1] [Medline: 38664653]
18. Reeves MM, Terranova CO, Winkler EA, et al. Effect of a remotely delivered weight loss intervention in early-stage breast cancer: randomized controlled trial. Nutrients 2021 Nov 15;13(11):4091. [doi: 10.3390/nu13114091] [Medline: 34836345]
19. Du X, Li Y, Fu L, et al. Strategies in activating lymphatic system to promote lymph flow on lymphedema symptoms in breast cancer survivors: a randomized controlled trial. Front Oncol 2022;12:1015387. [doi: 10.3389/fonc.2022.1015387]
20. Oswald LB, Morales-Cruz J, Eisel SL, et al. Pilot randomized controlled trial of eHealth cognitive-behavioral therapy for insomnia among Spanish-speaking breast cancer survivors. J Behav Med 2022 Jun;45(3):503-508. [doi: 10.1007/s10865-022-00313-6] [Medline: 35438442]
21. Shi N, Wong AK, Wong FK, Zhang N, Lai X, Gan L. A nurse-led mHealth self-management program (mChemotherapy) for breast cancer patients undergoing chemotherapy: study protocol of a randomized controlled pilot study. Cancer Control 2022;29:10732748221115469. [doi: 10.1177/10732748221115469] [Medline: 35921175]
22. Jing F, Zhu Z, Qiu J, et al. Symptom clusters and quality of life in patients with breast cancer receiving endocrine therapy in China. Semin Oncol Nurs 2023 Aug;39(4):151443. [doi: 10.1016/j.soncn.2023.151443] [Medline: 37173235]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097. [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
24. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: an R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. Campbell Syst Rev 2022;18(2):e1230. [doi: 10.1002/cl2.1230] [Medline: 36911350]
25. Krok-Schoen JL, Naughton MJ, Young GS, et al. Increasing adherence to adjuvant hormone therapy among patients with breast cancer: a smart phone app-based pilot study. Cancer Control 2019;26(1):1073274819883287. [doi: 10.1177/1073274819883287] [Medline: 31736324]
26. Lally P, May CN, Mitchell ES, McCallum M, Michaelides A, Fisher A. Prototype of an app designed to support self-management for health behaviors and weight in women living with breast cancer: qualitative user experience study. JMIR Cancer 2024 Dec 20;10:e48170. [doi: 10.2196/48170] [Medline: 39705674]
27. Brett J, Boulton M, Watson E. Development of an e-health app to support women prescribed adjuvant endocrine therapy after treatment for breast cancer. Patient Prefer Adherence 2018;12:2639-2647. [doi: 10.2147/PPA.S187692] [Medline: 30587936]
28. Harbeck N, Kates R, Schinköthe T, et al. Favorable impact of therapy management by an interactive eHealth system on severe adverse events in patients with hormone receptor-positive, HER2-negative locally advanced or metastatic breast cancer treated by palbociclib and endocrine therapy. Cancer Treat Rev 2023 Dec;121:102631. [doi: 10.1016/j.ctrv.2023.102631] [Medline: 37862832]
29. Bergqvist J, Lundström S, Wengström Y. Patient interactive digital support for women with adjuvant endocrine therapy in order to increase compliance and quality of life. Support Care Cancer 2021 Jan;29(1):491-497. [doi: 10.1007/s00520-020-05476-z] [Medline: 32405965]

XSL•FO
RenderX

30. Neuner J, Weil E, Fergestrom N, et al. Feasibility of a pharmacist-led symptom monitoring and management intervention to improve breast cancer endocrine therapy adherence. J Am Pharm Assoc (2003) 2022;62(4):1321-1328. [doi: 10.1016/j.japh.2022.03.001] [Medline: 35393248]
31. Takada F, Okuyama H, Nakamura S, Fujita KI. Application of personal health record in enhancing the quality of life in patients with breast cancer who received adjuvant hormonal therapy. Eur J Breast Health 2022 Apr;18(2):155-162. [doi: 10.4274/ejbh.galenos.2022.2021-12-2] [Medline: 35445175]
32. Mougalian SS, Epstein LN, Jhaveri AP, et al. Bidirectional text messaging to monitor endocrine therapy adherence and patient-reported outcomes in breast cancer. JCO Clin Cancer Inform 2017 Nov;1:1-10. [doi: 10.1200/CCI.17.00015] [Medline: 30657377]
33. Hershman DL, Unger JM, Hillyer GC, et al. Randomized trial of text messaging to reduce early discontinuation of adjuvant aromatase inhibitor therapy in women with early-stage breast cancer: SWOG S1105. J Clin Oncol 2020 Jul 1;38(19):2122-2129. [doi: 10.1200/JCO.19.02699] [Medline: 32369401]
34. Smith KL, Tsai HL, Lim D, et al. Feasibility of symptom monitoring during the first year of endocrine therapy for early breast cancer using patient-reported outcomes collected via smartphone app. JCO Oncol Pract 2023 Nov;19(11):981-989. [doi: 10.1200/OP.23.00038] [Medline: 37733984]
35. Okuyama H, Takada F, Taira N, Nakamura S. A randomized trial of the impact of symptom monitoring using an electronic patient-reported outcome app on health-related quality of life in postmenopausal breast cancer patients receiving adjuvant endocrine therapy. Breast Cancer 2024 Sep;31(5):787-797. [doi: 10.1007/s12282-024-01592-4] [Medline: 38796818]
36. Advani P, Brewster AM, Baum GP, Schover LR. A pilot randomized trial to prevent sexual dysfunction in postmenopausal breast cancer survivors starting adjuvant aromatase inhibitor therapy. J Cancer Surviv 2017 Aug;11(4):477-485. [doi: 10.1007/s11764-017-0606-3] [Medline: 28229275]
37. Jacobs JM, Walsh EA, Rapoport CS, et al. Development and refinement of a telehealth intervention for symptom management, distress, and adherence to adjuvant endocrine therapy after breast cancer. J Clin Psychol Med Settings 2021 Sep;28(3):603-618. [doi: 10.1007/s10880-020-09750-4] [Medline: 33219901]
38. Gomaa S, West C, Lopez AM, et al. A telehealth-delivered tai chi intervention (TaiChi4Joint) for managing aromatase inhibitor-induced arthralgia in patients with breast cancer during COVID-19: longitudinal pilot study. JMIR Form Res 2022;6(6):e34995. [doi: 10.2196/34995]
39. Chan WL, Wong YL, Tai YL, et al. Digital rehabilitation program for breast cancer survivors on adjuvant hormonal therapy: a feasibility study. Cancers (Basel) 2024 Dec 5;16(23):4084. [doi: 10.3390/cancers16234084] [Medline: 39682269]
40. Graetz I, Hu X, Kocak M, et al. Remote monitoring app for endocrine therapy adherence among patients with early-stage breast cancer: a randomized clinical trial. JAMA Netw Open 2024 Jun 3;7(6):e2417873. [doi: 10.1001/jamanetworkopen.2024.17873] [Medline: 38935379]
41. Walsh LE, Dunderdale L, Horick N, Temel JS, Greer JA, Jacobs JM. Intervention-related changes in coping ability drives improvements in mood and quality of life for patients taking adjuvant endocrine therapy. Psychooncology 2024 Dec;33(12):e70049. [doi: 10.1002/pon.70049] [Medline: 39706803]
42. Richardson D, Zhan L, Mahtani R, et al. A prospective observational study of patient-reported functioning and quality of life in advanced and metastatic breast cancer utilizing a novel mobile application. Breast Cancer Res Treat 2021 May;187(1):113-124. [doi: 10.1007/s10549-020-06082-7] [Medline: 33428072]
43. Harrell M, Fabbri D, Levy M. Analysis of adjuvant endocrine therapy in practice from electronic health record data of patients with breast cancer. JCO Clin Cancer Inform 2017 Nov;1:1-8. [doi: 10.1200/CCI.16.00044] [Medline: 30657375]
44. Yin Z, Harrell M, Warner JL, Chen Q, Fabbri D, Malin BA. The therapy is making me sick: how online portal communications between breast cancer patients and physicians indicate medication discontinuation. J Am Med Inform Assoc 2018 Nov 1;25(11):1444-1451. [doi: 10.1093/jamia/ocy118] [Medline: 30380083]
45. Aurell C, Haidar A, Giglio D. Follow-up routines matter for adherence to endocrine therapy in the adjuvant setting of breast cancer. Breast Cancer (Auckl) 2024;18:11782234241240171. [doi: 10.1177/11782234241240171] [Medline: 38628960]
46. Graetz I, McKillop CN, Stepanski E, Vidal GA, Anderson JN, Schwartzberg LS. Use of a web-based app to improve breast cancer symptom management and adherence for aromatase inhibitors: a randomized controlled feasibility trial. J Cancer Surviv 2018 Aug;12(4):431-440. [doi: 10.1007/s11764-018-0682-z] [Medline: 29492753]
47. Krukowski RA, Hu X, Arshad S, et al. Symptom monitoring app use associated with medication adherence among woman survivors of breast cancer on adjuvant endocrine therapy. JCO Clin Cancer Inform 2024 Dec;8:e2400179. [doi: 10.1200/CCI-24-00179] [Medline: 39642329]
48. Park HR, Kang HS, Kim SH, Singh-Carlson S. Effect of a smart pill bottle reminder intervention on medication adherence, self-efficacy, and depression in breast cancer survivors. Cancer Nurs 2022;45(6):E874-E882. [doi: 10.1097/NCC.0000000000001030] [Medline: 34661562]
49. Wheeler SB, Spencer J, Drier SW, Fray N, Reeder-Hayes KE. Motivational interviewing counseling to increase endocrine therapy adherence in diverse patients. Cancers (Basel) 2023 Mar 25;15(7):1973. [doi: 10.3390/cancers15071973] [Medline: 37046634]

XSL•FO
RenderX

50. Tan EH, Wong AL, Tan CC, et al. Improving medication adherence with adjuvant aromatase inhibitor in women with breast cancer: a randomised controlled trial to evaluate the effect of short message service (SMS) reminder. Breast 2020 Oct;53:77-84. [doi: 10.1016/j.breast.2020.06.012] [Medline: 32652462]
51. Jacobs JM, Post K, Massad K, et al. A telehealth intervention for symptom management, distress, and adherence to adjuvant endocrine therapy: a randomized controlled trial. Cancer 2022 Oct 1;128(19):3541-3551. [doi: 10.1002/cncr.34409] [Medline: 35924869]
52. Çınar D, Karadakovan A, Erdoğan AP. Effect of mobile phone app-based training on the quality of life for women with breast cancer. Eur J Oncol Nurs 2021 Jun;52:101960. [doi: 10.1016/j.ejon.2021.101960] [Medline: 33882446]
53. Ahlstedt Karlsson S, Henoch I, Olofsson Bagge R, Wallengren C. Person-centred support programme (RESPECT intervention) for women with breast cancer treated with endocrine therapy: a feasibility study. BMJ Open 2022 Oct 5;12(10):e060946. [doi: 10.1136/bmjopen-2022-060946] [Medline: 36198470]
54. Nathan M, Beatty L, Coborn J, et al. Targeting psychological and menopausal factors linked to non-adherence to anti-estrogen therapy in women with early-stage hormone-receptor positive breast cancer: a web-based feasibility study of Finding My Way. Psychooncology 2023 Sep;32(9):1461-1465. [doi: 10.1002/pon.6195] [Medline: 37496178]
55. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions: The Cochrane Collaboration; 2019:205-228. [doi: 10.1002/9781119536604]
56. Long HA, French DP, Brooks JM. Optimising the value of the critical appraisal skills programme (CASP) tool for quality appraisal in qualitative evidence synthesis. Res Methods Med Health Sci 2020;1(1):31-42. [doi: 10.1177/2632084320947559]
57. ÇInar D, Karadakovan A, Erdoğan AP. The effect of e-mobile education on the quality of life in women with breast cancer. Ann Oncol 2019 Oct;30(Supplement 5):v818-v819. [doi: 10.1093/annonc/mdz272.010]
58. Vrijens B, De Geest S, Hughes DA, et al. A new taxonomy for describing and defining adherence to medications. Br J Clin Pharmacol 2012 May;73(5):691-705. [doi: 10.1111/j.1365-2125.2012.04167.x] [Medline: 22486599]
59. Basit SA, Mathews N, Kunik ME. Telemedicine interventions for medication adherence in mental illness: a systematic review. Gen Hosp Psychiatry 2020;62:28-36. [doi: 10.1016/j.genhosppsych.2019.11.004] [Medline: 31775066]
60. Chilala CI, Kassavou A, Sutton S. Evaluating the effectiveness of remote behavioral interventions facilitated by health care providers at improving medication adherence in cardiometabolic conditions: a systematic review and meta-analysis. Ann Behav Med 2023 Feb 4;57(2):99-110. [doi: 10.1093/abm/kaac037] [Medline: 35916782]
61. Kim SK, Park SY, Hwang HR, Moon SH, Park JW. Effectiveness of mobile health intervention in medication adherence: a systematic review and meta-analysis. J Med Syst 2025 Jan 17;49(1):13. [doi: 10.1007/s10916-024-02135-2] [Medline: 39821698]
62. Cahir C, Bennett K, Dombrowski SU, et al. Informing interventions to improve uptake of adjuvant endocrine therapy in women with breast cancer: a theoretical-based examination of modifiable influences on non-adherence. Support Care Cancer 2023 Mar 4;31(3):200. [doi: 10.1007/s00520-023-07658-x] [Medline: 36869943]
63. Farokhzadian J, Khajouei R, Hasman A, Ahmadian L. Nurses' experiences and viewpoints about the benefits of adopting information technology in health care: a qualitative study in Iran. BMC Med Inform Decis Mak 2020 Sep 21;20(1):240. [doi: 10.1186/s12911-020-01260-5] [Medline: 32958042]
64. Paatela S, Kyytsönen M, Saranto K, Kinnunen UM, Vehko T. Experiences of electronic health records' and client information systems' use on a mobile device and factors associated with work time savings among practical nurses: cross-sectional study. J Med Internet Res 2024 May 29;26:e46954. [doi: 10.2196/46954] [Medline: 38809583]
65. Alipour J, Mehdipour Y, Zakerabasali S, Karimi A. Nurses' perspectives on using mobile health applications in southeastern Iran: awareness, attitude, and obstacles. PLoS One 2025;20(3):e0316631. [doi: 10.1371/journal.pone.0316631] [Medline: 40063563]
66. Yada H, Soejima K. Digital transformation in cardiology — mobile health. Circ J 2026;90(1):3-11. [doi: 10.1253/circj.CJ-24-0654]
67. Trojan A, Kühne C, Kiessling M, et al. Impact of electronic patient-reported outcomes on unplanned consultations and hospitalizations in patients with cancer undergoing systemic therapy: results of a patient-reported outcome study compared with matched retrospective data. JMIR Form Res 2024 May 6;8:e55917. [doi: 10.2196/55917] [Medline: 38710048]
68. van Steenbergen G, van Veghel D, van Lieshout D, Sperwer M, Ter Woorst J, Dekker L. Effects of video-based patient education and consultation on unplanned health care utilization and early recovery after coronary artery bypass surgery (IMPROV-ED): randomized controlled trial. J Med Internet Res 2022 Aug 26;24(8):e37728. [doi: 10.2196/37728] [Medline: 36018625]
69. Theodore BR, Whittington J, Towle C, et al. Transaction cost analysis of in-clinic versus telehealth consultations for chronic pain: preliminary evidence for rapid and affordable access to interdisciplinary collaborative consultation. Pain Med 2015 Jun;16(6):1045-1056. [doi: 10.1111/pme.12688] [Medline: 25616057]
70. van der Meij E, Anema JR, Otten RH, Huirne JA, Schaafsma FG. The effect of perioperative e-health interventions on the postoperative course: a systematic review of randomised and non-randomised controlled trials. PLoS One 2016;11(7):e0158612. [doi: 10.1371/journal.pone.0158612] [Medline: 27383239]

XSL•FO
RenderX

71. Theiler RN, Butler-Tobah Y, Hathcock MA, Famuyide A. OB Nest randomized controlled trial: a cost comparison of reduced visit compared to traditional prenatal care. BMC Pregnancy Childbirth 2021 Jan 21;21(1):71. [doi: 10.1186/s12884-021-03557-3] [Medline: 33478433]
72. Bok SK, Song Y, Lim A, Choi H, Shin H, Jin S. Korean translation and psychometric evaluation of Korean version EORTC QLQ-BRECON23. Int J Environ Res Public Health 2020 Dec 8;17(24):9163. [doi: 10.3390/ijerph17249163] [Medline: 33302470]
73. Clarijs ME, Oemrawsingh A, Bröker ME, Verhoef C, Lingsma H, Koppert LB. Quality of life of caregivers of breast cancer patients: a cross-sectional evaluation. Health Qual Life Outcomes 2022 Feb 19;20(1):29. [doi: 10.1186/s12955-022-01930-0] [Medline: 35183212]
74. Dewi IP, Gartika N, Sanusi S, Suryatiningsih D. Determinants of spiritual well-being that impact breast cancer patients' quality of life. J Holist Nurs 2024 Jun;42(2_suppl):S40-S46. [doi: 10.1177/08980101231181272] [Medline: 37287242]
75. Chopra I, Kamal KM. A systematic review of quality of life instruments in long-term breast cancer survivors. Health Qual Life Outcomes 2012 Jan 31;10:14. [doi: 10.1186/1477-7525-10-14] [Medline: 22289425]
76. Frennert S, Petersson L, Muhic M, et al. Materiality and the mediating roles of eHealth: a qualitative study and comparison of three cases. Digit Health 2022;8:20552076221116782. [doi: 10.1177/20552076221116782] [Medline: 35935713]
77. Frennert S, Petersson L, Erlingsdottir G. "More" work for nurses: the ironies of eHealth. BMC Health Serv Res 2023 Apr 27;23:411. [doi: 10.1186/s12913-023-09418-3]
78. Rouleau G, Wu K, Ramamoorthi K, et al. Mapping theories, models, and frameworks to evaluate digital health interventions: scoping review. J Med Internet Res 2024 Feb 5;26:e51098. [doi: 10.2196/51098] [Medline: 38315515]

Abbreviations

AET: adjuvant endocrine therapy
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QOL: quality of life
RCT: randomized controlled trial

Edited by M Balcarras; submitted 24.Jan.2026; peer-reviewed by A Trojan; revised version received 14.Apr.2026; accepted 15.Apr.2026; published 04.Jun.2026.

Please cite as:
Gu C, Shen L, Zhang S, Zhang N
The Application of eHealth in Symptom Management of Patients With Breast Cancer During Endocrine Therapy: Scoping Literature Review
JMIR Cancer 2026;12:e92030
URL: https://cancer.jmir.org/2026/1/e92030
doi:10.2196/92030

© Chengjia Gu, Lili Shen, Shiyu Zhang, Nan Zhang. Originally published in JMIR Cancer (https://cancer.jmir.org), 4.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multiturn Large Language Model–Based Conversational Agents for Patients With Cancer and Caregivers: Scoping Review

Yeongrok Jeong[1], MSN; Hyejeon Cha[2], MSN; Eunyoung E Suh[1], PhD

[1]Center for World-leading Human-care Nurse Leaders for the Future by Brain Korea 21 (BK21) Four Project, College of Nursing, Research Institute of Nursing Science, Seoul National University, 103 Daehak-ro, Jongno-gu, Seoul, Republic of Korea

[2]College of Nursing, Seoul National University, Seoul, Republic of Korea

Corresponding Author:
Eunyoung E Suh, PhD
Center for World-leading Human-care Nurse Leaders for the Future by Brain Korea 21 (BK21) Four Project, College of Nursing, Research Institute of Nursing Science, Seoul National University, 103 Daehak-ro, Jongno-gu, Seoul, Republic of Korea

Abstract

Background: Large language model (LLM)–based conversational agents are increasingly used in health care, yet their capacity to support genuine multiturn dialogue remains underexplored. In oncology, where patients and caregivers experience complex informational and emotional needs throughout the disease trajectory, conversational agents may support information provision, symptom consultation, and emotional assistance. However, research specifically examining multiturn conversational agents designed for patients with cancer and informal caregivers remains limited.

Objective: This scoping review aimed to map the research landscape of LLM-based multiturn conversational chatbots developed for patients with cancer and informal caregivers, focusing on system design, intervention purposes, evaluation approaches, safety considerations, and transparency of LLM-related components.

Methods: This scoping review followed the Joanna Briggs Institute methodology and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines. Six databases—PubMed, Embase, Scopus, Web of Science, CINAHL, and PsycINFO—were searched for studies published between January 2022 and January 2026, with supplementary searches conducted in IEEE Xplore Digital Library and ACM Digital Library in May 2026. Studies were included if they described LLM-based chatbots designed for patients with cancer or informal caregivers that supported multiturn conversational interaction. Two reviewers independently conducted the study selection and data extraction.

Results: Eight studies met the inclusion criteria. Most studies focused on prototype development, with limited research evaluating clinical outcomes. ChatGPT-based models were the most commonly used LLMs, and retrieval-augmented generation techniques were applied in several studies. Chatbots were primarily designed for emotional support or information provision. Evaluation approaches varied widely, including response quality, psychological outcomes, and user experience. However, no studies evaluated interaction-level characteristics such as conversational continuity or context retention, and only 2 studies reported any conversational memory mechanism. Reporting on safety risks, mitigation strategies, prompt design, model parameters, and adherence to LLM reporting guidelines was often limited or absent.

Conclusions: This scoping review identified only 8 studies on LLM-based multiturn conversational chatbots for patients with cancer and informal caregivers. The field remains at an early stage, characterized by prototype-oriented development, heterogeneous design and evaluation approaches, and inconsistent safety and transparency reporting. Future development should prioritize genuine conversational capability, safety management, and transparent reporting.

(JMIR Cancer 2026;12:e96241) doi:10.2196/96241

KEYWORDS

chatbot; large language model; LLM chatbot; conversational agent; artificial intelligence; generative artificial intelligence; generative AI chatbot; cancer; neoplasms

Introduction

Large language models (LLMs) represent an advanced subset of natural language processing (NLP), capable of generating contextually appropriate responses through probabilistic text generation [1]. Chatbots using LLMs can generate responses in multiturn interactions while maintaining context, without relying on predefined rules or response scenarios [2,3]. Such capability distinguishes them from traditional rule-based NLP-based chatbots, revealing fundamental differences in design principles and conversational approaches [4,5]. This generative nature, however, introduces novel risk structures that are more pronounced than those in rule-based NLP systems [6,7].

XSL•FO
RenderX

Following the proliferation of LLMs, LLM-based conversational agents for diverse purposes—such as patient education, symptom consultation, emotional support, and treatment decision-making assistance—are rapidly emerging in the medical field, with accumulating evaluation studies [8]. In mental health and chronic disease management, LLM-based conversational agents have increasingly been developed to improve patient outcomes, demonstrating clinical benefits across conditions including depression, anxiety, eating disorders, and diabetes [9-11].

In oncology, where information needs and uncertainty are high during diagnosis and treatment, the potential application of LLM-based conversational agents is gaining attention [12,13]. For patients with cancer and informal caregivers, whose communication, informational, and emotional needs evolve throughout the disease course, multiturn interactions are a critical characteristic necessitating treatment as an independent analytical category [14,15]. Chatbot studies applying diverse LLMs, such as ChatGPT, are rapidly proliferating in this context, alongside advances in development frameworks including retrieval-augmented generation (RAG), knowledge graphs, and LangChain. However, existing literature tends to focus on performance comparisons at the single-turn question-answer level, making it difficult to distinguish and understand the design and evaluation characteristics specific to patient-directed conversational agents [16].

Multiturn conversational interaction refers to systems that maintain context by integrating the current message and previous conversation history when generating responses [17,18]. Yet, in LLM-based conversational agents for patients with cancer or informal caregivers, multiturn capability is rarely evaluated or explicitly implemented in practice, despite being considered inherent to conversational agents.

Simultaneously, LLMs carry safety risks, including hallucinations and responses that may pose psychological risks to vulnerable patients [6,7,19]. Although expert oversight and escalation systems have been suggested as mitigation strategies, systematic reporting of safety risks and mitigation approaches in LLM-based oncology chatbot studies remains limited.

Furthermore, if reporting on LLM-specific components—such as prompts, model settings, and external knowledge integration methods—is insufficient, reproducibility and interpretability may be limited [20]. Reporting guidelines, such as TRIPOD-LLM (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Large Language Models) and CHART (Chatbot Assessment Reporting Tool), have been developed to address this concern [21,22]; however, the extent to which LLM-based chatbot studies in oncology adhere to these guidelines remains unclear.

Several reviews have examined related topics in oncology. Wang et al [23] focused primarily on rule-based systems predating the LLM era. Chen et al [24] evaluated the medical accuracy of LLM-based chatbots in oncology across diagnostic and management tasks, but centered on single-turn performance rather than patient-facing conversational interaction. Jiang et al [25] synthesized conversational agent interventions in cancer care without restricting to LLM-based systems, identifying LLMs as a key emerging technology to enhance personalization and sustained engagement. To our knowledge, none of these reviews specifically examined the unique characteristics of LLM-based chatbots such as system design, safety, and transparency in the context of multiturn conversational interaction with patients with cancer and informal caregivers.

Therefore, this scoping literature review aims to identify the overall research landscape of LLM-based conversational chatbot studies developed for patients with cancer and their caregivers and sets the following research questions:

1. What types of LLM-based conversational chatbots have been developed for patients with cancer and caregivers, and what are their key design elements?
2. What are the primary intervention goals of the developed LLM-based chatbots?
3. How and at what level were LLM-based chatbots evaluated?
4. What safety risks were reported for LLM-based chatbots, and what mitigation strategies were used?
5. To what extent were the key LLM-related components of the chatbot development process transparently reported to support reproducibility?

Methods

This scoping systematic review was conducted according to the Joanna Briggs Institute (JBI) scoping systematic review methodology [26] and reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines (Checklist 1) [27]. The protocol for this study was preregistered in the Open Science Framework (OSF) [28], and a critical appraisal of individual studies was not performed.

Search Strategy

Search Sources

A pilot search for relevant literature was conducted in 3 databases—PubMed, Embase, and Scopus—from January 20, 2026, to January 31, 2026. Based on the pilot search results, the search strategy and inclusion or exclusion criteria were refined. An expanded search was then conducted from February 9, 2026, to February 13, 2026, across 6 databases initially: PubMed, Embase, Scopus, Web of Science, CINAHL, and PsycINFO. IEEE Xplore Digital Library and ACM Digital Library were subsequently searched on May 10, 2026, to ensure comprehensive coverage of computer science and AI literature.

Search Terms

Search terms were designed to comprehensively identify studies on LLM-based interactive chatbots targeting patients with cancer and their informal caregivers. The search query comprised four blocks: (1) cancer, (2) conversational agent or chatbot, (3) LLM or generative AI, and (4) patient or informal caregiver. Each block combined controlled vocabulary (eg, MeSH/EMTREE) with free-text terms using AND operators. The same search structure was applied across all databases; the full search terms for each database are presented in Multimedia Appendix 1.

XSL•FO
RenderX

Study Eligibility Criteria

This review included peer-reviewed journal articles and conference proceedings, encompassing various research designs such as experimental studies, observational studies, mixed methods studies, system development and evaluation studies, and review studies. Conversely, literature lacking systematic research design or results reporting, such as editorials, opinion pieces, and conference abstracts, was excluded. The literature was limited to English-language articles published between January 2022 and January 2026.

Participants

The participants for this scoping review were defined as studies addressing LLM-based chatbots developed primarily for patients with cancer and their informal caregivers, including family members, spouses, and other unpaid supporters involved in the care of patients with cancer. Participants were restricted to adult patients with cancer and adult caregivers. Studies including pediatric participants were included in this review if adult patients were also a primary focus.

Concept

The core concept of this review was interactive chatbots developed using LLMs. Multiturn conversational interaction was operationally defined as systems designed to support iterative dialogue, including those that (1) explicitly generate responses based on previous conversational context, (2) are architecturally configured to enable context-maintaining exchanges, or (3) are developed with the explicit intent of facilitating ongoing conversational interaction with users. Conversely, conversational systems implemented solely using rule-based NLP techniques or predefined response scenarios were excluded from this study's scope. Additionally, single-turn chatbots evaluated solely for accuracy, readability, and so forth, without presupposing interaction, were excluded.

Context

The purpose of this review is to identify the overall research landscape of LLM-based conversational chatbot studies developed for patients with cancer and their caregivers. Therefore, as it systematically organizes the purpose, design approach, application environment, and evaluation methods of the technology, it includes studies conducted in various contexts such as hospitals, homes, and online.

Study Selection

This systematic review identified relevant studies from databases, removed duplicates, performed an initial screening based on titles and abstracts, and reviewed full texts to determine final inclusion according to the predefined inclusion and exclusion criteria. Two researchers (YJ, HC) independently performed the literature screening and full-text review. Disagreements in selection were resolved through discussion and consensus. To ensure efficiency and transparency in the selection process, the web-based systematic review program Rayyan (Rayyan Systems Inc.) was used. The selection process and results were reported using the PRISMA-ScR flowchart.

Some included studies were identified as borderline cases. Given the limited number of studies explicitly addressing multiturn LLM-based conversational agents in oncology, an inclusive approach was adopted to comprehensively capture the emerging research landscape in this field. The initial judgments of each author and the consensus process for all included studies are documented in Multimedia Appendix 2.

Data Extraction and Data Synthesis

Data extraction was performed independently by 2 researchers using a predefined data extraction form, which was refined after pilot testing on a subset of studies. Data extraction was recorded using an electronic database tool (Notion), with a third researcher (ES) mediating when necessary.

Data extraction items included the following: (1) authors and publication year, (2) study country and setting, (3) study design and type, (4) participant characteristics, (5) primary purpose of the chatbot, (6) type of LLM used and system configuration, (7) evaluation design and key evaluation results, (8) reporting on safety and risk management, and (9) reporting on transparency or reproducibility elements related to the LLM.

The level of reporting on transparency and reproducibility of the LLM-related components was classified according to predefined criteria, adapted from reporting elements specified in TRIPOD-LLM [21]. The level of prompt reporting was assessed based on whether (1) system prompts, (2) user prompts, and (3) specific descriptions of model settings or parameters were provided. Complete reporting was defined as cases where all 3 elements were explicitly described, while partial reporting was defined as cases where only some elements were reported. Cases in which relevant information was not explicitly described were classified as not reported.

RAG was classified as applied when the term was explicitly used in the study or when the study described integrating information from external documents prior to response generation. Additionally, safety risk factors were categorized based on the risk factors described in each study.

The extracted data were organized into tables and diagrams according to study characteristics and key concepts, and the results were presented using descriptive mapping.

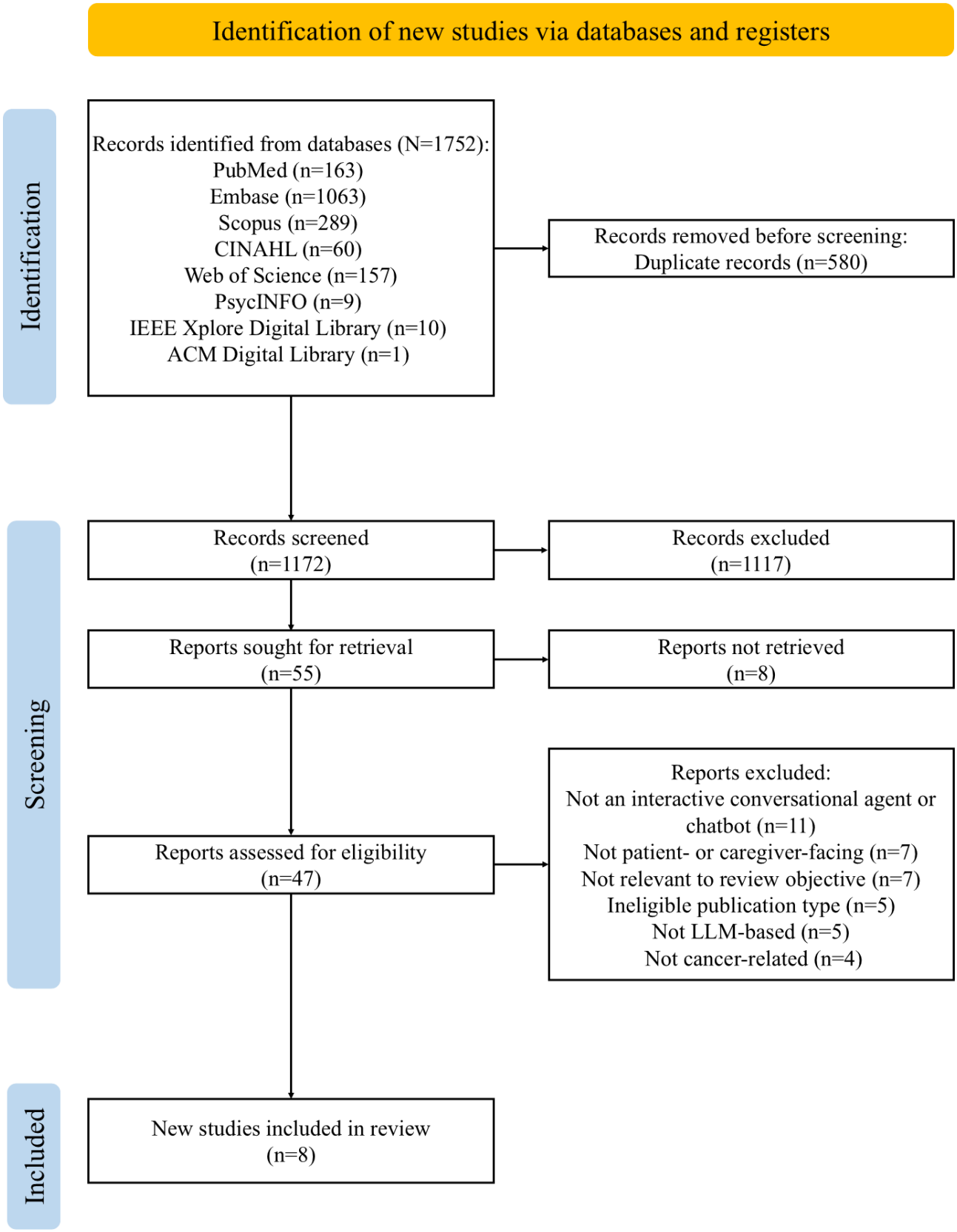
Results

Search Results

This review searched a total of 8 databases according to a predefined search protocol. The search and selection process is presented in Figure 1. The initial search identified a total of 1752 documents, including books, conference proceedings, and journal articles. After removing 580 duplicate documents, 1172 documents were included for title and abstract screening.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram showing the study selection process.



Following the title and abstract screening, a total of 1117 documents were excluded, and 55 were selected for full-text review. Excluded documents were categorized based on predefined inclusion and exclusion criteria, with some documents meeting multiple exclusion criteria simultaneously. The specific distribution of exclusion reasons at the title and abstract stage is presented in Multimedia Appendix 3.

During the full-text review stage, 47 of the 55 articles were reviewed, excluding 8 articles with restricted full-text access. These comprised 1 oral abstract, 3 conference abstracts, 2

XSL•FO
RenderX

ePosters, 1 book chapter, and 1 news or views article, none of which met our inclusion criteria. Articles excluded during the full-text review were categorized based on 1 primary exclusion reason. The main exclusion reasons were as follows: not being an interactive chatbot (11 studies; eg, studies generating tailored health education materials using LLMs without conversational interaction, or evaluating the accuracy of single-turn question-answering systems), not primarily targeting patients or caregivers (7 studies), not addressing a cancer-related context (4 studies), not using an LLM (5 studies), publication type not meeting inclusion criteria (5 studies), and insufficient relevance to the study objectives (7 studies).

Ultimately, a total of 8 studies were included in this scoping review. The limited number of included studies reflects that research applying LLM-based interactive chatbots to patients with cancer and caregivers is a relatively recent field of study.

General Description of the Studies

A total of 8 studies were included in this systematic review (Table 1). Two studies were reported in 2024 and 6 in 2025. Countries where studies were conducted included Germany, Japan, South Korea, the United States, Turkey, India, China, and the United Kingdom, with 1 study each from these countries. Publication types comprised 5 journal articles and 3 conference proceedings.

Regarding study design, development studies were the most common (5), followed by clinical evaluation studies (2) and usability evaluation studies (1). Seven studies developed chatbots exclusively for patients with cancer, while 1 study developed a chatbot exclusively for informal caregivers. No included study developed a chatbot for joint use by both patients and caregivers. Three studies specified breast cancer as the cancer type, while the remaining 5 studies did not restrict to a specific cancer type.

Two studies targeted newly diagnosed patients with cancer within 1 year of diagnosis, while 6 studies did not restrict by disease stage or timing of diagnosis. One study included a control group, 2 applied a pre-post design, and 5 did not use a comparison group, indicating varied comparative designs across studies. Overall, the included studies showed heterogeneous characteristics in terms of research design, participant characteristics, and comparison methods.

XSL•FO
RenderX

Table . General characteristics of the included studies.

Characteristics	Number of studies
Year of publication	
2024	2 [29,30]
2025	6 [31-36]
Country	
Germany	1 [33]
Japan	1 [34]
Republic of Korea	1 [30]
The United States	1 [36]
Turkey	1 [31]
India	1 [32]
China	1 [29]
The United Kingdom	1 [35]
Type of publication	
Journal article	5 [29-31,33,34]
Conference paper	3 [32,35,36]
Study design	
Clinical evaluation	2 [31,34]
Development study	5 [29,30,32,33,35]
Usability study	1 [36]
Target population	
Patients	7 [29-35]
Caregivers	1 [36]
Cancer type	
Breast cancer	3 [29,33,35]
Not restricted	5 [30-32,34,36]
Disease status	
Newly diagnosed (≤1 year since diagnosis)	2 [31,34]
Not restricted	6 [29,30,32,33,35,36]
Comparator design	
Controlled comparative design	1 [31]
Pre-post design	2 [34,36]
No comparator applied	5 [29,30,32,33,35]

Types and Design Characteristics of LLM-Based Chatbots

Analysis of the types of LLMs used in the studies showed that ChatGPT 4.0 accounted for the largest share with 5 studies. ChatGPT 3.5, ChatGPT 3.5 Turbo, and Gemini were each used in 1 study. Additionally, some studies used various LLMs such as Gemini, Llama 3.1, Mistral 7B, and PHI 3.5 to perform model comparisons (Table 2).

Five studies applied RAG techniques, while 3 studies did not report on RAG application. Regarding knowledge sources used by chatbots, clinical practice guidelines were the most common source, used in 3 studies. Educational materials, books, and published literature were each used in 1 study. Conversely, 3 studies did not explicitly specify their knowledge sources.

Three studies explicitly used frameworks like LangChain or knowledge graphs for chatbot development, while the remaining 5 did not report on frameworks. Overall, the studies included in this literature review centered on ChatGPT-based models. However, heterogeneity was observed in model composition, with some studies using open-weight models for comparative evaluations. Differences were also noted in the level of reporting regarding knowledge sources and system implementation methods.

XSL•FO
RenderX

Meanwhile, the level of reporting on conversational memory handling was also limited. Only 1 study implemented embedding-based memory retrieval, and 1 reported history-based memory retention. Conversely, the remaining 6 studies did not present any method for maintaining context.

Table . Large language model (LLM)–specific design characteristics of included studies[a,b].

Characteristics	Number of studies
LLM profile (multiple models possible)	
ChatGPT 3.5	1 [30]
ChatGPT 3.5 Turbo	1 [29]
ChatGPT 4.0	5 [31-34,36]
Gemini	1 [32]
Llama 3.1	1 [35]
Mistral 7B	1 [35]
PHI 3.5	1 [35]
RAG[c]	
Applied	5 [30,32,33,35,36]
Not reported	3 [29,31,34]
Knowledge source	
Clinical guideline	3 [30,32,33]
Educational material	1 [33]
Published literature	1 [35]
Textbook-based sources	1 [30]
Not reported	3 [29,31,34]
Development framework	
LangChain	3 [30,32,36]
Knowledge graph	1 [32]
Not reported	5 [29,31,33-35]
Conversational memory handling	
Embedding-based memory retrieval	1 [29]
History-based memory retention	1 [34]
Not reported	6 [30-33,35,36]

[a]"Not reported" indicates that the item was not explicitly described in the study.

[b]Study [35] evaluated 3 large language models comparatively, and all models assessed are listed.

[c]RAG: retrieval-augmented generation.

Primary Purposes of LLM-Based Chatbots

Analysis of the primary purposes of the developed chatbots revealed 4 categories: emotional support, information provision, enhancing user convenience, and analyzing user experience. Three studies focused on emotional support as their primary purpose, aiming to provide psychological support centered on reducing anxiety, depression, and stress. Three studies focused on information provision as their primary purpose, developing chatbots that provide personalized responses to cancer-related questions and deliver medical information.

Meanwhile, 1 study primarily aimed to support clinical research by automating the collection and organization of patient-reported outcomes (PRO) data. Another study primarily focused on exploring the interaction experience with conversational AI support agents, targeting caregivers as the main subjects.

Overall, the included studies used LLM-based chatbots as tools for information provision or emotional support. Some studies focused on enhancing clinical convenience or understanding user experience and interaction patterns.

Evaluation Approaches and Outcome Domains

Analysis of evaluation methods and domains revealed that the assessment approaches and measured outcome domains varied significantly across studies (Table 3). One study evaluated participants' psychological outcomes using validated measurement tools, while another reported psychological states

XSL•FO
RenderX

using a predefined scoring scale. Additionally, 1 study focused its evaluation on user experience and perceptions.

Four studies evaluated the response performance of LLM-based chatbots, focusing on chatbot response accuracy, completeness, comprehension, consistency, and security or safety. One study focused on chatbot development and did not report separate evaluation results.

The studies included in this literature review were evaluated across diverse domains such as clinical outcomes, response quality, and user experience, revealing heterogeneous characteristics in the evaluation design and outcome domains. However, despite being multiturn interactive systems, no studies were identified that evaluated the continuity of conversation or the interaction process itself.

Table . Evaluation methods and outcome measures used in included studies.

Characteristics	Number of studies
General evaluation outcomes	
Psychological outcomes (validated instruments)	1 [31]
User-reported psychological states (nonvalidated measures)	1 [34]
User experience and perception	1 [36]
LLM[a]-specific evaluation outcomes	
Response accuracy	4 [29,31,33,35]
Response completeness	2 [29,33]
Response comprehension	3 [29,30,33]
Response consistency	1 [29]
Response security	2 [29,35]
No formal evaluation	1 [32]

[a]LLM: large language model.

Reported Safety Risks and Mitigation Strategies

Analysis of safety risks and mitigation strategies reported in the included studies revealed that the level of safety-related reporting varied across studies (Table 4). Risks related to the reliability of chatbot responses were reported in 4 studies, while risks related to physical and psychological safety were identified in 1 study each. Conversely, 3 studies did not provide specific reports on safety risk factors.

Two studies specified expert intervention as a strategy to mitigate safety risks. Two studies applied an automated escalation system that provided emergency contacts to participants or connected them to external support systems during crisis situations. Additionally, 1 study applied a strategy to exclude models with low performance or potential harm. However, the remaining 3 studies did not report on risk mitigation strategies. While some studies proposed safety management strategies, systematic reporting of safety risk factors and mitigation strategies was limited.

Table . Reported safety risks and mitigation strategies in included studies.

Characteristics	Number of studies
Safety risks	
Response reliability	4 [29,31,34,35]
Psychological safety risk	1 [34]
Physical safety risk	1 [32]
Not reported	3 [30,33,36]
Mitigation strategies	
Expert involvement	2 [31,34]
Escalation systems	2 [32,34]
Removal of poorly performing models	1 [35]
Not reported	3 [30,33,36]

XSL•FO
RenderX

Transparency and Reproducibility of LLM-Specific Components

Analysis of the reporting level for LLM-related components revealed that transparency in prompt design and system configuration varied across studies (Table 5). No study reported prompts in full, while 5 studies reported only partial components. Conversely, 3 studies did not report prompt design at all.

Only 1 study explicitly reported the system prompt as a prompt component, 2 reported the user prompt, and 2 reported model parameters. However, 3 studies did not report specific information about prompt components. Furthermore, only 1 study explicitly referenced LLM-based research reporting guidelines, while the remaining 7 studies made no mention of such guidelines.

Overall, the included studies provided basic information on LLM selection and use, but reporting on reproducibility aspects, such as prompt design and system configuration, was limited. The research landscape across the 5 research questions is summarized in Figure 2.

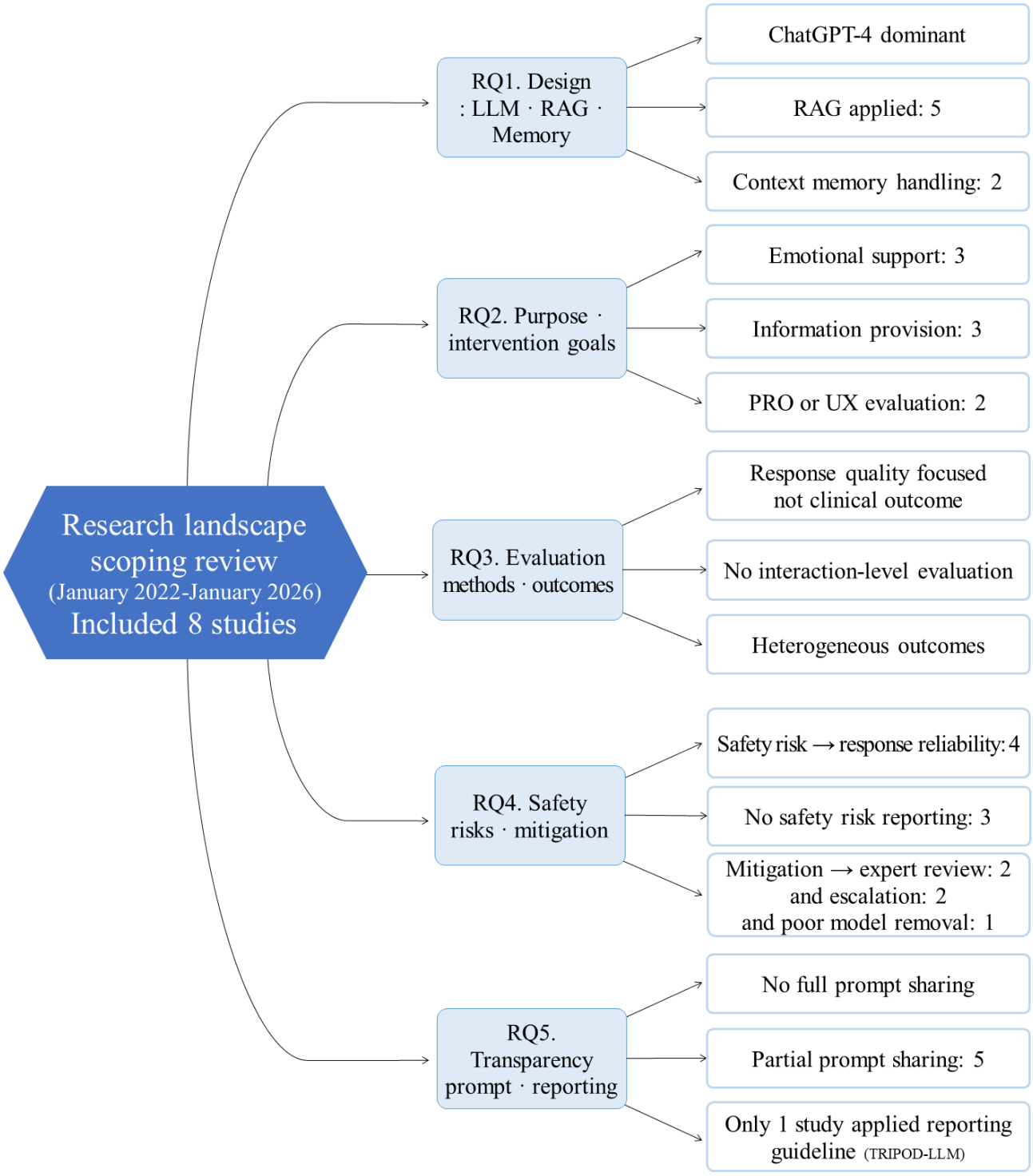
Table . Transparency and reproducibility of large language model (LLM)–specific components.

Characteristics	Number of studies
Prompt reporting completeness	
Complete	0
Partial	5 [29,30,33-35]
Not reported	3 [31,32,36]
Prompt components reported	
System prompt	1 [33]
User prompt	2 [29,34]
Model parameters	2 [30,35]
Not reported	3 [31,32,36]
Reporting guideline adherence	
TRIPOD-LLM[a]	1 [33]
Not reported	7 [29-32,34-36]

[a]TRIPOD-LLM: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Large Language Models.

XSL•FO
RenderX

Figure 2. Conceptual map of the research landscape across 5 research questions (RQs). LLM: large language model; PRO: patient-reported outcome; RAG: retrieval-augmented generation; TRIPOD-LLM: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Large Language Models; UX: user experience.



Discussion

Principal Findings

The most significant finding of this review is the immaturity of the multiturn chatbot field in oncology. Multiturn interactions are often regarded as an inherent and fundamental feature of conversational agents. Nevertheless, studies that could be rigorously classified as genuinely multiturn were markedly limited. This scoping review identified 8 studies, of which only 2 explicitly reported any conversational memory mechanism. Most studies included in this review remained at the stage of system development and prototyping, with limited research applying comparative designs or validating clinical efficacy. While systematic assessments of patient-centered clinical outcomes were relatively scarce, most of the studies focused on response quality and safety verification.

In terms of design, ChatGPT-based models were primarily used, with some studies applying RAG techniques or external

XSL•FO
RenderX

knowledge integration methods. The level of reporting on key design elements—such as knowledge sources, retrieval architecture, model configuration, and prompt composition—varied significantly across studies. Furthermore, reported safety risk factors, mitigation strategies, and transparency of LLM components were inconsistent across studies. Only 1 study explicitly referenced reporting guidelines.

Research on LLM-based oncology multiturn conversational chatbots is rapidly expanding but has not yet reached maturity. Key areas requiring further development include clarity of design architecture, validation of clinical efficacy, safety management strategies, and reporting standardization.

Design and Intervention Characteristics of LLM-Based Oncology Chatbots

Overall, the included studies focused on prototype-level implementations based on commercial LLMs, exhibiting heterogeneity in model selection, knowledge integration strategies, and system design approaches. This design diversity can be interpreted as reflecting the exploratory application phase of the technology [37]. However, such diversity in chatbot design also complicates comparability between studies [20,38].

In oncology, LLM-based chatbots were often developed for emotional support and information provision. Some studies presented automation of collecting and organizing PRO data or exploring user experience and interaction patterns as primary objectives. This aligns with prior research indicating that LLM-based chatbots in oncology are applied for education and patient support [39,40].

Meanwhile, studies specifying a particular cancer type were primarily focused on breast cancer, and several studies targeted patients in the early stages of diagnosis. These patterns suggest that LLM-based chatbot interventions are initially being applied to patient groups with relatively standardized clinical guidelines and high informational and emotional support needs [13,41]. However, this bias toward specific cancer types and disease stages limits generalizability across oncology. Furthermore, despite being premised on multiturn interactions, studies explicitly addressing conversational memory handling were limited.

Most studies identified in this literature review focused on individual response-level performance metrics, such as chatbot response accuracy or completeness. No studies were found that evaluated the multiturn interaction process itself. Multiturn dialogue systems generate responses by considering previous utterances and conversation history. They are reported as a crucial technical element for providing sophisticated information and building psychological rapport by posing follow-up questions in real time [42,43]. Furthermore, the ability to generate consistent responses while maintaining long-term conversational context is presented as a major research challenge for LLM-based conversational systems [18]. Therefore, evaluating interaction-level factors, such as conversational continuity and context retention capability, may also be important considerations.

Evaluation, Safety, and Transparency of LLM-Based Conversational Agents

Included studies tended to focus on response quality and safety verification, with limited comparative designs evaluating clinical efficacy. Furthermore, while some studies mentioned potential psychological risks, safety reporting was limited in a significant number of studies, and specific safety threat factors were not explicitly reported. Additionally, while evaluating the “quality” and “stability” of interactions is crucial for multisession interactive chatbots, reporting on this aspect was also limited.

LLMs are generative models based on probabilistic next-word prediction engines that can produce plausible-looking but factually inconsistent hallucinations [44,45]. Response accuracy varies across studies and models, but reported accuracy levels around 80% still imply the potential for generating inaccurate information [46-48]. Particularly, cancer treatment requires a multidisciplinary approach and complex decision-making. Therefore, the accuracy and reliability of information provided by LLM-based chatbots can significantly impact the decision-making and treatment processes of patients with cancer [24,49,50].

Mitigation strategies to reduce hallucinations and improve accuracy, such as RAG or knowledge graphs that connect LLMs with external evidence sources, have been proposed. At the clinical level, some studies proposed expert prereview as a safety measure [31,34] or introduced escalation mechanisms connecting to external support systems during crisis situations [34]. However, these mitigation strategies were not consistently applied or systematically evaluated across studies, and many studies did not sufficiently describe specific risk management strategies. Therefore, establishing a standardized framework to systematically define and report safety threats in LLM-based conversational chatbots is required.

With respect to transparency, reporting on LLM-related components was generally limited across the included studies, as noted in existing literature [20,51,52]. While some studies partially disclosed prompts, key elements such as system prompts, user prompts, and model parameters were often insufficiently described. Prompt engineering, the process of adjusting LLM responses through the input structure and instruction design, impacts response accuracy [53-56]. Furthermore, since LLMs respond sensitively to prompt structure without revealing their internal reasoning processes, prompt design is a critical variable in determining response appropriateness in medical settings [21,57-59]. Therefore, the failure to clearly report prompt structure and model settings may limit the interpretability and reproducibility of research findings [60].

Notably, only 1 included study explicitly referenced a reporting guideline—specifically, TRIPOD-LLM [21]—underscoring the limited adoption of standardized reporting frameworks in this emerging field. Standardized and transparent reporting is essential for building trust in the use of generative AI models in clinical practice [22]. Reporting guidelines such as TRIPOD-LLM, CHART, and CONSORT-AI have been developed to address this gap [21,22,61]. Future studies on

XSL•FO
RenderX

LLM-based chatbots in oncology should explicitly adhere to these frameworks.

Implications for Future Research and Practice

Future research should report model versions, knowledge update cycles, retrieval strategies, prompt systems, and system pipelines in standardized formats to enhance comparability and reproducibility across studies. For systems designed for multiturn interactions, it is necessary to describe conversational state management and context retention (conversational memory handling) strategies and apply interaction-level evaluation metrics such as dialogue consistency and safety. Furthermore, the applicability across diverse cancer types, disease stages, and user groups (patients and caregivers) should be systematically validated, and comparative designs enabling clinical efficacy assessment should be expanded. Finally, adherence to established reporting guidelines is strongly recommended to promote transparency and reproducibility in future studies.

Limitations

This study has the following limitations. First, the search was restricted to English-language publications in 8 databases, so the possibility of omitting relevant studies cannot be ruled out. Second, due to the rapidly evolving nature of this research field, conference papers were included; however, some studies did not sufficiently report detailed methods and components, limiting interpretation. Third, consistent with the nature of a scoping review, methodological quality assessments of individual studies were not performed, preventing a systematic comparison and evaluation of the evidence levels across the included studies. Fourth, given the rapid pace of LLM research, studies published after the search cutoff (February 2026 for the primary databases and May 2026 for the supplementary AI or computer science databases) may not be captured in this review, and periodic updates are recommended to reflect the latest developments in this field.

Conclusion

This scoping review systematically summarized the overall research landscape of LLM-based multiturn conversational chatbot studies developed for patients with cancer and their informal caregivers. Only 8 studies were included, most of which were recently reported prototype-focused development research, reflecting the emerging nature of this field. Heterogeneity among studies was found in design architecture, evaluation methods, safety management strategies, and the level of transparency. Notably, despite the conversational nature of the included systems, reporting on interaction-level evaluations, such as context retention strategies, was limited.

LLM-based oncology conversational chatbots show potential as promising tools for patients with cancer and caregivers. Genuine multiturn conversational capability should therefore be prioritized in future development. However, safety management and standardization of design and reporting must proceed concurrently. Future research should explicitly identify potential patient safety risks and propose corresponding mitigation strategies. Furthermore, LLM-based chatbot development in oncology should explicitly document prompt design and system configuration in accordance with validated reporting guidelines.

Acknowledgments

The authors declare the use of generative artificial intelligence (Claude Sonnet 4.6, Anthropic) for language editing during manuscript preparation. Responsibility for the final manuscript lies entirely with the authors.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search string.
[DOCX File, 30 KB - cancer_v12i1e96241_app1.docx]

Multimedia Appendix 2
Full-text screening decision at full-text screening.
[DOCX File, 18 KB - cancer_v12i1e96241_app2.docx]

Multimedia Appendix 3
Reasons for exclusion during title or abstract screening (counts not mutually exclusive).
[DOCX File, 14 KB - cancer_v12i1e96241_app3.docx]

Checklist 1
PRISMA checklist.

XSL•FO
RenderX

[PDF File, 170 KB - cancer_v12i1e96241_app4.pdf]

References

1. Chang Y, Wang X, Wang J, et al. A survey on evaluation of large language models. ACM Trans Intell Syst Technol 2024 Jun 30;15(3):1-45. [doi: 10.1145/3641289]
2. Clusmann J, Kolbinger FR, Muti HS, et al. The future landscape of large language models in medicine. Commun Med (Lond) 2023 Oct 10;3(1):141. [doi: 10.1038/s43856-023-00370-1] [Medline: 37816837]
3. Lu Z, Peng Y, Cohen T, Ghassemi M, Weng C, Tian S. Large language models in biomedicine and health: current research landscape and future directions. J Am Med Inform Assoc 2024 Sep 1;31(9):1801-1811. [doi: 10.1093/jamia/ocae202] [Medline: 39169867]
4. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med 2023 Aug;29(8):1930-1940. [doi: 10.1038/s41591-023-02448-8] [Medline: 37460753]
5. Tudor Car L, Dhinagaran DA, Kyaw BM, et al. Conversational agents in health care: scoping review and conceptual analysis. J Med Internet Res 2020 Aug 7;22(8):e17158. [doi: 10.2196/17158] [Medline: 32763886]
6. Huang L, Yu W, Ma W, et al. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. ACM Trans Inf Syst 2025 Mar 31;43(2):1-55. [doi: 10.1145/3703155]
7. Omiye JA, Gui H, Rezaei SJ, Zou J, Daneshjou R. Large language models in medicine: the potentials and pitfalls: a narrative review. Ann Intern Med 2024 Feb;177(2):210-220. [doi: 10.7326/M23-2772] [Medline: 38285984]
8. Haltaufderheide J, Ranisch R. The ethics of ChatGPT in medicine and healthcare: a systematic review on Large Language Models (LLMs). NPJ Digit Med 2024 Jul 8;7(1):183. [doi: 10.1038/s41746-024-01157-x] [Medline: 38977771]
9. Choi R, Kim T, Park S, Kim JG, Lee SJ. Private yet social: how LLM chatbots support and challenge eating disorder recovery. Presented at: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems; Apr 26 to May 1, 2025. [doi: 10.1145/3706598.3713485]
10. Heinz MV, Mackin DM, Trudeau BM, et al. Randomized trial of a generative AI chatbot for mental health treatment. NEJM AI 2025 Mar 27;2(4). [doi: 10.1056/AIoa2400802]
11. Jeon S, Lee S, Kim EH, et al. Generative AI chatbot for diabetes management: formative 2-part qualitative study using DTalksBot involving patients and clinicians. JMIR Form Res 2025 Nov 12;9:e72553. [doi: 10.2196/72553] [Medline: 41223424]
12. Lewandowska A, Rudzki G, Lewandowski T, Rudzki S. The problems and needs of patients diagnosed with cancer and their caregivers. Int J Environ Res Public Health 2020 Dec 24;18(1):87. [doi: 10.3390/ijerph18010087] [Medline: 33374440]
13. Osowiecka K, Rucińska M, Abe T, et al. Cancer related information that cancer patients need. Sci Rep 2025 May 6;15(1):15811. [doi: 10.1038/s41598-025-99498-4] [Medline: 40328851]
14. Stiefel F, Bourquin C, Salmon P, et al. Communication and support of patients and caregivers in chronic cancer care: ESMO Clinical Practice Guideline. ESMO Open 2024 Jul;9(7):103496. [doi: 10.1016/j.esmoop.2024.103496] [Medline: 39089769]
15. Thorne S, Hislop TG, Kim-Sing C, Oglov V, Oliffe JL, Stajduhar KI. Changing communication needs and preferences across the cancer care trajectory: insights from the patient perspective. Support Care Cancer 2014 Apr;22(4):1009-1015. [doi: 10.1007/s00520-013-2056-4] [Medline: 24287506]
16. Tu T, Schaekermann M, Palepu A, et al. Towards conversational diagnostic artificial intelligence. Nature 2025 Jun;642(8067):442-450. [doi: 10.1038/s41586-025-08866-7] [Medline: 40205050]
17. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 1;25(9):1248-1258. [doi: 10.1093/jamia/ocy072] [Medline: 30010941]
18. Yi Z, Ouyang J, Xu Z, et al. A survey on recent advances in LLM-based multi-turn dialogue systems. ACM Comput Surv 2026 Apr 30;58(6):1-38. [doi: 10.1145/3771090]
19. Pichowicz W, Kotas M, Piotrowski P. Performance of mental health chatbot agents in detecting and managing suicidal ideation. Sci Rep 2025 Aug 27;15(1):31652. [doi: 10.1038/s41598-025-17242-4] [Medline: 40866537]
20. Huo B, Boyle A, Marfo N, et al. Large language models for chatbot health advice studies: a systematic review. JAMA Netw Open 2025 Feb 3;8(2):e2457879. [doi: 10.1001/jamanetworkopen.2024.57879] [Medline: 39903463]
21. Gallifant J, Afshar M, Ameen S, et al. The TRIPOD-LLM reporting guideline for studies using large language models. Nat Med 2025 Jan;31(1):60-69. [doi: 10.1038/s41591-024-03425-5] [Medline: 39779929]
22. CHART Collaborative, Huo B, Collins GS, et al. Reporting guideline for chatbot health advice studies: the CHART statement. JAMA Netw Open 2025 Aug 1;8(8):e2530220. [doi: 10.1001/jamanetworkopen.2025.30220] [Medline: 40747871]
23. Wang A, Qian Z, Briggs L, Cole AP, Reis LO, Trinh QD. The use of chatbots in oncological care: a narrative review. Int J Gen Med 2023;16:1591-1602. [doi: 10.2147/IJGM.S408208] [Medline: 37152273]
24. Chen D, Avison K, Alnassar S, Huang RS, Raman S. Medical accuracy of artificial intelligence chatbots in oncology: a scoping review. Oncologist 2025 Apr 4;30(4):oyaf038. [doi: 10.1093/oncolo/oyaf038] [Medline: 40285677]
25. Jiang XH, Yuan XH, Zhao H, Peng JS. The effectiveness and feasibility of conversational agents in supporting care for patients with cancer: systematic review and meta-analysis. J Med Internet Res 2025 Aug 8;27:e76968. [doi: 10.2196/76968] [Medline: 40736462]

XSL•FO
RenderX

26. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
27. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
28. Jeong Y, Cha H, Suh E. LLM-based conversational agents for patients with cancer and caregivers: a scoping review. JMIR Preprints. Preprint posted online on Feb 5, 2026. [doi: 10.2196/preprints.96241]
29. Chen Z, Wang Q, Sun Y, Cai H, Lu X. Chat-ePRO: development and pilot study of an electronic patient-reported outcomes system based on ChatGPT. J Biomed Inform 2024 Jun;154:104651. [doi: 10.1016/j.jbi.2024.104651] [Medline: 38703936]
30. Lee JW, Yoo IS, Kim JH, et al. Development of AI-generated medical responses using the ChatGPT for cancer patients. Comput Methods Programs Biomed 2024 Sep;254:108302. [doi: 10.1016/j.cmpb.2024.108302] [Medline: 38996805]
31. Akdogan O, Uyar GC, Yesilbas E, et al. Effect of a ChatGPT-based digital counseling intervention on anxiety and depression in patients with cancer: a prospective, randomized trial. Eur J Cancer 2025 May 15;221:115408. [doi: 10.1016/j.ejca.2025.115408] [Medline: 40215593]
32. Bharambe U, Patil K, Ingle P, Bhangale U. Exploring opportunities and addressing challenges in designing knowledge graph-enhanced RAG-based chatbots for managing radiation toxicity in cancer care. Presented at: 2025 International Conference for Artificial Intelligence, Applications, Innovation and Ethics (AI2E); Feb 3-5, 2025; Muscat, Oman p. 1-6. [doi: 10.1109/AI2E64943.2025.10983078]
33. Boie SD, Glastetter E, Lux MP, et al. Evaluating a chatbot as a companion for patients with breast cancer: collaborative pilot study. JMIR Cancer 2025 Aug 13;11:e68426. [doi: 10.2196/68426] [Medline: 40802992]
34. Hasei J, Hanzawa M, Nagano A, et al. Empowering pediatric, adolescent, and young adult patients with cancer utilizing generative AI chatbots to reduce psychological burden and enhance treatment engagement: a pilot study. Front Digit Health 2025;7:1543543. [doi: 10.3389/fdgth.2025.1543543] [Medline: 40070545]
35. McInerney S, Nash T, Lee R, Falis M, Gruber F, Casey A. AI chatbot for cancer patient support: development and evaluation using Llama 3.1, Mistral 7B, and PHI 3B. Stud Health Technol Inform 2025 May 15;327:890-891. [doi: 10.3233/SHTI250494] [Medline: 40380602]
36. Sugan M, Nadkarni S, Sanaeipoor S, et al. Exploring caregivers' acceptance of conversational AI in pediatric cancer caregiving: a mixed-methods study. Presented at: Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems; Apr 26 to May 1, 2025; Yokohama, Japan p. 1-8. [doi: 10.1145/3706599.3721270]
37. Carl N, Schramm F, Haggenmüller S, et al. Large language model use in clinical oncology. NPJ Precis Oncol 2024 Oct 23;8(1):240. [doi: 10.1038/s41698-024-00733-4] [Medline: 39443582]
38. Lai X, Lai Y, Chen J, Huang S, Gao Q, Huang C. Evaluation strategies for large language model-based models in exercise and health coaching: scoping review. J Med Internet Res 2025 Oct 14;27:e79217. [doi: 10.2196/79217] [Medline: 41086432]
39. Chen D, Parsa R, Swanson K, et al. Large language models in oncology: a review. BMJ Oncol 2025;4(1):e000759. [doi: 10.1136/bmjonc-2025-000759] [Medline: 40519217]
40. Raynaud C, Wu D, Levy J, Marengo M, Bibault JE. Patients facing large language models in oncology: a narrative review. JCO Clin Cancer Inform 2024 Nov;8:e2400149. [doi: 10.1200/CCI-24-00149] [Medline: 39514825]
41. Priambodo AP, Trisyani Y, Nuraeni A, Anna A, Sugiharto F. The information needs of breast cancer patients at all stages of their journey: a scoping review. Patient Prefer Adherence 2026;20:578182. [doi: 10.2147/PPA.S578182] [Medline: 41783122]
42. Ning B, Zhao D, Liu X, Li G. EAGS: an extracting auxiliary knowledge graph model in multi-turn dialogue generation. World Wide Web 2022 Sep 30:1-22. [doi: 10.1007/s11280-022-01100-8] [Medline: 36196376]
43. Vatsal S, Dubey H, Singh A. Agentic AI in healthcare and medicine: a seven-dimensional taxonomy for empirical evaluation of LLM-based agents. IEEE Access 2026;14:4840-4863. [doi: 10.1109/ACCESS.2026.3651218]
44. Howell MD. Generative artificial intelligence, patient safety and healthcare quality: a review. BMJ Qual Saf 2024 Oct 18;33(11):748-754. [doi: 10.1136/bmjqs-2023-016690] [Medline: 39048292]
45. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel) 2023 Mar 19;11(6):887. [doi: 10.3390/healthcare11060887] [Medline: 36981544]
46. Hao Y, Qiu Z, Holmes J, et al. Large language model integrations in cancer decision-making: a systematic review and meta-analysis. NPJ Digit Med 2025 Jul 17;8(1):450. [doi: 10.1038/s41746-025-01824-7] [Medline: 40676129]
47. Kaboudi N, Firouzbakht S, Shahir Eftekhar M, et al. Diagnostic accuracy of ChatGPT for patients' triage; a systematic review and meta-analysis. Arch Acad Emerg Med 2024;12(1):e60. [doi: 10.22037/aaem.v12i1.2384] [Medline: 39290765]
48. Deng J, Lin Y. The benefits and challenges of ChatGPT: an overview. Front Comput Intell Syst 2023;2(2):81-83. [doi: 10.54097/fcis.v2i2.4465]
49. Beeram M, Kennedy A, Hales N. Barriers to comprehensive multidisciplinary head and neck care in a community oncology practice. Am Soc Clin Oncol Educ Book 2021 Mar;41:1-10. [doi: 10.1200/EDBK_320967] [Medline: 34010055]
50. Guo Z, Lai A, Thygesen JH, Farrington J, Keen T, Li K. Large language models for mental health applications: systematic review. JMIR Ment Health 2024 Oct 18;11:e57400. [doi: 10.2196/57400] [Medline: 39423368]

XSL•FO
RenderX

51. Ko JS, Heo H, Suh CH, Yi J, Shim WH. Adherence of studies on large language models for medical applications published in leading medical journals according to the MI-CLEAR-LLM checklist. Korean J Radiol 2025 Apr;26(4):304-312. [doi: 10.3348/kjr.2024.1161] [Medline: 40015560]
52. Ibrahim H, Liu X, Zariffa N, Morris AD, Denniston AK. Health data poverty: an assailable barrier to equitable digital health care. Lancet Digit Health 2021 Apr;3(4):e260-e265. [doi: 10.1016/S2589-7500(20)30317-4] [Medline: 33678589]
53. Liu P, Yuan W, Fu J, Jiang Z, Hayashi H, Neubig G. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing. ACM Comput Surv 2023 Sep 30;55(9):1-35. [doi: 10.1145/3560815]
54. Patil R, Heston TF, Bhuse V. Prompt engineering in healthcare. Electronics (Basel) 2024;13(15):2961. [doi: 10.3390/electronics13152961]
55. Russe MF, Reisert M, Bamberg F, Rau A. Improving the use of LLMs in radiology through prompt engineering: from precision prompts to zero-shot learning. Rofo 2024 Nov;196(11):1166-1170. [doi: 10.1055/a-2264-5631] [Medline: 38408477]
56. Shah K, Xu AY, Sharma Y, et al. Large language model prompting techniques for advancement in clinical medicine. J Clin Med 2024 Aug 28;13(17):5101. [doi: 10.3390/jcm13175101] [Medline: 39274316]
57. Harrer S. Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. EBioMedicine 2023 Apr;90:104512. [doi: 10.1016/j.ebiom.2023.104512] [Medline: 36924620]
58. Meskó B. Prompt engineering as an important emerging skill for medical professionals: tutorial. J Med Internet Res 2023 Oct 4;25:e50638. [doi: 10.2196/50638] [Medline: 37792434]
59. Xu L, Sanders L, Li K, Chow JCL. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR Cancer 2021 Nov 29;7(4):e27850. [doi: 10.2196/27850] [Medline: 34847056]
60. Park SH, Suh CH, Lee JH, et al. Minimum reporting items for CLEAR evaluation of accuracy reports of large language models in healthcare (MI-CLEAR-LLM): 2025 updates. Korean J Radiol 2025 Dec;26(12):1123-1132. [doi: 10.3348/kjr.2025.1522] [Medline: 41199132]
61. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020 Sep;26(9):1364-1374. [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]

Abbreviations

CHART: Chatbot Assessment Reporting Tool
JBI: Joanna Briggs Institute
LLM: large language model
NLP: natural language processing
OSF: Open Science Framework
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
PRO: patient-reported outcome
RAG: retrieval-augmented generation
TRIPOD-LLM: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Large Language Models

Edited by M Balcarras; submitted 26.Mar.2026; peer-reviewed by D Shin, Y Zhang; revised version received 31.May.2026; accepted 17.Jun.2026; published 21.Jul.2026.

Please cite as:
Jeong Y, Cha H, Suh EE
Multiturn Large Language Model–Based Conversational Agents for Patients With Cancer and Caregivers: Scoping Review
JMIR Cancer 2026;12:e96241
URL: https://cancer.jmir.org/2026/1/e96241
doi:10.2196/96241

© Yeongrok Jeong, Hyejeon Cha, Eunyoung E Suh. Originally published in JMIR Cancer (https://cancer.jmir.org), 21.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Interventions to Reduce Fear of Cancer Recurrence Among People With Cancer: Scoping Review

Niu Li[1], MD; Cynthia Wensley[1], PhD; Lisa Reynolds[2], PhD; Wal Baraza[3,4], MPhil, FRCS; Tatiana Osorio Leyton[1], MD; Andrew Jull[1,5], PhD

[1]School of Nursing, Faculty of Medicine and Health Science, University of Auckland, 85 Park Road, Grafton, Auckland, New Zealand

[2]Department of Psychological Medicine, University of Auckland, Auckland, New Zealand

[3]Department of Surgery, Faculty of science, University of Auckland, Auckland, New Zealand

[4]Department of Surgery, Auckland City Hospital, Auckland, New Zealand

[5]Centre for Translational Research in Health, University of Auckland, Auckland, New Zealand

Corresponding Author:
Niu Li, MD
School of Nursing, Faculty of Medicine and Health Science, University of Auckland, 85 Park Road, Grafton, Auckland, New Zealand

Abstract

Background: Fear of cancer recurrence (FCR) is prevalent among cancer survivors, affecting between 39% and 97% of patients. FCR is associated with impaired concentration, sleep disturbances, decreased quality of life, and increased psychological distress and health care use. To date, the literature lacks a review that summarizes the breadth of psychological interventions available for reducing fear of recurrence.

Objective: This review aims to identify and summarize the evidence on psychological interventions for addressing FCR across all cancers.

Methods: The Joanna Briggs Institute method for scoping reviews guided the processes, and we reported the review following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist. We searched 5 databases (CINAHL, PsycInfo, the Cochrane Central Register of Controlled Trials, Embase, and MEDLINE) and 2 gray literature sources (ProQuest Dissertations & Theses Global and the World Health Organization International Clinical Trials Registry). Eligible studies included adults (≥18 years) diagnosed with cancer and evaluated psychological interventions aimed at reducing FCR. Data extraction captured study characteristics, intervention details, outcome effectiveness, and follow-up durations. We synthesized the findings using descriptive summaries and narrative analysis.

Results: Overall, 5131 articles were screened, and 122 were included in this review; 48 (39.3%) involved patients with breast cancer, 47 (38.5%) focused on patients with multiple cancer types; over half of the studies (n=64, 52.5%) were randomized controlled trials. Only 28 (23%) studies explicitly reported the definition of FCR. Eighteen different measurement tools were used. Blended interventions (different combinations of cognitive behavioral therapy, mindfulness, acceptance and commitment therapy, and other strategies) formed the largest intervention category (n=38, 31.1%), followed by cognitive behavioral therapy interventions (n=26, 21.3%) and mindfulness-based interventions (n=24, 19.7%). Of the included studies, 104 (85.2%) demonstrated significant reductions in FCR. Most interventions were delivered face-to-face by disciplinary specialists (n=75, 61.5%), while some were delivered remotely (n=34, 27.9%), with the majority of these delivered via the website (n=18, 52.9%). Follow-up duration ranged from postintervention to 3 years.

Conclusions: FCR has been the focus of an increasing number of studies since 2009, with the majority being randomized controlled trials. Most interventions are delivered face-to-face and rely on trained specialists. Most have had statistically significant results. However, the included studies demonstrated heterogeneity in terms of delivery, duration, and dose, requiring cautious interpretation of intervention effects. Future research should develop consistent guidelines to standardize the definition of FCR, the measurement tools used, and the timing of follow-up assessments. Long-term follow-up data are needed to evaluate the sustained effects.

Trial Registration: OSF Registries 10.17605/OSF.IO/R5C43; https://osf.io/10.17605/OSF.IO/R5C43

(JMIR Cancer 2026;12:e81579) doi:10.2196/81579

KEYWORDS

fear of cancer recurrence; fear of progression; cancer survivor; psychological interventions; cancer survivors; scoping review; mobile health interventions; mental health; digital health intervention

XSL•FO
RenderX

Introduction

In 1981, Northouse first introduced the concept of fear of cancer recurrence (FCR) to describe the concerns of cancer patients in remission who feared that their cancer might return [1]. FCR is defined as fear, worry, or concern about the possibility of cancer returning [2]. FCR arises when cancer-related cues or bodily sensations are interpreted as signs of threat. These threat appraisals, combined with difficulties managing distress, can lead to patterns of hypervigilance, worry, reassurance seeking, or avoidance [3-5]. In addition, the idea that staying alert prevents recurrence further maintains persistent monitoring and rumination. Ongoing uncertainty about symptoms or the course of illness can amplify anxiety and contribute to the persistence of FCR over time [5,6].

Previous studies report that between 39% and 97% of cancer survivors experience FCR, with 22% to 87% experiencing moderate to high levels of fear [7]. This wide range may be due to heterogeneity in cancer types and survivorship stages, differences in measurement instruments, and variation in the timing of assessment [7]. It is common for cancer survivors to experience short-term or low levels of FCR. In some cases, such fear may be protective by increasing vigilance to bodily symptoms and promoting positive health behaviors, such as attending regular follow-up appointments and maintaining a healthy lifestyle [8]. However, when it becomes persistent and severe, it can interfere with a person's mental health and daily functioning [9,10]. FCR may disrupt concentration, decision-making, sleep, and social functioning [11,12]. It also decreases overall well-being and quality of life and increases psychological distress and resistance to follow-up appointments [13-15]. High levels of FCR are significantly associated with a range of psychological comorbidities, including depression, anxiety, impaired emotional functioning, and insomnia [6]. FCR is linked to frequent checkups and additional examinations [16], and high levels can increase health care use, particularly in primary care [17,18]. FCR may also significantly delay return to work or even lead to premature withdrawal from employment [19]. If untreated, FCR can continue unabated throughout a patient's survival journey [20,21]. In a large cohort study, 44% - 56% of patients continued to report clinical levels of FCR up to 18 months post surgery [22].

To date, only one scoping review has summarized psychological interventions for FCR [23]. However, it was solely focused on cognitive behavioral–based approaches [23]. In our review, we aim to map studies of the psychological interventions designed to reduce FCR in people with cancer, detailing the study characteristics, intervention details, outcome measurement, and follow-up durations to provide an overview of the interventions and highlight evidence gaps. These insights could inform the development of potential strategies to address FCR in specific populations.

Methods

Study Design

The review followed the Joanna Briggs Institute method for scoping reviews [24] and was reported by the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist [25]. We preregistered the scoping review protocol on the Open Science Framework (registration number: 29YA3) [26].

Review Objective

The primary objective of this scoping review is to map and summarize the evidence on the effects of psychological interventions designed to reduce FCR in people with cancer.

The scoping review addressed the following questions:

1. How has FCR been defined?
2. What instruments have been used to measure FCR?
3. What psychological interventions have been investigated to reduce FCR?
4. What are the intervention characteristics (types, setting, duration of intervention content, mode of delivery, personnel involved, and number of sessions)?
5. What is the effect of interventions in reducing FCR?
6. What is the follow-up duration of interventions addressing FCR?

Eligibility Criteria

We used the population/participant-concept-context framework [24] to define eligibility.

- Population: studies focused on FCR in all people who have been histologically and clinically diagnosed with cancer, aged 18 years or older at diagnosis.
- Concept: any type of intervention with a psychological component aiming to reduce FCR. There were no limitations on the intervention's duration, frequency, delivery method, or type of provider.
- Context: interventions provided in any type of setting.

Type of Evidence Sources

We included any quantitative study design that reported data relevant to the research questions, including randomized controlled trials (RCTs), quasi-experimental studies, nonrandomized intervention studies, observational studies, and mixed methods studies with extractable intervention data. Review articles were excluded to avoid data duplication, but the reference lists were screened for relevant studies. We also excluded ongoing trials and protocol papers.

Search Strategy

The search strategy was initially focused on colorectal cancer, but due to an insufficient number of studies, it was broadened to include all people with cancer. The search strategy was developed based on three key concepts: (1) neoplasms, including synonyms such as neoplas*, cancer, carcinom*, malignant*, tumor*, and oncolog*; (2) fear, with related terms including anxiety, worr*, concern, and distress; and (3) recurrence, with associated terms such as relapse, progress*, reappearance, and return. These terms were combined using the Boolean operator "OR" within each concept group and "AND" between the concept groups. However, the "fear" and "recurrence" concepts were combined using a proximity operator (*adj4*) in most databases to ensure co-occurrence within a 4-word context. The research team designed the search strategy in collaboration with

XSL•FO
RenderX

an experienced librarian, and it was adapted for each database. We searched the reference lists of all studies included in the review to identify additional studies.

We searched 5 electronic databases: CINAHL, PsycInfo, the Cochrane Central Register of Controlled Trials (CENTRAL), Embase, and MEDLINE. To capture relevant gray literature, we also searched 2 gray literature sources: ProQuest Dissertations & Theses Global and the World Health Organization International Clinical Trials Registry Platform. These sources were searched from January 1980 to April 2024, given that the concept of FCR was initially raised in the early 1980s [1]. The detailed search strategy and results are provided in Table S1 in Multimedia Appendix 1. We imported the search results from all databases into Covidence [27] and removed duplicates.

Study Selection

Two authors (NL and TOL) screened titles and abstracts independently for assessment against the inclusion criteria. A third author (CW) resolved any disagreements. A flowchart was created to standardize article selection decisions between screeners (Figure S1 in Multimedia Appendix 1). Two independent reviewers (NL and CW) assessed the full texts in detail against the inclusion criteria. Any disagreements between the reviewers were resolved through discussion with a third author (LR). Reasons for exclusion at the full-text stage were recorded and are presented in the PRISMA-ScR flow diagram (see the Results section).

Data Extraction

We piloted the data extraction form on the first 5 eligible studies and modified it accordingly. A data dictionary was developed with collective input from the author team to specify all elements of the data extraction form (Table S2 in Multimedia Appendix 1). One reviewer (NL) then extracted the data in Microsoft Excel under the supervision of the research team to ensure adherence to the predefined extraction protocol. The extracted data included:

- General information (author, year of publication, title, journal, country, and study aim);
- Participant characteristics (sample size, age, gender, diagnosis, inclusion criteria, and cancer stage);
- Concept-related information (definition of FCR/progression, study design, intervention timepoint, intervention components, intervention provider, provider qualification, recruitment method, intervention frequency/intervals/duration, definition of control group (if applicable), follow-up timepoints, theoretical framework, outcomes, measurement tool, and result);
- Context (setting and format of delivery);
- Study methodology, key findings, and research gaps.

Data Analysis and Presentation

Given the heterogeneity across included studies, we conducted a descriptive analysis approach to synthesize the findings. Descriptive numerical summaries were presented through tables and figures to map the characteristics of the included studies, and the findings were synthesized narratively [28]. The synthesis was structured around the review questions.

We summarized the characteristics of the identified studies by reporting frequencies and percentages. We also extracted how FCR was defined and which measurement tools were used in each study and grouped these definitions and measures into tables separately. Key intervention content (ie, components, sample sizes, measurement tools, study design, and reported statistical significance) was summarized. To clearly present the diverse intervention approaches for reducing FCR, we classified the interventions according to their theoretical basis (including blended interventions, cognitive behavioral therapy [CBT], mindfulness-based intervention [MBI], acceptance and commitment therapy [ACT], tailored psychoeducational interventions, and others) with categories developed inductively from the included studies. We created a harvest plot to visually present the distribution of interventions across cancer types and study designs.

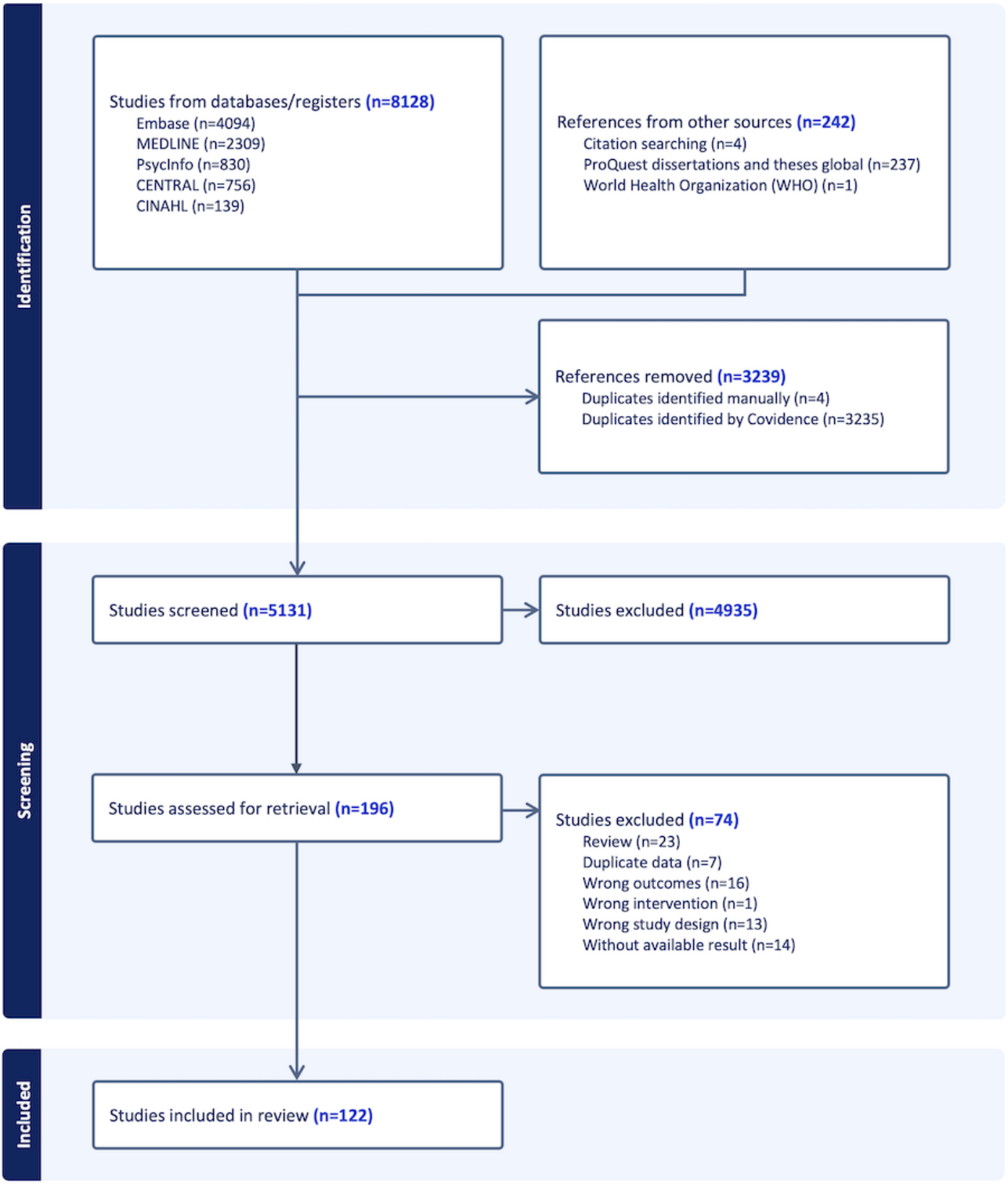
Results

Overview

We identified 8128 citations through database searches and 242 additional records from gray literature and reference tracking. After removing duplicates, 5131 titles and abstracts were screened, with 4935 excluded. We assessed the full texts of the remaining 196 articles and excluded 74 of them. A total of 122 studies were included in the review (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flowchart.



Sixty-five (53.3%) studies were published within the last 5 years (Table 1). The United States accounted for the highest number of studies (n=35, 28.7%). Forty-eight (39.3%) studies focus on breast cancer, followed by 47 (38.5%) studies that included participants with more than one cancer type. The majority of studies (n=64, 52.5%) were RCTs.

XSL•FO
RenderX

Table . Characteristics of the included studies (n=122).

Category	Values, n (%)
Publish date	
2009 - 2014	14 (11.5)
2015 - 2019	43 (35.2)
2020 - 2024	65 (53.3)
Geographic region (n=126)[a]	
North America	42 (33.3)
Europe	40 (31.7)
Asia	22 (17.5)
Oceania	20 (15.9)
South America	2 (1.6)
Targeted population	
Breast cancer	48 (39.3)
Gynecologic cancer	8 (6.6)
Melanoma	4 (3.2)
Colorectal cancer	3 (2.5)
Glioma	3 (2.5)
Prostate cancer	2 (1.7)
Renal	2 (1.7)
Other types of cancer	5 (4.0)
Stomach	1 (0.8)
Hypopharyngeal	1 (0.8)
Testicular	1 (0.8)
Oral	1 (0.8)
Lung	1 (0.8)
Any type (more than one type of cancer)	47 (38.5)
Most breast cancer	29 (23.8)
Most hematological cancers	3 (2.5)
Most head and neck cancers	1 (0.8)
Most lung cancer	1 (0.8)
Most lymphoma cancer	1 (0.8)
Mixed	12 (9.8)
Study design	
Randomized controlled trial	64 (52.5)
Pre-post study	44 (36.1)
Mixed methods research	11 (9.0)
Nonrandomized trial study	2 (1.6)
Observational design	1 (0.8)
Intervention (by theoretical base)	
Blended interventions	38 (31.1)
Cognitive behavioral therapy	26 (21.3)
Mindfulness-based interventions	24 (19.7)
Acceptance and commitment therapy	15 (12.3)

XSL•FO
RenderX

Category	Values, n (%)
Tailored psychoeducational intervention	13 (10.6)
Others	6 (5.0)
Mode of delivery	
In-person	75 (61.5)
Remote (app/web/videoconference/phone call)	34 (27.9)
Mixed (in-person+ remote)	13 (10.6)
Disciplinary specialist (ie, person delivering the intervention)	
Psychologists	68 (55.8)
Other researchers (eg, social workers/graduate students)	20 (16.4)
Any occupation	15 (12.3)
Nurses	7 (5.7)
Doctors	7 (5.7)
Not available	5 (4.1)

[a]Although 122 studies were included, some were conducted in more than one country; therefore, the total geographic region count is 126.

The interventions were categorized according to their theoretical basis (Table 1). Blended interventions (different combinations of CBT, mindfulness, ACT, and other strategies) formed the largest category (n=38, 31.1%), followed by CBT (n=26, 21.3%) and MBIs (n=24, 19.7%). All studies implemented interventions post diagnosis (Table S3 in Multimedia Appendix 1). In 38 (31.1%) studies, researchers did not report the exact clinical timepoint of intervention delivery. Thirty-six (29.5%) trials included "people who had completed or were still undergoing cancer treatment." The remaining 48 (39.3%) targeted "people who were disease-free," with the timing of intervention delivery ranging from 3 months to 15 years following primary cancer treatment. Most interventions were delivered by disciplinary specialists (n=117, 95.9%), with psychologists being the most common providers (n=68, 55.8%).

The Definition of FCR

Only 28 (23%) studies identified which FCR definition they used. Among them, 9 (32.1%) studies used Vickberg's definition [28-36], 18 (64.3%) studies used Lebel et al definition [37-54], and 1 (3.6%) study [55] cited the FCR definition by Herschbach and Dinkel [56]. These three definitions are summarized in Table 2. Vickberg's [57] definition, developed for breast cancer patients, describes FCR as the fear that cancer may return either in the same location or in another part of the body. Lebel et al [2] proposed a definition based on a Delphi consensus process, defining FCR as "fear, worry, or concern relating to the possibility that cancer will come back or progress." Herschbach and Dinkel [56] introduced the concept of fear of progression (FoP), which specifically highlights concerns about disease progression and its associated psychosocial consequences.

Table . Definition of FCR[a].

Author	Definition
Vickberg (2003) [57]	"Fear or concern about the return or progression of the disease in the same organ or other parts of the body."
Lebel et al (2016) [2]	"Fear, worry, or concern relating to the possibility that cancer will come back or progress."
Herschbach and Dinkel (2014) [56]	"Fear of progression (FoP) is defined as the fear and anxiety of patients whose disease will progress or recur with all its social and biological consequences."

[a]FCR: fear of cancer recurrence.

FCR Measurement Tools

Eighteen validated measurement tools were identified (Table S4 in Multimedia Appendix 1). Thirteen (72.2%) were specifically designed to assess FCR/FoP, while the remaining 5 (27.8%) included FCR-related items as part of broader assessments. The Fear of Cancer Recurrence Inventory [58] was the most frequently used tool, appearing in 40 (31.7%) studies. Other commonly used tools included the Concerns About Recurrence Scale [57], used in 20 (15.9%) studies, and the Fear of Cancer Recurrence Inventory-Short Form [59], used in 18 (14.3%) studies.

Intervention Characteristics and Effects

Summary

We summarized author, year, intervention, sample size, measurement tools, statistical significance, and study design (Table 3). The distribution and reported effects of interventions across cancer types and study designs are summarized visually

XSL•FO
RenderX

(Figure 2). Detailed characteristics of statistically significant interventions are outlined (Table S5 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . The intervention and effect.

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
AhmadiQaragezlou et al 2020 [29]	38	Mindfulness-based stress reduction (MBSR) program	FCRI	Yes	RCT[c]
Akkol-Solakoglu and Hevey 2023 [60]	72	Internet-delivered cognitive behavioral therapy (CBT)	CWS	No	RCT
Akechi et al 2021 [61]	59	Collaborative care intervention—problem-solving treatment (PST) and behavioral activation program (BA) (modified version)	CARS	No	RCT
Akechi et al 2023 [62]	444	Smartphone-based PST (Kaiketsu-App) and behavioral activation program (Genki-App)—problem-solving treatment (PST) and behavioral activation program (BA) (modified version)	CARS and FCRI-SF	Yes	RCT
Almeida et al 2022 [37]	17	Emotion-focused therapy (EFT)—humanistic-experiential psychotherapy stemming from person-centered, gestalt, and existential traditions	PQ-FCR	Yes	Pre-post study
Ananeva 2020 [63]	N/A[d]	ConquerFear (CF)—Common-Sense Model (CSM) of illness, the Self-Regulatory Executive Function model (S-REF; targets metacognitions), and Relational Frame Theory (RFT; theoretical basis for acceptance and commitment therapy)	FCRI	Yes	Pre-post study
Arch and Mitchell 2016 [64]	42	Valued Living (VL) intervention	CARS	Yes	Pre-post study
Arch et al 2021 [65]	134	Valued Living (VL) intervention	CARS	Yes	RCT
Arch et al 2024 [66]	29	Written exposure-based coping intervention (EASE)	CARS and FOP-Q-SF	Yes	Pre-post study

XSL•FO
RenderX

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
Beith et al 2017 [67]	222	ConquerFear (CF)—Common-Sense Model (CSM) of illness, the Self-Regulatory Executive Function model (S-REF; targets metacognitions), and Relational Frame Theory (RFT; theoretical basis for acceptance and commitment therapy)	FCRI	Yes	RCT
Bergerot et al 2022 [68]	23	Smartphone app–based mindfulness program	FCR-7	Yes	Pre-post study
Bergerot et al 2023 [69]	41	Mindfulness-Based Cancer Survivorship Journey	FCR-7	Yes	Pre-post study
Bin et al 2023 [70]	98	Observation group based on the Rosenthal effect	FCRI	No	RCT
Brooker et al 2020 [71]	173	Mindful Self-Compassion (MSC) program	FCRI-SF	Yes	Pre-post study
Burm et al 2019 [38]	88	Blended cognitive behavioral therapy (CBT)	CWS	Yes	RCT
Butow et al 2017 [72]	222	ConquerFear (CF)—Common-Sense Model (CSM) of illness, the Self-Regulatory Executive Function model (S-REF; targets metacognitions), and Relational Frame Theory (RFT; theoretical basis for acceptance and commitment therapy)	FCRI	Yes	RCT
Chambers et al 2012 [73]	19	Mindfulness-based cognitive therapy group intervention (MBCT)	MAX-PC	Yes	Mixed methods research
Chang 2025 [74]	50	Integrated Mindfulness-Based Fitness Training (MBFT) program	ASC	Yes	RCT
Cheng et al 2021 [75]	172	Mindfulness-Based Fitness Training (MBFT) program	FCRS	Yes	RCT
Cillessen et al 2018 [76]	245	Group face-to-face and individual internet-based mindfulness-based cognitive therapy (MBCT and eMBCT)	FCRI-SF	No	RCT
Cohen et al 2022 [31]	19	Mindfulness-based cognitive therapy (MBCT)	FCRI	Yes	Pre-post study
Compen et al 2018 [77]	245	Face-to-face and internet-based mindfulness-based cognitive therapy (MBCT and eMBCT)	FCRI	Yes	RCT

XSL•FO
RenderX

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
Crane-Okada et al 2012 [78]	49	Mindful Movement Program (MMP)	FCRS	Yes	RCT
Davidson et al 2018 [30]	16	Mini - Adjustment to Fear, Threat, and Expectation of Recurrence (Mini - AFTERc) intervention	FCRI-SF	Yes	Pre-post study
Deuning-Smit et al 2024 [39]	81	Cognitive behavioral therapy (CBT)	CWS-6 and FCRI-SF	Yes	Mixed methods research
Dieng et al 2016 [79]	164	Melanoma care program	FCRI	Yes	RCT
Dieng et al 2019 [80]	164	Melanoma care program	FCRI	Yes	RCT
Dieng et al 2020 [81]	164	Melanoma care program	FCRI	Yes	RCT
Dirkse et al 2020 [82]	86	Internet-delivered cognitive behavioral therapy	FCRI-SF	Yes	RCT
Dodds et al 2015 [83]	33	Cognitively Based Compassion Training (CBCT) intervention	FCRI	Yes	RCT
Döking et al 2021 [84]	1	Combined face-to-face and online cognitive behavioral therapy	CWS	Yes	Mixed methods research
Eckert et al 2020 [85]	15	Mindfulness-enhanced cognitive stress management (MECSM) group intervention	CWS	No	Pre-post study
Fishbein et al 2023 [86]	113	Valued Living (VL) intervention	CARS	No	RCT
Fishbein et al 2022 [87]	134	Valued Living (VL) intervention	CARS	Yes	RCT
Fishbein and Arch 2022 [88]	73	Valued Living (VL) intervention	CARS	No	RCT
Fisher et al 2019 [89]	27	Metacognitive therapy (MCT)—trans-diagnostic theory of psychopathology, the Self-Regulatory Executive Function (S-REF) mode	FCRI	No	Pre-post study
Fisher et al 2017 [90]	4	Metacognitive therapy (MCT)—trans-diagnostic theory of psychopathology, the Self-Regulatory Executive Function (S-REF) mode	FCRI	Yes	Nonconcurrent multiple baseline design
Frangou et al 2021 [91]	182	Psychological support intervention primarily incorporated cognitive behavioral therapy (CBT), Mindfulness and acceptance and commitment therapy (ACT) elements	FOP-Q-SF	Yes	RCT

XSL•FO
RenderX

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
Gonzalez-Hernandez et al 2018 [92]	56	Cognitively-Based Compassion Training (CBCT)	FCRI	Yes	RCT
Hall et al 2022 [93]	23	A mind-body resiliency intervention—relaxation response (RR), cognitive behavioral based	FCRI-SF	Yes	Mixed methods research
Hall et al 2020 [94]	23	A mind-body resiliency intervention—relaxation response (RR), cognitive behavioral based	FCRI	Yes	Pre-post study
Hasannezhad Reskati et al 2020 [55]	30	Acceptance and commitment therapy (ACT)	FOP-Q	Yes	RCT
Herschbach 2012 [95]	174	Cognitive behavioral group therapy (CBT)	FOP-Q	Yes	RCT
Herschbach et al 2010 [95]	265	Group Psychotherapy CBT intervention	FOP-Q	Yes	RCT
Howells et al 2019 [91]	107	Cognitive behavioral therapy (CBT)	FOP-Q-SF	Yes	RCT
Humphris and Rogers 2012 [96]	90	AFTER (Adjustment to the Fear, Expectation or Threat of Recurrence) intervention	WOC	Yes	RCT
Imai et al 2019 [97]	38	Problem-solving therapy (PST)	CARS	Yes	Pre-post study
Johns et al 2020 [98]	91	Acceptance and commitment therapy (ACT)	FCRI-SF	Yes	RCT
Kacel 2019 [99]	32	Mindfulness yoga program	FCRI	No	Pre-post study
Lebel et al 2014 [32]	56	Cognitive-existential (CE) group intervention-Leventhal's common sense model, Mishel's uncertainty in illness theory and cognitive models of worry	FCRS	Yes	Pre-post study
Lee et al 2022 [100]	54	Mindfulness stress management (MSM) intervention	FCRI	Yes	RCT
Lengacher et al 2016 [101]	322	Mindful Movement Program (MMP)	CARS	Yes	RCT
Lengacher et al 2021 [102]	322	Mindfulness-based stress reduction (MBSR) intervention	CARS	Yes	RCT
Lengacher et al 2009 [103]	84	Mindfulness-based stress reduction (MBSR) intervention	CARS	Yes	RCT
Lengacher et al 2018 [104]	15	Mobile mindfulness-based stress reduction (mMBSR) intervention	CARS	Yes	Pre-post study

XSL•FO
RenderX

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
Lichtenthal et al 2017 [105]	110	Attention and Interpretation Modification for Fear of Breast Cancer Recurrence (AIM-FBCR)	CARS	Yes	RCT
Li et al 2021 [106]	140	Psychological nursing intervention	FoP-Q-SF	Yes	RCT
Liu et al 2021 [107]	61	Clinician Intervention to address Fear of Cancer Recurrence (CIFeR)—FCR theoretical models (self-regulation model of illness, self-regulatory executive functioning model, family-based model, uncertainty in illness theory, social-cognitive processing theory, and terror management theory)	FCRI	Yes	Pre-post study
Loughan et al 2022 [108]	12	Managing Cancer and Living Meaningfully (CALM) intervention	FCR-7	No	Pre-post study
Loughan et al 2021 [109]	10	Managing Cancer and Living Meaningfully (CALM) intervention	FCR-7	No	Mixed methods research
Luberto et al 2019 [110]	1	Mindfulness-based cognitive therapy (MBCT)	CWS	Yes	Mixed methods research
Luigjes‐Huizer et al 2023 [40]	167	Online primary care intervention—cognitive behavioral therapy and the model on FCR by Lee‐Jones	FCRI-SF	Yes	RCT
Maheu et al 2023 [111]	164	Fear of cancer recurrence therapy (FORT) based on cognitive existential (CE) approach, which combines elements of modern existential therapies and cognitive behavioral therapy (CBT)	FCRI	Yes	RCT
Manne et al 2017 [112]	352	Communication enhancing intervention (CCI)/supportive counseling (SC)—cognitive-affective social processing theory/encouraging emotional expression, supporting existing coping behaviors, and enhancing self-esteem and autonomy	CARS	Yes	RCT
Martin et al 2020 [113]	114	iHOPE (Help to Overcome Problems Effectively) intervention	QLACS	Yes	Pre-post study

XSL•FO
RenderX

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
McHale et al 2024 [41]	92	Mini - Adjustment to Fear, Threat, and Expectation of Recurrence (Mini - AFTERc) intervention	FCR-4	Yes	Nonrandomized trial study
Merckaert et al 2017 [114]	170	Multiple - component structured manualized group intervention (MGI)—cognitive behavioral and hypnosis components	FCRI	Yes	RCT
Momino et al 2017 [115]	40	Collaborative care intervention—problem-solving therapy (PST) and behavioral activation therapy (BAT)	CARS	No	Pre-post study
Montesinos and Francisco 2016 [116]	12	Abridged version of acceptance and commitment therapy (ACT)	Mini-MAC	Yes	Pre-post study
Murphy et al 2020 [117]	114	Internet - delivered cognitive behavioral therapy (iCBT)	FCRI	Yes	RCT
Neubert et al 2023 [118]	157	Video Sequence-Based Intervention	FoP-Q-SF	No	RCT
Nguyen et al 2022 [119]	57	Virtual telephone coaching program	FCRI	Yes	Pre-post study
Otto 2015 [120]	67	Gratitude intervention	CARS	Yes	Pre-post study
Park et al 2020 [121]	74	Mindfulness-based cognitive therapy (MBCT) program	CARS	Yes	RCT
Peng et al 2022 [33]	65	Mindfulness-based stress reduction (MBSR) intervention	FCRI-SF	Yes	RCT
Pradhan et al 2021 [122]	62	Online Fear of Cancer Recurrence Booklet—adapted from the ConquerFear (Common Sense Model of Illness, Self-Regulatory Executive Function Model (S-REF), and Relational Frame Theory)	FOP-Q-SF	No	Pre-post study
Reb et al 2020 [43]	31	ConquerFear (CF)—contemporary cognitive-processing model from Metacognitive Therapy (MCT) and acceptance and commitment therapy (ACT) principles	FOP-Q-SF	Yes	Pre-post study
Reb et al 2020 [44]	2	Day-by-day (DBD) intervention—ConquerFear based (Common Sense Model of Illness, Self-Regulatory Executive Function Model (S-REF), and Relational Frame Theory)	FOP-Q-SF	Yes	Pre-post study

XSL•FO
RenderX

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
Smith et al 2020 [123]	44	Online psychological self-management intervention—cognitive behavioral, acceptance and commitment, metacognitive, and relaxation therapy	FCRI-SF	Yes	Mixed methods research
Rudolph et al 2018 [124]	10	Acceptance and commitment therapy (ACT)	FoP-Q	Yes	Pre-post study
Russell et al 2019 [125]	69	Online mindfulness-based program (MBIs)	FCRI	Yes	RCT
Sajadian et al 2021[126]	30	Psycho-spiritual therapy	FCRI	Yes	RCT
Saltbæk et al 2024 [127]	503	Nurse-led individualized follow-up	CARQ	Yes	RCT
Salazar-Alejo et al 2023 [128]	97	Online mindfulness-based stress reduction (MBSR) intervention	CWS	Yes	RCT
Sarizadeh et al 2018 [129]	20	Acceptance and commitment therapy (ACT)	FCRI	Yes	Pre-post study
Sauer and Maatouk 2021 [130]	30	Acceptance and commitment therapy (ACT)	N/A	Yes	Pre-post study
Savard et al 2018 [45]	38	Group cognitive behavioral therapy (CBT)	FCRI	Yes	Mixed methods research
Sakai et al 2024 [131]	1	Acceptance and commitment therapy (ACT)	CARS	Yes	Pre-post study
Schlecht et al 2023 [132]	155	Video-based intervention—ACT, mindfulness-based	FOP-Q-SF	Yes	Observational design
Sharpe et al 2019 [46]	152	ConquerFear (CF)—a combination of metacognitive therapy, the self-regulation theory, and acceptance and commitment therapy	FCRI	Yes	Pre-post study
Shih et al 2014 [133]	96	ConquerFear (CF)—Self-Regulatory Executive Function Model and Relational Frame Theory	FCRI	Yes	RCT
Shumay et al 2013 [134]	28	Acceptance and commitment therapy (ACT)	FCRI	No	RCT
Sinclair et al 2023 [135]	97	Acceptance and commitment therapy (ACT)	FCRI-SF	Yes	Pre-post study
Sinclair et al 2020 [136]	79	Acceptance and commitment therapy (ACT) and patient education intervention (adapted from ConquerFear intervention)—Common Sense Model of Illness, Self-Regulatory Executive Function Model (S-REF), and Relational Frame Theory	FACT	Yes	Mixed methods research

XSL•FO
RenderX

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
Smith et al 2024 [47]	54	ConquerFear (CF)—Common Sense Model of Illness, Self-Regulatory Executive Function Model (S-REF), and Relational Frame Theory	FCRI-SF	Yes	Mixed methods research
Smith et al 2015 [34]	8	ConquerFear (CF)—Common Sense Model of Illness, Self-Regulatory Executive Function Model (S-REF), and Relational Frame Theory	FCRI	Yes	Pre-post study
Steinecke et al 2022 [137]	33	Psychobiological group therapy program—behavioral strategies as well as mindfulness	N/A	Yes	Pre-post study
Subnis 2014 [138]	40	Online expressive writing intervention (EW)	FCRI-SF	Yes	RCT
Sun et al 2023 [139]	40	Traditional Chinese Medicine (TCM) interventions combined group psychotherapy, mindfulness, and cognitive behavioral therapy-based	FCRI	Yes	Pre-post study
Tauber et al 2023 [48]	85	ConquerFear (CF)—Common Sense Model of Illness, Self-Regulatory Executive Function Model (S-REF), and Relational Frame Theory	FCRI	Yes	RCT
Tauber et al 2022 [49]	27	ConquerFear (CF)—Common Sense Model of Illness, Self-Regulatory Executive Function Model (S-REF), and Relational Frame Theory	FCRI-SF	Yes	RCT
Thewes et al 2012 [140]	8	ConquerFear (CF)—Common Sense Model of Illness, Self-Regulatory Executive Function Model (S-REF), and Relational Frame Theory	FCRI-SF	Yes	Pre-post study
Tomei et al 2018 [50]	25	Cognitive-existential (CE) psychotherapy intervention—Leventhal's common sense model, Mishel's uncertainty in illness, and cognitive models of worry	FCRI	Yes	RCT

XSL•FO
RenderX

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
Tomei et al 2016 [35]	3	Cognitive-existential (CE) psychotherapy intervention—Leventhal's common sense model, Mishel's uncertainty in illness, and cognitive models of worry	FCRI	Yes	Pre-post study
Tomei et al 2014 [141]	1	Cognitive-existential (CE) psychotherapy intervention—Leventhal's common sense model, Mishel's uncertainty in illness, and cognitive models of worry	FCRI	No	Pre-post study
Tran 2025 [142]	53	Fear-Less self-management intervention (stepped-care framework)	FCRI-SF	Yes	Pre-post study
Ulfig 2020 [36]	35	Cognitive behavioral therapy (CBT)	FACT	Yes	RCT
Ulfig 2020 [36]	115	Cognitive behavioral therapy (CBT)	FACT	Yes	Pre-post study
Valcu et al 2023 [143]	9	Cognitive behavioral therapy (CBT)	FoP-Q-SF	Yes	Pre-post study
van de Wal et al 2018 [51]	1	Cognitive behavioral therapy (CBT)	CWS and FCRI	Yes	Mixed methods research
van de Wal et al 2015 [144]	1	Blended cognitive behavioral therapy (CBT)	CWS	Yes	Pre-post study
van de Wal et al 2017 [145]	88	Blended cognitive behavioral therapy (CBT)	CWS	Yes	RCT
van Helmondt et al 2023 [52]	262	Cognitive behavioral therapy (CBT)	FCRI	No	RCT
Victorson et al 2012 [146]	115	Mindfulness-based stress reduction intervention (MBCT)	MAX-PC	Yes	RCT
Wagner et al 2021 [147]	196	Cognitive behavioral therapy (CBT)	FCRI	Yes	RCT
Wang et al 2023 [53]	103	Managing cancer and living meaningfully (CALM)	CWS	Yes	RCT
Wang et al 2023 [148]	98	Meaning-centered group psychotherapy (MCGP)—Viktor Frankl's theory and principles, and contains some existential elements	FOP-Q-SF	Yes	RCT
Weis et al 2020 [149]	50	Psychoeducational (PE) group intervention	FoP-Q-SF	Yes	Nonrandomized trial study
Yoon et al 2023 [150]	41	Interactive coaching intervention	FCRI-SF	No	Pre-post study
Zhang et al 2022 [151]	98	Managing cancer and living meaningfully (CALM)	CARS	No	RCT

XSL•FO
RenderX

Authors (year)	Sample, n	Intervention	FCR tools[a]	Statistically significant[b]	Research design
Zhao et al 2023 [54]	80	Managing cancer and living meaningfully (CALM) intervention	FCRI	Yes	RCT
Zhao and Xu 2021 [152]	258	Comprehensive nursing based on cognitive behavioral intervention	FoP-Q-SF	Yes	RCT

[a]Fear of cancer recurrence (FCR) tools used are as follows—ASC: Assessment of Survivor Concerns; CARQ: Concerns About Recurrence Questionnaire; CARS: Concerns About Recurrence Scale; CWS: Cancer Worry Scale; CWS-6: 6-Item Cancer Worry Scale; FACT: Functional Assessment of Cancer Therapy; FCR-4: Fears of Cancer Recurrence, 4-Item Version; FCR-7: Fears of Cancer Recurrence, 7-Item Version; FCRI: Fear of Cancer Recurrence Inventory; FCRI-SF: Fear of Cancer Recurrence Inventory—Short Form; FCRS: Fear of Cancer Recurrence Scale; FoP-Q: Fear of Progression Questionnaire; FOP-Q-SF: Short-Form Fear of Progression Scale; MAX-PC: Memorial Anxiety Scale for Prostate Cancer; Mini-MAC: Mini-Mental Adjustment to Cancer Scale; PQ-FCR: Personal Questionnaire—Fear of Recurrence; QLACS: Quality of Life in Adult Cancer Survivors Scale; WOC: Worry of Cancer Scale.

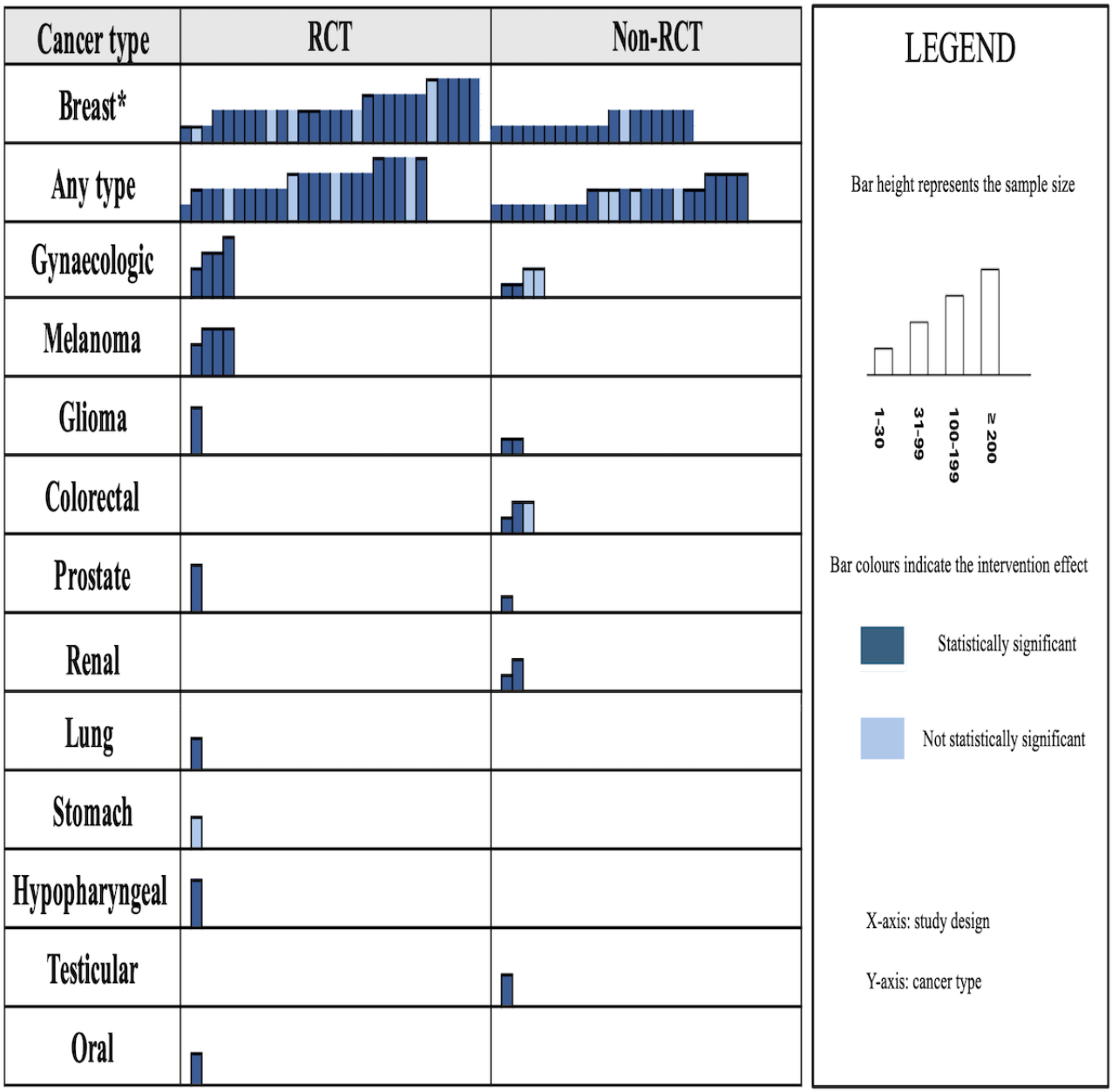
[b]"Yes" means that the intervention had a statistically significant mitigating effect on fear of cancer recurrence (FCR); "No" means that the intervention did not have a statistically significant mitigating effect on FCR.

[c]RCT: randomized controlled trial.

[d]N/A: not available.

XSL•FO
RenderX

Figure 2. Harvest plot showing the distribution of interventions by study design and cancer type. Each bar represents one study. Bar colors indicate the intervention effect, and bar height reflects the sample size. “Any type” refers to studies that include more than one type of cancer. Blue bars indicate a statistically significant effect (P<.05) in reducing fear of cancer recurrence; light blue bars indicate no statistically significant effect. Studies are grouped by cancer type (y-axis) and study design (randomized controlled trial [RCT] vs non-RCT; x-axis). * represents one study with missing sample size data. Within each group, studies are ordered from left to right by sample size, starting with the smallest sample size on the left.



Blended Interventions

Thirty-eight (31.1%) studies evaluated the effect of blended interventions [31,32,34,35,37,40,42-44,46,50,53,61-63,67,72,89,91,93,94,107,111,112,114,115,123,132,133,136,137,139-141]. Most of them (n=23, 60.5%) combined CBT with other interventions (eg, mindfulness, ACT, and other strategies). Eighteen (47.4%) studies used a pre-post study design [31,32,34,35,37,42-44,46,63,89,94,107,115,137,139-141]. Fourteen (36.8%) studies were RCTs [40,48-50,53,61,62,67,72,91,111,112,133,153].

The majority of studies (n=33, 86.8%) reported statistically significant effects. Among them, 22 (66.7%) were conducted in person by disciplinary specialists [31,32,34,37,46,48,49,63,67,72,90,91,93,94,107,111,114,133,136,137,140,148], 8 (24.2%) were delivered remotely [40,43,44,47,62,123,132,139], and 3 (9.1%) used a hybrid format, combining both in-person and remote delivery [35,50,112]. These interventions typically followed a weekly format, lasting from 5 to 12 weeks. Individual sessions ranged from 10 minutes to 2 hours. Moreover, the length of the follow-up period varied, with assessments conducted from immediately after the intervention up to 2 years.

CBT

Twenty-six (21.3%) studies evaluated the effect of CBT [30,36,38,39,41,45,51,52,60,82-84,91,92,95,96,105,119,143-145,147,152,154-156].

XSL•FO
RenderX

The majority of the studies (n=16, 61.5%) were RCTs [36,38,52,60,82,83,91,92,95,96,105,117,145,147,152,155], followed by 5 (19.2%) pre-post studies [30,119,143,144,154], 4 (15.4%) mixed methods studies [39,45,51,84], and 1 (3.9%) non-RCT [41].

Twenty-three (88.5%) studies reported statistically significant effects. Among them, 11 (47.8%) were delivered face to face by disciplinary specialists [36,39,45,83,95,96,105,152,154,155,157], while 7 (30.4%) studies used remote delivery methods [30,41,82,117,119,143,147], followed by 5 (21.8%) delivered both in-person and online [38,51,84,144,145]. These interventions typically followed a weekly format, lasting from 1 to 12 weeks. Individual sessions ranged from 10 minutes to 2.5 hours, with follow-up periods spanning from immediately after the intervention to 15 months.

Mindfulness-Based Interventions

Twenty-four (19.7%) studies evaluated the effect of MBIs [29,33,68,69,71,73-78,85,99-104,110,121,125,126,128,146]. Seven (29.2%) studies [29,33,100,102-104,128] used mindfulness-based stress reduction interventions. Six (25%) [73,76,77,110,121,146] applied mindfulness-based cognitive therapy. Four (16.7%) studies [74,75,78,101] evaluated mindfulness-based fitness or movement programs, and 1 (4.2%) study [71] conducted a mindful self-compassion program. Additionally, 6 (25%) studies [68,69,85,99,125,126] incorporated a mix of mindfulness components. Most studies (n=16, 66.7%) were RCTs [29,33,74-78,100-103,121,125,126,128,146], followed by 6 (25%) pre-post studies [68,69,71,85,99,104] and 2 (8.3%) mixed methods studies [73,110].

The majority (n=21, 87.5%) reported statistically significant effects. Disciplinary specialists delivered most effective interventions (n=14, 66.7%) face-to-face, while 7 (33.3%) studies delivered interventions remotely via websites [74,104,125,128], and apps [62,68,69]. The majority of interventions (n=19, 90.5%) involved weekly sessions lasting between 5 and 12 weeks, with each session lasting between 90 minutes and 2.5 hours. All studies measured outcomes immediately post intervention, and the longest follow-up lasted 1 year.

Acceptance and Commitment Therapy

Fifteen (12.3%) studies evaluated the effect of ACT [55,64,65,86-88,98,116,118,124,129-131,134,135]. Among them, 8 (53.3%) studies were RCTs [55,65,86-88,98,118,134], while 7 (46.7%) followed a pre-post design [64,116,124,129-131,135].

Most studies (n=11, 73.3%) reported statistically significant effects. All effective interventions were delivered in person by disciplinary specialists, following a weekly format that ranged from a single session to 8 weeks. Each session lasted between 50 and 120 minutes. Follow-up periods varied widely, from immediately post intervention to up to 8 months.

Tailored Psychoeducational Interventions

Thirteen (10.6%) trials evaluated the effect of tailored psychoeducational interventions. Among them, 5 (38.5%) studies [53,54,108,109,151] evaluated the effect of an intervention named the Managing Cancer and Living Meaningfully intervention. Three (23.1%) studies evaluated the effect of a melanoma care program [79-81]. The majority of studies were RCTs (n=9, 69.2%) [53,54,70,79-81,106,127,151].

Most studies (n=9, 69.2%) reported statistically significant effects. All interventions were delivered by disciplinary specialists, with 8 (88.9%) delivered face-to-face. The duration of effective interventions ranged from 4 to 24 weeks, with individual sessions lasting between 30 minutes and 4 hours. Follow-up periods ranged from immediately post intervention to up to 3 years.

Others

In addition to the widely used interventions described above, 6 (4.9%) studies explored alternative approaches, including writing interventions [66,138], a holistic biopsychosocial intervention [150], a problem-solving intervention [97], an intervention grounded in hope theories [113], and an intervention inspired by broadening and building theories [158]. Among them, 5 (83.3%) studies used a pre-post design [66,97,113,120,150], and 1 used an RCT design [138]. Five (83.3%) reported statistically significant effects and were delivered online by disciplinary specialists. Follow-up periods ranged from immediately post intervention to 4.5 months.

Discussion

Principal Findings

This scoping review is the first to map all psychological intervention strategies for alleviating FCR across cancer types and to identify key gaps in existing evidence. From our review, only about one-fifth of the studies reported the definition of FCR used. Eighteen validated FCR measurement tools were identified. Blended interventions, CBT, and MBIs were the most common approaches, typically delivered face-to-face by disciplinary specialists. Breast cancer populations dominated the evidence base, while other cancer types received limited attention. Most interventions reported statistically significant effects in reducing FCR, although the wide variation in study designs, intervention components, outcome measures, and follow-up durations makes meaningful interpretation difficult.

A clear definition of FCR is critically important, as it underpins the identification of its clinical manifestations, supports accurate diagnosis in clinical settings, and informs the design of targeted interventions that address the specific fears and cognitive patterns associated with cancer recurrence [21,32]. Much of the existing literature treats the definitions of FCR and FoP interchangeably [159,160]. However, Coutts-Bain et al [161] conducted a factor analysis and found that, although FCR and FoP are related, they are structurally distinct psychological constructs. It challenges the previous assumption that the two are the same construct and emphasizes the importance of clearly distinguishing between FCR and FoP in clinical practice and research. Conceptually, FCR is more relevant to cancer survivors

XSL•FO
RenderX

who have completed curative treatment and are in remission, encompassing concerns about the cancer returning after a period of being disease-free. In contrast, individuals with metastatic cancer are less likely to fear recurrence and more likely to experience fear of progression, worrying about the cancer worsening or spreading further due to the ongoing nature of their illness [162]. Consequently, clearly distinguishing between FCR and FoP in clinical practice and research not only facilitates the development of more effective, targeted interventions but also enhances the accuracy of research findings and enables the provision of more tailored psychological support to cancer survivors.

Blended interventions were the most commonly investigated for reducing FCR among cancer survivors, with numerous studies reporting positive outcomes. This finding aligns with another review of 13 studies [163], which also identified that interventions combining CBT and MBI strategies were effective in reducing FCR. In recent years, researchers have proposed that traditional CBT may have limitations in managing FCR, particularly among survivors living with ongoing uncertainty [21,164]. An overemphasis on challenging "irrational" thoughts may leave some people feeling misunderstood or emotionally invalidated, especially when their fears are grounded in real risk rather than cognitive distortion [165]. However, evidence supporting these concerns remains limited. Some scholars regard ACT as a contextual extension of CBT, incorporating core mindfulness-based practices to enhance individuals' awareness and acceptance of their emotions [166]. Similarly, mindfulness techniques are increasingly integrated into traditional CBT. A prime example of this integration is mindfulness-based cognitive therapy, which combines the cognitive restructuring techniques of CBT with mindfulness practices to help individuals break the cycle of negative thinking and reduce the FCR [121].

The diversity in FCR interventions reflects the lack of a universally agreed theoretical foundation and the variety of disciplinary approaches to understanding cancer-related psychological experiences, which creates challenges for their application in clinical practice. Despite growing research on FCR interventions, cancer survivors still rarely receive support for managing FCR outside the research environment [167]. It is therefore important to identify the core effective components across FCR interventions and to examine whether tailored approaches for specific cancer populations offer advantages over more general interventions [168]. Furthermore, for widespread adoption, considerations such as cost-effectiveness and resource requirements will be essential to ensure the feasibility of dissemination [169,170].

The majority of existing FCR interventions were resource-intensive, requiring in-person delivery over multiple sessions by trained specialists. These findings are also supported by a previous review by Pradhan et al [122]. As FCR is a persistent and commonly unmet concern across cancer survivor populations, there is a pressing need for scalable solutions [47,164]. In addition, a recently published consensus statement [165] indicates that approximately 80% of cancer survivors do not require high-intensity therapy, and that lower-intensity approaches, such as self-guided online programs, psychoeducational resources, or brief interventions delivered via SMS text messages or telephone, can reduce FCR [165]. In this context, mobile health interventions offer greater scalability and accessibility with reduced demands on health care resources, making them a promising alternative [171]; text messaging–based programs have shown positive outcomes in the management of mental health disorders [172] and in improving self-management of chronic conditions [173]. However, there remains a lack of mobile health programs in managing FCR.

In this review, the majority of studies (85%) reported statistically significant effects of interventions on FCR. This trend is consistent with evidence from previous systematic reviews and meta-analyses, which also demonstrated small-to-moderate reductions in FCR across specific psychological interventions [164,174]. However, this predominance of positive findings may be partially explained by publication bias, whereby studies with null or negative results are less likely to be published [175]. Researchers and journals often favor significant or "positive" outcomes, which can distort the evidence base and lead to an overestimation of intervention effectiveness [176]. Selective outcome reporting may further contribute to this bias, as authors might emphasize significant results while downplaying or omitting nonsignificant findings [177]. Small sample sizes in some studies may increase the risk of false positives and inflated effect sizes [178]. Collectively, these factors could mislead clinical decision-making and contribute to the adoption of interventions with uncertain or limited effects. In addition, while most interventions reduced FCR, these findings should be interpreted with caution due to the heterogeneity across the included studies. The interventions varied widely in terms of dose, delivery format, sample size, measurement time points, and study design, all of which pose significant challenges for synthesizing the findings. Such variability limits the comparability of studies and makes it difficult to have a specific conclusion regarding the overall effect of interventions for managing FCR. These challenges are inherent to scoping reviews, which aim to map diverse and broad evidence [28,179]. It is necessary to conduct a systematic review with high-quality RCTs to evaluate the effectiveness of different interventions in the future.

Although FCR was recognized as a clinical concern in the 1980s [1], our review found that interventions only began to appear from around 2009. FCR was underrecognized as a distinct psychological construct and was often assessed as part of broader measures of distress, anxiety, or quality of life previously [7]. The lack of validated tools specifically designed to assess FCR also limited early research and hindered the earlier evaluation of intervention outcomes [58]. Over the past two decades, the growing number of cancer survivors has shifted the focus of cancer care from survival alone to long-term psychological well-being, highlighting the need for FCR-specific interventions [2].

Limitations

First, this review used a vote-counting approach to summarize intervention effects because of the heterogeneity across studies. However, this means all studies were weighted equally, regardless of their quality, sample size, or effect size, which

XSL•FO
RenderX

limits our ability to compare effect estimates or assess the clinical meaningfulness of intervention outcomes. In addition, we did not conduct a quality appraisal of the included studies, consistent with the methodological nature of a scoping review. Future research would benefit from high-quality systematic reviews and meta-analyses that incorporate rigorous risk-of-bias assessments and enable quantitative comparisons of the effectiveness of different intervention modalities.

Second, this review was limited to English-language articles or English-language abstracts, and the exclusion of non-English publications may have reduced the completeness and representativeness of the evidence. This restriction may introduce language bias and underrepresent studies conducted in non-English-speaking regions. Nonetheless, our review included studies from 5 continents, indicating some geographic diversity despite the language limitation. Future reviews should consider removing language restrictions to ensure broader inclusion.

Third, this review relied solely on published information and did not seek additional details from study authors, which may have resulted in incomplete reporting of methodological or contextual elements, such as the specific definitions of FCR used. Future studies would benefit from more detailed reporting and from contacting authors when necessary to obtain missing information.

Conclusion and Implications

Our scoping review mapped the full range of psychological intervention strategies aimed at alleviating FCR across cancer types. We identified 122 studies encompassing diverse FCR interventions. FCR has been the focus of an increasing number of studies since 2009, with the majority being RCTs. Most interventions are delivered face-to-face and rely on trained specialists, limiting their scalability and accessibility. Most have had statistically significant results. However, the included studies demonstrated heterogeneity in terms of delivery, duration, and dose, requiring cautious interpretation of intervention effects. Future research should develop consistent guidelines to standardize the definition of FCR, the measurement tools used, and the timing of follow-up assessments, thereby improving comparability across studies and strengthening the evidence base. Long-term follow-up data are needed to evaluate the sustained effects. As most studies have focused on breast cancer survivors, further research is needed to provide evidence for other cancer populations.

Acknowledgments

We are grateful to Ms Rayna Dewar for assistance with the search strategy.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

This scoping review is based entirely on data extracted from publicly available published literature. No new or unpublished data were generated or analyzed in this study.

Authors' Contributions

Conceptualization: AJ, CW, LR, NL, WB
Data curation: CW, NL, TOL
Formal analysis: NL
Methodology: AJ, CW, LR, NL, WB
Supervision: AJ, CW, LR, WB
Writing – original draft: NL
Writing – review & editing: AJ, CW, LR, WB

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies, data dictionary, intervention characteristics, and screening flowchart.
[PDF File, 582 KB - cancer_v12i1e81579_app1.pdf]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 124 KB - cancer_v12i1e81579_app2.docx]

XSL•FO
RenderX

References

1. Northouse LL. Mastectomy patients and the fear of cancer recurrence. Cancer Nurs 1981 Jun;4(3):213-220. [Medline: 6909039]
2. Lebel S, Ozakinci G, Humphris G, et al. From normal response to clinical problem: definition and clinical features of fear of cancer recurrence. Support Care Cancer 2016 Aug;24(8):3265-3268. [doi: 10.1007/s00520-016-3272-5] [Medline: 27169703]
3. Deuning-Smit E, Custers JAE, Miroševič Š, et al. Prospective longitudinal study on fear of cancer recurrence in patients newly diagnosed with head and neck cancer: course, trajectories, and associated factors. Head Neck 2022 Apr;44(4):914-925. [doi: 10.1002/hed.26985] [Medline: 35084079]
4. Lee-Jones C, Humphris G, Dixon R, Hatcher MB. Fear of cancer recurrence—a literature review and proposed cognitive formulation to explain exacerbation of recurrence fears. Psychooncology 1997 Jun;6(2):95-105. [doi: 10.1002/(sici)1099-1611(199706)6:23.0.Co;2-b] [Medline: 9205967]
5. Babadostu MK, Eyrenci A. Investigation of the relationship between intolerance of uncertainty, metacognitions, maladaptive coping, and fear of cancer recurrence via moderated mediation model. Psychooncology 2025 Jan;34(1):e70076. [doi: 10.1002/pon.70076] [Medline: 39799468]
6. Fritzson E, Salafia C, Bellizzi KM, Park CL. Cascading pathways from physical symptom burden to distress in adults with cancer. Health Psychol 2025 Jan;44(1):57-65. [doi: 10.1037/hea0001420] [Medline: 39298210]
7. Simard S, Thewes B, Humphris G, et al. Fear of cancer recurrence in adult cancer survivors: a systematic review of quantitative studies. J Cancer Surviv 2013 Sep;7(3):300-322. [doi: 10.1007/s11764-013-0272-z] [Medline: 23475398]
8. Faraji A, Dehghani M, Khatibi A. Familial aspects of fear of cancer recurrence: current insights and knowledge gaps. Front Psychol 2023;14:1279098. [doi: 10.3389/fpsyg.2023.1279098] [Medline: 38034286]
9. Tran TXM, Jung SY, Lee EG, et al. Fear of cancer recurrence and its negative impact on health-related quality of life in long-term breast cancer survivors. Cancer Res Treat 2022 Oct;54(4):1065-1073. [doi: 10.4143/crt.2021.835] [Medline: 34883553]
10. Rha SY, Park JS, Choi YY, Hong B, Lee J. Fear of cancer recurrence and its predictors and outcomes among cancer survivors: a descriptive correlational study. Eur J Oncol Nurs 2022 Jun;58:102138. [doi: 10.1016/j.ejon.2022.102138] [Medline: 35490660]
11. Perndorfer C, Soriano EC, Siegel SD, Spencer RMC, Otto AK, Laurenceau JP. Fear of cancer recurrence and sleep in couples coping with early-stage breast cancer. Ann Behav Med 2022 Nov 5;56(11):1131-1143. [doi: 10.1093/abm/kaac018] [Medline: 35551585]
12. Butow PN, Fardell JE, Smith AB. Fear of cancer recurrence. In: Holland JC, Breitbart WS, Jacobsen PB, Loscalzo MJ, McCorkle R, Butow PN, editors. Psycho-Oncology: Oxford University Press; 2015.
13. Mehnert A, Koch U, Sundermann C, Dinkel A. Predictors of fear of recurrence in patients one year after cancer rehabilitation: a prospective study. Acta Oncol 2013 Aug;52(6):1102-1109. [doi: 10.3109/0284186X.2013.765063] [Medline: 23384721]
14. Zhang X, Sun D, Qin N, Liu M, Jiang N, Li X. Factors correlated with fear of cancer recurrence in cancer survivors: a meta-analysis. Cancer Nurs 2022;45(5):406-415. [doi: 10.1097/NCC.0000000000001020] [Medline: 34560707]
15. Butow PN, Sharpe LA, Thewes B, Turner J, Gilchrist J, Beith J. Fear of cancer recurrence: a practical guide for clinicians. Oncology (Williston Park) 2018 Jan 15;32(1):32-38. [Medline: 29447419]
16. Williams JTW, Pearce A. A systematic review of fear of cancer recurrence related healthcare use and intervention cost-effectiveness. Psychooncology 2021 Aug;30(8):1185-1195. [doi: 10.1002/pon.5673] [Medline: 33880822]
17. Prins JB, Deuning-Smit E, Custers JAE. Interventions addressing fear of cancer recurrence: challenges and future perspectives. Curr Opin Oncol 2022 Jul 1;34(4):279-284. [doi: 10.1097/CCO.0000000000000837] [Medline: 35730523]
18. Champagne A, Ivers H, Savard J. Utilization of health care services in cancer patients with elevated fear of cancer recurrence. Psychooncology 2018 Aug;27(8):1958-1964. [doi: 10.1002/pon.4748] [Medline: 29719072]
19. Zhu X, Lei J, Chen R, et al. Cancer recurrence fear and return to work in breast cancer survivors: the mediating effects of health literacy. J Multidiscip Healthc 2025;18:1031-1041. [doi: 10.2147/JMDH.S498387] [Medline: 40008288]
20. Schapira L, Zheng Y, Gelber SI, et al. Trajectories of fear of cancer recurrence in young breast cancer survivors. Cancer 2022 Jan 15;128(2):335-343. [doi: 10.1002/cncr.33921] [Medline: 34614212]
21. Smith AB, Fardell JE, Butow PN, et al. Fear of cancer recurrence. In: Breitbart W, Butow P, Jacobsen P, editors. Psycho-Oncology: Oxford University Press; 2021:755-768.
22. Savard J, Ivers H. The evolution of fear of cancer recurrence during the cancer care trajectory and its relationship with cancer characteristics. J Psychosom Res 2013 Apr;74(4):354-360. [doi: 10.1016/j.jpsychores.2012.12.013] [Medline: 23497839]
23. Cincidda C, Pizzoli SFM, Pravettoni G. Remote psychological interventions for fear of cancer recurrence: scoping review. JMIR Cancer 2022 Jan 11;8(1):e29745. [doi: 10.2196/29745] [Medline: 35014956]
24. Peters MDJ, Godfrey C, McInerney P, et al. Best practice guidance and reporting items for the development of scoping review protocols. JBI Evid Synth 2022 Apr 1;20(4):953-968. [doi: 10.11124/JBIES-21-00242] [Medline: 35102103]
25. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]

XSL•FO
RenderX

26. Li N. Interventions to reduce fear of cancer recurrence (FCR) among patients with colorectal cancer (CRC): a scoping review protocol. : OSF; 2024. [doi: 10.17605/OSF.IO/29YA3]
27. Covidence systematic review software. Veritas Health Innovation. 2024. URL: https://www.covidence.org/reviewers/ [accessed 2026-04-07]
28. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
29. AhmadiQaragezlou N, Rahimian Boogar I, Asadi J, Vojdani R. The effectiveness of mindfulness-based stress reduction on fear of cancer recurrence in colorectal and breast cancer patients: a randomized controlled trial. Middle East J Cancer 2020;11(2):216-224. [doi: 10.30476/mejc.2019.78695.0]
30. Davidson J, Malloch M, Humphris G. A single-session intervention (the Mini-AFTERc) for fear of cancer recurrence: a feasibility study. Psychooncology 2018 Nov;27(11):2668-2670. [doi: 10.1002/pon.4724] [Medline: 29624779]
31. Cohen PA, Musiello T, Jeffares S, Bennett K. Mindfulness-based cognitive therapy for fear of recurrence in ovarian cancer survivors (FROCS): a single-arm, open-label, pilot study. Support Care Cancer 2022 Mar;30(3):2317-2325. [doi: 10.1007/s00520-021-06659-y] [Medline: 34727225]
32. Lebel S, Maheu C, Lefebvre M, et al. Addressing fear of cancer recurrence among women with cancer: a feasibility and preliminary outcome study. J Cancer Surviv 2014 Sep;8(3):485-496. [doi: 10.1007/s11764-014-0357-3] [Medline: 24756313]
33. Peng L, Yang Y, Chen M, et al. Effects of an online mindfulness-based intervention on fear of cancer recurrence and quality of life among Chinese breast cancer survivors. Complement Ther Clin Pract 2022 Nov;49:101686. [doi: 10.1016/j.ctcp.2022.101686] [Medline: 36347151]
34. Smith A, Thewes B, Turner J, et al. Pilot of a theoretically grounded psychologist-delivered intervention for fear of cancer recurrence (Conquer Fear). Psychooncology 2015 Aug;24(8):967-970. [doi: 10.1002/pon.3775] [Medline: 25728489]
35. Tomei C, Lebel S, Maheu C, Mutsaers B. Addressing fear of recurrence: improving psychological care in cancer survivors. Support Care Cancer 2016 Jul;24(7):2815-2818. [doi: 10.1007/s00520-016-3103-8] [Medline: 26838024]
36. Ulfig CM. Cognitive behavioral effects on emotional well-being, worry about cancer progression, proinflammatory cytokines, and cancer recurrence among women with confirmed gynecologic malignancies [Dissertation]. : University of Florida; 2020 URL: https://www.proquest.com/openview/b154771b795c89ee6e04e72344ab0d06/1?pq-origsite=gscholar&cbl=18750&diss=y [accessed 2026-03-25]
37. Almeida SN, Elliott R, Silva ER, Sales CMD. Emotion-focused therapy for fear of cancer recurrence: a hospital-based exploratory outcome study. Psychotherapy (Chic) 2022 Jun;59(2):261-270. [doi: 10.1037/pst0000389] [Medline: 34990161]
38. Burm R, Thewes B, Rodwell L, et al. Long-term efficacy and cost-effectiveness of blended cognitive behavior therapy for high fear of recurrence in breast, prostate and colorectal cancer survivors: follow-up of the SWORD randomized controlled trial. BMC Cancer 2019 May 16;19(1):462. [doi: 10.1186/s12885-019-5615-3] [Medline: 31096934]
39. Deuning-Smit E, Custers JAE, Braam CIW, Hermens R, Prins JB. Toward implementation of an evidence-based intervention for fear of cancer recurrence: feasibility in real-world psycho-oncology practice. Psychooncology 2024 Jan;33(1):e6297. [doi: 10.1002/pon.6297] [Medline: 38282226]
40. Luigjes-Huizer YL, Helsper CW, de Wit NJ, van der Lee ML. Effectiveness of a guided online primary care intervention for fear of cancer recurrence: a randomised controlled trial. Psychooncology 2023 Dec;32(12):1839-1847. [doi: 10.1002/pon.6231] [Medline: 37850876]
41. McHale C, Cruickshank S, Brown T, et al. Mini-AFTERc: a controlled pilot trial of a nurse-led psychological intervention for fear of breast cancer recurrence. Pilot Feasibility Stud 2024 Jan 8;10(1):3. [doi: 10.1186/s40814-023-01431-x] [Medline: 38191445]
42. Pradhan P, Sharpe L, Butow PN, Smith AB, Russell H. Is a brief online booklet sufficient to reduce fear of cancer recurrence or progression in women with ovarian cancer? Front Psychol 2021;12(634136):634136. [doi: 10.3389/fpsyg.2021.634136] [Medline: 33716902]
43. Reb AM, Borneman T, Economou D, et al. A nurse-led intervention for fear of cancer progression in advanced cancer: a pilot feasibility study. Eur J Oncol Nurs 2020 Dec;49:101855. [doi: 10.1016/j.ejon.2020.101855] [Medline: 33120211]
44. Reb AM, Borneman T, Economou D, Cangin MA, Patel SK, Sharpe L. Fear of cancer progression: findings from case studies and a nurse-led intervention. Clin J Oncol Nurs 2020 Aug 1;24(4):400-408. [doi: 10.1188/20.CJON.400-408] [Medline: 32678373]
45. Savard J, Savard MH, Caplette-Gingras A, Casault L, Camateros C. Development and feasibility of a group cognitive-behavioral therapy for fear of cancer recurrence. Cogn Behav Pract 2018 May;25(2):275-285. [doi: 10.1016/j.cbpra.2017.08.001]
46. Sharpe L, Turner J, Fardell JE, et al. Psychological intervention (ConquerFear) for treating fear of cancer recurrence: mediators and moderators of treatment efficacy. J Cancer Surviv 2019 Oct;13(5):695-702. [doi: 10.1007/s11764-019-00788-4] [Medline: 31347010]
47. Smith A, Bamgboje-Ayodele A, Jegathees S, et al. Feasibility and preliminary efficacy of iConquerFear: a self-guided digital intervention for fear of cancer recurrence. J Cancer Surviv 2024 Apr;18(2):425-438. [doi: 10.1007/s11764-022-01233-9] [Medline: 35876964]

XSL•FO
RenderX

48. Tauber NM, O'Toole MS, Jensen AB, et al. ConquerFear-Group: a randomized controlled trial of an online-delivered group-based psychological intervention for fear of cancer recurrence in breast cancer survivors. Psychooncology 2023 Sep;32(9):1424-1432. [doi: 10.1002/pon.6193] [Medline: 37489745]
49. Tauber NM, Zachariae R, Jensen AB, et al. ConquerFear-Group: feasibility study with pilot results of a psychological intervention for fear of cancer recurrence delivered in groups. Psychooncology 2022 Jan;31(1):30-38. [doi: 10.1002/pon.5772] [Medline: 34289212]
50. Tomei C, Lebel S, Maheu C, Lefebvre M, Harris C. Examining the preliminary efficacy of an intervention for fear of cancer recurrence in female cancer survivors: a randomized controlled clinical trial pilot study. Support Care Cancer 2018 Aug;26(8):2751-2762. [doi: 10.1007/s00520-018-4097-1] [Medline: 29500582]
51. van de Wal M, Servaes P, Berry R, Thewes B, Prins J. Cognitive behavior therapy for fear of cancer recurrence: a case study. J Clin Psychol Med Settings 2018 Dec;25(4):390-407. [doi: 10.1007/s10880-018-9545-z] [Medline: 29468568]
52. van Helmondt SJ, Lodder P, van Woezik R, de Vries J, van der Lee ML. CBT-based online self-help training to reduce fear and distress after cancer (CAREST randomized trial): 24 months follow-up using latent growth models and latent class analysis. Ann Behav Med 2023 Aug 21;57(9):787-799. [doi: 10.1093/abm/kaac078] [Medline: 37078920]
53. Wang M, Gan C, Zhao J, et al. Positive influence of managing cancer and living meaningfully (CALM) on fear of cancer recurrence in breast cancer survivors. Am J Cancer Res 2023;13(7):3067-3079. [Medline: 37559986]
54. Zhao J, Wang M, Huang R, et al. Effects of CALM intervention on neutrophil-to-lymphocyte ratio (NLR), fear of cancer recurrence and quality of life in patients with lung cancer. Support Care Cancer 2023 Jul 7;31(7):447. [doi: 10.1007/s00520-023-07929-7] [Medline: 37414980]
55. Hasannezhad Reskati M, Hosseini S, Alizadeh-Navaei R, Khosravi S, Mirzaian B. Effectiveness of acceptance and commitment therapy management on self-differentiation and fear of disease progression in patients with breast cancer. J Babol Univ Med Sci 2020;22(1):110-118. [doi: 10.22088/jbums.22.1.110]
56. Herschbach P, Dinkel A. Fear of progression. Recent Results Cancer Res 2014;197(11-29):11-29. [doi: 10.1007/978-3-642-40187-9_2] [Medline: 24305766]
57. Vickberg SMJ. The Concerns About Recurrence Scale (CARS): a systematic measure of women's fears about the possibility of breast cancer recurrence. Ann Behav Med 2003;25(1):16-24. [doi: 10.1207/S15324796ABM2501_03] [Medline: 12581932]
58. Simard S, Savard J. Fear of Cancer Recurrence Inventory: development and initial validation of a multidimensional measure of fear of cancer recurrence. Support Care Cancer 2009 Mar;17(3):241-251. [doi: 10.1007/s00520-008-0444-y] [Medline: 18414902]
59. Fardell JE, Jones G, Smith AB, et al. Exploring the screening capacity of the Fear of Cancer Recurrence Inventory-Short Form for clinical levels of fear of cancer recurrence. Psychooncology 2018 Feb;27(2):492-499. [doi: 10.1002/pon.4516] [Medline: 28755462]
60. Akkol-Solakoglu S, Hevey D. Internet-delivered cognitive behavioural therapy for depression and anxiety in breast cancer survivors: results from a randomised controlled trial. Psychooncology 2023 Mar;32(3):446-456. [doi: 10.1002/pon.6097] [Medline: 36635249]
61. Akechi T, Momino K, Katsuki F, et al. Brief collaborative care intervention to reduce perceived unmet needs in highly distressed breast cancer patients: randomized controlled trial. Jpn J Clin Oncol 2021 Feb 8;51(2):244-251. [doi: 10.1093/jjco/hyaa166] [Medline: 32914169]
62. Akechi T, Yamaguchi T, Uchida M, et al. Smartphone psychotherapy reduces fear of cancer recurrence among breast cancer survivors: a fully decentralized randomized controlled clinical trial (J-SUPPORT 1703 Study). J Clin Oncol 2023 Feb 10;41(5):1069-1078. [doi: 10.1200/JCO.22.00699] [Medline: 36322882]
63. Ananeva E. Conquer fear of cancer recurrence program for breast cancer patients in Russia. Presented at: Annual Meeting and Supportive Care in Cancer; Jun 25-27, 2020. [doi: 10.1007/s00520-021-06285-8]
64. Arch JJ, Mitchell JL. An acceptance and commitment therapy (ACT) group intervention for cancer survivors experiencing anxiety at re-entry. Psychooncology 2016 May;25(5):610-615. [doi: 10.1002/pon.3890] [Medline: 26130586]
65. Arch JJ, Mitchell JL, Genung SR, et al. Randomized trial of acceptance and commitment therapy for anxious cancer survivors in community clinics: outcomes and moderators. J Consult Clin Psychol 2021 Apr;89(4):327-340. [doi: 10.1037/ccp0000630] [Medline: 34014694]
66. Arch JJ, Slivjak ET, Finkelstein LB. A novel intervention to reduce fear of progression and trauma symptoms in advanced cancer using written exposure to worst-case scenarios. J Palliat Med 2024 Aug;27(8):1009-1017. [doi: 10.1089/jpm.2023.0658] [Medline: 38579139]
67. Beith JM, Thewes B, Turner J, et al. Long-term results of a phase II randomized controlled trial (RCT) of a psychological intervention (Conquer Fear) to reduce clinical levels of fear of cancer recurrence in breast, colorectal, and melanoma cancer survivors. J Clin Oncol 2017;35(15 Suppl):LBA10000. [doi: 10.1200/JCO.2017.35.15_suppl.LBA1000]
68. Bergerot CD, Malhotra J, Bergerot PG, et al. Prospective assessment of a smartphone-app based mindfulness program for patients with metastatic renal cell carcinoma (mRCC). J Clin Oncol 2022 Feb 20;40(6_suppl):324-324. [doi: 10.1200/JCO.2022.40.6_suppl.324]

XSL•FO
RenderX

69. Bergerot CD, Malhotra J, Bergerot PG, et al. Improved health-related quality of life (HR-QOL) with use of an online mindfulness tool in patients with metastatic renal cell carcinoma (mRCC) receiving immunotherapy (IO). J Clin Oncol 2023 Feb 20;41(6_suppl):660-660. [doi: 10.1200/JCO.2023.41.6_suppl.660]
70. Bin Z, Ming Y, Puyu Z, Zhenkun K. In patients undergoing surgery for stomach cancer, the Rosenthal effect's impact on hope level, anxiety, and depression. Arch Clin Psychiatry 2023;50(3):184-188. [doi: 10.15761/0101-60830000000610]
71. Brooker J, Julian J, Millar J, et al. A feasibility and acceptability study of an adaptation of the Mindful Self-Compassion program for adult cancer patients. Palliat Support Care 2020 Apr;18(2):130-140. [doi: 10.1017/S1478951519000737] [Medline: 31595861]
72. Butow PN, Turner J, Gilchrist J, et al. Randomized trial of ConquerFear: a novel, theoretically based psychosocial intervention for fear of cancer recurrence. J Clin Oncol 2017 Dec 20;35(36):4066-4077. [doi: 10.1200/JCO.2017.73.1257] [Medline: 29095681]
73. Chambers SK, Foley E, Galt E, Ferguson M, Clutton S. Mindfulness groups for men with advanced prostate cancer: a pilot study to assess feasibility and effectiveness and the role of peer support. Support Care Cancer 2012 Jun;20(6):1183-1192. [doi: 10.1007/s00520-011-1195-8] [Medline: 21625914]
74. Chang YC. Effectiveness of integrated mindfulness-based fitness training program in ameliorating physical symptoms and survivor concerns in patients with breast cancer: a randomized clinical trial. Cancer Nurs 2025;48(3):159-170. [doi: 10.1097/NCC.0000000000001285] [Medline: 38158601]
75. Cheng KKF, Siah R, Zhang M, et al. Effect of an e-home based symptom management and mindfulness training program on quality of life in breast cancer survivors: A randomized clinical trial. J Clin Oncol 2021 May 20;39(15_suppl):12080-12080. [doi: 10.1200/JCO.2021.39.15_suppl.12080]
76. Cillessen L, Schellekens MPJ, Van de Ven MOM, et al. Consolidation and prediction of long-term treatment effect of group and online mindfulness-based cognitive therapy for distressed cancer patients. Acta Oncol 2018 Oct;57(10):1293-1302. [doi: 10.1080/0284186X.2018.1479071] [Medline: 29932784]
77. Compen F, Bisseling E, Schellekens M, et al. Face-to-face and internet-based mindfulness-based cognitive therapy compared with treatment as usual in reducing psychological distress in patients with cancer: a multicenter randomized controlled trial. J Clin Oncol 2018 Aug 10;36(23):2413-2421. [doi: 10.1200/JCO.2017.76.5669] [Medline: 29953304]
78. Crane-Okada R, Kiger H, Sugerman F, et al. Mindful movement program for older breast cancer survivors: a pilot study. Cancer Nurs 2012;35(4):E1-13. [doi: 10.1097/NCC.0b013e3182280f73] [Medline: 22705939]
79. Dieng M, Butow PN, Costa DSJ, et al. Psychoeducational intervention to reduce fear of cancer recurrence in people at high risk of developing another primary melanoma: results of a randomized controlled trial. J Clin Oncol 2016 Dec 20;34(36):4405-4414. [doi: 10.1200/JCO.2016.68.2278] [Medline: 27998215]
80. Dieng M, Khanna N, Kasparian NA, et al. Cost-effectiveness of a psycho-educational intervention targeting fear of cancer recurrence in people treated for early-stage melanoma. Appl Health Econ Health Policy 2019 Oct;17(5):669-681. [doi: 10.1007/s40258-019-00483-6] [Medline: 31228015]
81. Dieng M, Morton RL, Costa DSJ, et al. Benefits of a brief psychological intervention targeting fear of cancer recurrence in people at high risk of developing another melanoma: 12-month follow-up results of a randomized controlled trial. Br J Dermatol 2020 Apr;182(4):860-868. [doi: 10.1111/bjd.17990] [Medline: 30965384]
82. Dirkse D, Hadjistavropoulos HD, Alberts NA, et al. Making Internet-delivered cognitive behaviour therapy scalable for cancer survivors: a randomized non-inferiority trial of self-guided and technician-guided therapy. J Cancer Surviv 2020 Apr;14(2):211-225. [doi: 10.1007/s11764-019-00810-9] [Medline: 31853727]
83. Dodds SE, Pace TWW, Bell ML, et al. Feasibility of Cognitively-Based Compassion Training (CBCT) for breast cancer survivors: a randomized, wait list controlled pilot study. Support Care Cancer 2015 Dec;23(12):3599-3608. [doi: 10.1007/s00520-015-2888-1] [Medline: 26275769]
84. Döking S, Koulil SSV, Thewes B, Braamse AMJ, Custers JAE, Prins JB. Combined face-to-face and online cognitive-behavioral therapy for high distress of colorectal cancer survivors: a case study. Cogn Behav Pract 2021 Feb;28(1):107-123. [doi: 10.1016/j.cbpra.2020.06.008]
85. Eckert S, Morean W, Bentley J, McGregor B. Bridging the gap: testing the effectiveness of an empirically-based 11-week stress management intervention among cancer survivors in a non-academic setting. Psychooncology 2020;29(S1). [doi: 10.1016/j.brat.2023.104266]
86. Fishbein JN, Haslbeck J, Arch JJ. Network intervention analysis of anxiety-related outcomes and processes of acceptance and commitment therapy (ACT) for anxious cancer survivors. Behav Res Ther 2023 Mar;162:104266. [doi: 10.1016/j.brat.2023.104266] [Medline: 36739856]
87. Fishbein JN, Judd CM, Genung S, Stanton AL, Arch JJ. Intervention and mediation effects of target processes in a randomized controlled trial of acceptance and commitment therapy for anxious cancer survivors in community oncology clinics. Behav Res Ther 2022 Jun;153:104103. [doi: 10.1016/j.brat.2022.104103] [Medline: 35504144]
88. Fishbein JN, Arch JJ. Examining the effects of prior acceptance and commitment therapy (ACT) treatment among anxious cancer survivors during the COVID-19 pandemic: evidence from a randomized controlled trial. J Contextual Behav Sci 2022 Apr;24:31-37. [doi: 10.1016/j.jcbs.2022.03.006] [Medline: 35310680]

XSL•FO
RenderX

89. Fisher PL, Byrne A, Fairburn L, Ullmer H, Abbey G, Salmon P. Brief metacognitive therapy for emotional distress in adult cancer survivors. Front Psychol 2019;10:162. [doi: 10.3389/fpsyg.2019.00162] [Medline: 30766505]
90. Fisher PL, Byrne A, Salmon P. Metacognitive therapy for emotional distress in adult cancer survivors: a case series. Cognit Ther Res 2017;41(6):891-901. [doi: 10.1007/s10608-017-9862-9] [Medline: 29104332]
91. Frangou E, Bertelli G, Love S, et al. OVPSYCH2: a randomized controlled trial of psychological support versus standard of care following chemotherapy for ovarian cancer. Gynecol Oncol 2021 Aug;162(2):431-439. [doi: 10.1016/j.ygyno.2021.05.024] [Medline: 34059348]
92. Gonzalez-Hernandez E, Romero R, Campos D, et al. Cognitively-Based Compassion Training (CBCT®) in breast cancer survivors: a randomized clinical trial study. Integr Cancer Ther 2018 Sep;17(3):684-696. [doi: 10.1177/1534735418772095] [Medline: 29681185]
93. Hall DL, Yeh GY, O'Cleirigh C, et al. A multi-step approach to adapting a mind-body resiliency intervention for fear of cancer recurrence and uncertainty in survivorship (IN FOCUS). Glob Adv Health Med 2022;11. [doi: 10.1177/21649561221074690] [Medline: 35237466]
94. Hall DL, Park ER, Cheung T, Davis RB, Yeh GY. A pilot mind-body resiliency intervention targeting fear of recurrence among cancer survivors. J Psychosom Res 2020 Oct;137:110215. [doi: 10.1016/j.jpsychores.2020.110215] [Medline: 32818720]
95. Herschbach P, Berg P, Waadt S, et al. Group psychotherapy of dysfunctional fear of progression in patients with chronic arthritis or cancer. Psychother Psychosom 2010;79(1):31-38. [doi: 10.1159/000254903] [Medline: 19887889]
96. Humphris GM, Rogers SN. AFTER and beyond: cancer recurrence fears and a test of an intervention in oral and oropharyngeal patients. Soc Sci Dent 2012;2(1):29-38 [FREE Full text]
97. Imai F, Momino K, Katsuki F, et al. Smartphone problem-solving therapy to reduce fear of cancer recurrence among breast cancer survivors: an open single-arm pilot study. Jpn J Clin Oncol 2019 Jun 1;49(6):537-544. [doi: 10.1093/jjco/hyz005] [Medline: 30793156]
98. Johns SA, Stutz PV, Talib TL, et al. Acceptance and commitment therapy for breast cancer survivors with fear of cancer recurrence: a 3-arm pilot randomized controlled trial. Cancer 2020 Jan 1;126(1):211-218. [doi: 10.1002/cncr.32518] [Medline: 31539169]
99. Kacel EL. Feasibility of a comprehensive yoga intervention for fear of cancer recurrence in gynecologic cancers [Thesis]. : University of Florida; 2019 URL: https://ufdc.ufl.edu/es/UFE0055693/00001/citation [accessed 2026-03-25]
100. Lee JT, Lee YH, Chang YP. Mindfulness stress management for female cancer survivors facing the uncertainty of disease progression: a randomized controlled study. Int J Environ Res Public Health 2022 Apr 8;19(8):4497. [doi: 10.3390/ijerph19084497] [Medline: 35457366]
101. Lengacher CA, Reich RR, Paterson CL, et al. Examination of broad symptom improvement resulting from mindfulness-based stress reduction in breast cancer survivors: a randomized controlled trial. J Clin Oncol 2016 Aug 20;34(24):2827-2834. [doi: 10.1200/JCO.2015.65.7874] [Medline: 27247219]
102. Lengacher CA, Gruss LF, Kip KE, et al. Mindfulness-based stress reduction for breast cancer survivors (MBSR(BC)): evaluating mediators of psychological and physical outcomes in a large randomized controlled trial. J Behav Med 2021 Oct;44(5):591-604. [doi: 10.1007/s10865-021-00214-0] [Medline: 33963420]
103. Lengacher CA, Johnson-Mallard V, Post-White J, et al. Randomized controlled trial of mindfulness-based stress reduction (MBSR) for survivors of breast cancer. Psychooncology 2009 Dec;18(12):1261-1272. [doi: 10.1002/pon.1529] [Medline: 19235193]
104. Lengacher CA, Reich RR, Ramesar S, et al. Feasibility of the mobile mindfulness-based stress reduction for breast cancer (mMBSR(BC)) program for symptom improvement among breast cancer survivors. Psychooncology 2018 Feb;27(2):524-531. [doi: 10.1002/pon.4491] [Medline: 28665541]
105. Lichtenthal WG, Corner GW, Slivjak ET, et al. A pilot randomized controlled trial of cognitive bias modification to reduce fear of breast cancer recurrence. Cancer 2017 Apr 15;123(8):1424-1433. [doi: 10.1002/cncr.30478] [Medline: 28055119]
106. Li X, Chen L, Lei B, Xie C. Home-based psychological nursing interventions for improvement of sleep quality and psychological health in patients with hypopharyngeal carcinoma undergoing surgical resections: a randomized trial. Ann Palliat Med 2021 Dec;10(12):12347-12357. [doi: 10.21037/apm-21-3029] [Medline: 35016426]
107. Liu J, Butow P, Bui KT, et al. Novel clinician-lead intervention to address fear of cancer recurrence in breast cancer survivors. JCO Oncol Pract 2021 Jun;17(6):e774-e784. [doi: 10.1200/OP.20.00799] [Medline: 33571035]
108. Loughan AR, Willis KD, Braun SE, et al. Managing cancer and living meaningfully (CALM) in adults with malignant glioma: a proof-of-concept phase IIa trial. J Neurooncol 2022 May;157(3):447-456. [doi: 10.1007/s11060-022-03988-8] [Medline: 35437687]
109. Loughan A, Willis K, Lanoye A, et al. Exploring CALM in patients with high-grade glioma. Neuro-oncology 2021;23(Suppl 6). [doi: 10.1093/neuonc/noab196.728]
110. Luberto CM, Hall DL, Chad-Friedman E, Park ER. Theoretical rationale and case illustration of mindfulness-based cognitive therapy for fear of cancer recurrence. J Clin Psychol Med Settings 2019 Dec;26(4):449-460. [doi: 10.1007/s10880-019-09610-w] [Medline: 30756278]

XSL•FO
RenderX

111. Maheu C, Lebel S, Bernstein LJ, et al. Fear of cancer recurrence therapy (FORT): a randomized controlled trial. Health Psychol 2023 Mar;42(3):182-194. [doi: 10.1037/hea0001253] [Medline: 36862474]
112. Manne SL, Virtue SM, Ozga M, et al. A comparison of two psychological interventions for newly-diagnosed gynecological cancer patients. Gynecol Oncol 2017 Feb;144(2):354-362. [doi: 10.1016/j.ygyno.2016.11.025] [Medline: 27887806]
113. Martin F, Wright H, Moody L, et al. Help to overcome problems effectively for cancer survivors: development and evaluation of a digital self-management program. J Med Internet Res 2020 May 19;22(5):e17824. [doi: 10.2196/17824] [Medline: 32209529]
114. Merckaert I, Lewis F, Delevallez F, et al. Improving anxiety regulation in patients with breast cancer at the beginning of the survivorship period: a randomized clinical trial comparing the benefits of single-component and multiple-component group interventions. Psychooncology 2017 Aug;26(8):1147-1154. [doi: 10.1002/pon.4294] [Medline: 27718533]
115. Momino K, Mitsunori M, Yamashita H, et al. Collaborative care intervention for the perceived care needs of women with breast cancer undergoing adjuvant therapy after surgery: a feasibility study. Jpn J Clin Oncol 2017 Mar 1;47(3):213-220. [doi: 10.1093/jjco/hyw189] [Medline: 28003321]
116. Montesinos M, Francisco L. Acceptance of relapse fears in breast cáncer patients: effects of an ACT-based abridged intervention [Spanish]. Psychooncology 2016;13(1):7-21. [doi: 10.5209/rev_PSIC.2016.v13.n1.52484]
117. Murphy MJ, Newby JM, Butow P, et al. Randomised controlled trial of internet-delivered cognitive behaviour therapy for clinical depression and/or anxiety in cancer survivors (iCanADAPT Early). Psychooncology 2020 Jan;29(1):76-85. [doi: 10.1002/pon.5267] [Medline: 31659822]
118. Neubert S, Schlecht S, Meng K, Rabe A, Jentschke E. Effects of a video sequence based intervention on anxiety, fatigue and depression in cancer patients: results of a randomized controlled trial. Integr Cancer Ther 2023;22. [doi: 10.1177/15347354231153172] [Medline: 36799503]
119. Nguyen P, Heisey R, Quenneville C, et al. An examination of depression, anxiety, and fear of recurrence among cancer survivors who participated in a virtual cognitive behavioral therapy (CBT)-based telephone coaching program. Support Care Cancer 2022 Sep;30(9):7323-7332. [doi: 10.1007/s00520-022-07148-6] [Medline: 35606479]
120. Otto AK. Goal pursuit mediates effects of a gratitude intervention on death-related fear of recurrence in breast cancer survivors [Thesis]. : University of Delaware; 2015 URL: https://www.proquest.com/pqdtglobal1/dissertations-theses/goal-pursuit-mediates-effects gratitude/docview/1718184550/sem-2?accountid=8424 [accessed 2026-04-07]
121. Park S, Sato Y, Takita Y, et al. Mindfulness-based cognitive therapy for psychological distress, fear of cancer recurrence, fatigue, spiritual well-being, and quality of life in patients with breast cancer – a randomized controlled trial. J Pain Symptom Manage 2020 Aug;60(2):381-389. [doi: 10.1016/j.jpainsymman.2020.02.017] [Medline: 32105790]
122. Pradhan P, Sharpe L, Menzies RE. Towards a stepped care model for managing fear of cancer recurrence or progression in cancer survivors. Cancer Manag Res 2021;13:8953-8965. [doi: 10.2147/CMAR.S294114] [Medline: 34880676]
123. Smith AB, Rincones O, Heiniger L, et al. Online psychological self-management intervention for testicular cancer survivors (e-TC 2.0): promising acceptability and preliminary efficacy, but limited feasibility. J Psychosoc Oncol Res Pract 2020;2(3):e33. [doi: 10.1097/OR9.0000000000000033]
124. Rudolph B, Wünsch A, Herschbach P, Dinkel A. Ambulante verhaltenstherapeutische Gruppentherapie zur Behandlung von Progredienzangst bei Krebspatienten [Cognitive-behavioral group therapy addressing fear of progression in cancer out-patients]. Psychother Psych Med Psychol 2018 Jan;68(1):38-43. [doi: 10.1055/s-0043-107774] [Medline: 28718864]
125. Russell L, Ugalde A, Orellana L, et al. A pilot randomised controlled trial of an online mindfulness-based program for people diagnosed with melanoma. Support Care Cancer 2019 Jul;27(7):2735-2746. [doi: 10.1007/s00520-018-4574-6] [Medline: 30506103]
126. Sajadian A, Zahrakar K, Asadpour E. Effectiveness of transpersonal therapy (spiritual therapy, yoga-meditation) in reducing fear of cancer recurrence in breast cancer survivors: a randomized controlled trial [Iran]. Iran Q J Breast Dis 2021;14(2):50-63. [doi: 10.30699/ijbd.14.2.50]
127. Saltbæk L, Bidstrup PE, Karlsen RV, et al. Nurse-led individualized follow-up versus regular physician-led visits after early breast cancer (MyHealth): a phase III randomized, controlled trial. J Clin Oncol 2024 Jun 10;42(17):2038-2049. [doi: 10.1200/JCO.23.01447] [Medline: 38498781]
128. Salazar-Alejo MG, Ferrigno AS, Gutiérrez-Ornelas J, et al. 339P An online mindfulness-based stress-reduction (MBSR) intervention for breast cancer (BC) survivors: a randomized trial. Ann Oncol 2023 Oct;34(Suppl 2):S318. [doi: 10.1016/j.annonc.2023.09.1068]
129. Sarizadeh MS, Mozaffari S, Boogar IR. Effectiveness of acceptance and commitment therapy on the fear of cancer recurrence and post-traumatic growth among patients with breast cancer. Koomesh 2018;20(4):e153003 [FREE Full text]
130. Sauer C, Maatouk I. Entwicklung und Evaluation einer Akzeptanz- und Commitment-Therapie (ACT)-basierten Gruppenintervention für psychisch belastete onkologische Patienten [in German]. Psychother Psychosom Med Psychol 2021 Feb;71(2):90-98. [doi: 10.1055/a-1174-0283] [Medline: 32634840]
131. Sakai M, Kondo M, Sugiura T, Akechi T. Acceptance and commitment therapy in the transdiagnostic treatment of a breast cancer survivor: a case study. Jpn Psychol Res 2024 Jul;66(3):264-275. [doi: 10.1111/jpr.12422]

XSL•FO
RenderX

132. Schlecht S, Neubert S, Meng K, Rabe A, Jentschke E. Changes of symptoms of anxiety, depression, and fatigue in cancer patients 3 months after a video-based intervention. Int J Environ Res Public Health 2023 Oct 17;20(20):6933. [doi: 10.3390/ijerph20206933] [Medline: 37887671]
133. Shih TF, Turner J, Butow P, et al. They say the odds are good, but I can't stop thinking about it: responding to fear of cancer recurrence. Presented at: Proceedings of the 41st Annual Scientific Meeting of the Clinical Oncology Society of Australia; Dec 2-4, 2014 URL: https://hdl.handle.net/10536/DRO/DU:30087918 [accessed 2026-04-07]
134. Shumay D, Fujiye MT, Berman M, Gregg J, Melisko M, Dunn L. WiLLoW: an acceptance and commitment therapy intervention for breast cancer survivors to address anxiety, worry and fear of cancer recurrence. Psychooncology 2013;22:26-27. [doi: 10.10002/pon.3245]
135. Sinclair F, Gillanders D, Rooney N, et al. Real-world evaluation of an acceptance and commitment therapy-based group programme for breast cancer survivors with fear of cancer recurrence. Support Care Cancer 2023 Nov 15;31(12):700. [doi: 10.1007/s00520-023-08179-3] [Medline: 37966674]
136. Sinclair F, Rooney N, Hewitt C, McLoone P. Therapeutic radiographer and clinical psychology group-based pilot study: acceptance and commitment therapy (ACT) and patient education intervention to address fear of cancer recurrence following treatment for a primary breast cancer. Radiography (Lond) 2020 Jan;26:S3-S4. [doi: 10.1016/j.radi.2019.11.012]
137. Steinecke M, Joachims J, Kotzerke M, Lueke C, Berg J, Zimmermann T. Dealing with fear of cancer recurrence – evaluation of a psychooncological group therapy program. Oncol Res Treat 2022;45(Suppl 1):203. [doi: 10.1159/000521004]
138. Subnis UB. Stress management for cancer survivors using a technologically adapted psychosocial intervention: a randomized trial determining the effect of expressive writing on psychoneuroimmunology based outcomes [Dissertation]. : Virginia Commonwealth University; 2014. [doi: 10.25772/Y47T-7550]
139. Sun L, Pang Y, Wang Z, et al. Effect of traditional Chinese medicine combined group psychotherapy on psychological distress management and gut micro-biome regulation for colorectal cancer survivors: a single-arm phase I clinical trial. Support Care Cancer 2023 Nov 15;31(12):698. [doi: 10.1007/s00520-023-08131-5] [Medline: 37964024]
140. Thewes B, Bell M, Smith A, Fardell J, Turner J, Butow P. Conquer Fear: the development and pilot testing of a novel psychological treatment for fear of cancer recurrence. Asia Pac J Clin Oncol 2012;8(Suppl 3):207. [doi: 10.1111/ajco.12029]
141. Tomei C, Lebel S, Maheu C, et al. Examining the preliminary efficacy of an intervention for fear of cancer recurrence in female cancer survivors: a randomized controlled clinical trial pilot study. Support Care Cancer 2018;26:2751-2762. [doi: 10.1007/s00520-018-4097-1]
142. Tran MJ, Jefford M, Fua T, et al. Feasibility and acceptability of the fear-less screening and stratified-care model for fear of cancer recurrence among people affected by early-stage cancer. Psychooncology 2025 Feb;34(2):e70070. [doi: 10.1002/pon.70070] [Medline: 39953998]
143. Vâlcu EG, Firescu D, Nechita A, et al. Considerations regarding online group psychotherapy sessions for breast cancer patients in active phase of oncological treatment. Healthcare (Basel) 2023 Aug 16;11(16):2311. [doi: 10.3390/healthcare11162311] [Medline: 37628508]
144. van de Wal M, Servaes P, Berry R, et al. Cognitive behavior therapy for fear of cancer recurrence: a case study. J Clin Psychol Med Settings 2018;25:390-407. [doi: 10.1007/s10880-018-9545-z]
145. van de Wal M, Thewes B, Gielissen M, Speckens A, Prins J. Efficacy of blended cognitive behavior therapy for high fear of recurrence in breast, prostate, and colorectal cancer survivors: the SWORD study, a randomized controlled trial. J Clin Oncol 2017 Jul 1;35(19):2173-2183. [doi: 10.1200/JCO.2016.70.5301] [Medline: 28471726]
146. Victorson D, Du H, Hankin V, et al. Mindfulness based stress reduction decreases fear of progression over time for men with prostate cancer on active surveillance: results from a randomized clinical trial. J Urol 2012;187(4S):e157. [doi: 10.1016/j.juro.2012.02.447]
147. Wagner LI, Tooze JA, Hall DL, et al. Targeted eHealth intervention to reduce breast cancer survivors' fear of recurrence: results from the FoRtitude randomized trial. J Natl Cancer Inst 2021 Nov 2;113(11):1495-1505. [doi: 10.1093/jnci/djab100] [Medline: 34057469]
148. Wang S, Zhu Y, Wang Z, et al. Efficacy of meaning-centered group psychotherapy in Chinese patients with cancer: a randomized controlled trial. Palliat Support Care 2023 Oct;21(5):773-781. [doi: 10.1017/S1478951523000998] [Medline: 37558651]
149. Weis JB, Gschwendtner K, Giesler JM, Adams L, Wirtz MA. Psychoeducational group intervention for breast cancer survivors: a non-randomized multi-center pilot study. Support Care Cancer 2020 Jul;28(7):3033-3040. [doi: 10.1007/s00520-019-05076-6] [Medline: 31576471]
150. Yoon J, Lee H, Son H. Effects of an interactive coaching intervention on quality of life and psychological factors for colorectal cancer survivors: a single group pre and posttest design. Eur J Oncol Nurs 2023 Oct;66:102413. [doi: 10.1016/j.ejon.2023.102413] [Medline: 37776600]
151. Zhang X, Yao S, Wang M, et al. The impact of VR-CALM intervention based on VR on psychological distress and symptom management in breast cancer survivors. J Oncol 2022;2022:1012813. [doi: 10.1155/2022/1012813] [Medline: 35712124]
152. Zhao YH, Xu Y. Effect of comprehensive nursing based on cognitive behavior on psychological function of glioma patients. Neuropsychiatr Dis Treat 2021;17:777-785. [doi: 10.2147/NDT.S286164] [Medline: 33731997]

XSL•FO
RenderX

153. Martinez MR, Pasha A. Prioritizing mental health research in cancer patients and survivors. AMA J Ethics 2017 May 1;19(5):486-492. [doi: 10.1001/journalofethics.2017.19.5.msoc2-1705] [Medline: 28553906]
154. Ulfig CM. Cognitive behavioral effects on emotional well-being, worry about cancer progression, proinflammatory cytokines, and cancer recurrence among women with confirmed gynecologic malignancies [Dissertation]. : University of Florida; 2020 URL: https://ufdc.ufl.edu/ufe0056666/00001 [accessed 2026-04-07]
155. Herschbach P, Berg P, Waadt S, et al. Group psychotherapy of dysfunctional fear of progression in patients with chronic arthritis or cancer. Psychother Psychosom 2009;79(1):31-38. [doi: 10.1159/000254903]
156. Nissen ER, O'Connor M, Kaldo V, et al. Internet-delivered mindfulness-based cognitive therapy for anxiety and depression in cancer survivors: a randomized controlled trial. Psychooncology 2020 Jan;29(1):68-75. [doi: 10.1002/pon.5237] [Medline: 31600414]
157. Frangou E, Bertelli G, Love S, et al. OVPSYCH2: a randomised study of psychological support versus standard of care following chemotherapy for ovarian cancer. Gynecol Oncol 2021;162:431-439. [doi: 10.1016/j.ygyno.2021.05.024]
158. Otto AK, Szczesny EC, Soriano EC, Laurenceau JP, Siegel SD. Effects of a randomized gratitude intervention on death-related fear of recurrence in breast cancer survivors. Health Psychol 2016 Dec;35(12):1320-1328. [doi: 10.1037/hea0000400] [Medline: 27513475]
159. Dinkel A, Marten-Mittag B, Kremsreiter K. Association between daily worry, pathological worry, and fear of progression in patients with cancer. Front Psychol 2021;12(648623):648623. [doi: 10.3389/fpsyg.2021.648623] [Medline: 34456783]
160. Kenny-Jones M, Nankivell P, Mehanna H, Ozakinci G, PETNECK2 Research Team. Fear of cancer recurrence and fear of cancer progression, digital resource engagement and health literacy: a review. Curr Oncol 2024 Nov 29;31(12):7586-7602. [doi: 10.3390/curroncol31120559] [Medline: 39727682]
161. Coutts-Bain D, Sharpe L, Pradhan P, Russell H, Heathcote LC, Costa D. Are fear of cancer recurrence and fear of progression equivalent constructs? Psychooncology 2022 Aug;31(8):1381-1389. [doi: 10.1002/pon.5944] [Medline: 35470502]
162. Lebel S, Zwaal C, Craig L, et al. Fear of cancer recurrence guideline. : Ontario Health (Cancer Care Ontario); 2024 Mar 1 Program in Evidence-Based Care Guideline No. 19-7 URL: https://www.breastcancer.org.nz/sites/default/files/FCR%20guideline%20Ontario%202024.pdf [accessed 2026-03-26]
163. Paperák P, Javůrková A, Raudenská J. Therapeutic intervention in fear of cancer recurrence in adult oncology patients: a systematic review. J Cancer Surviv 2023 Aug;17(4):1017-1035. [doi: 10.1007/s11764-022-01277-x] [Medline: 36307611]
164. Tauber NM, O'Toole MS, Dinkel A, et al. Effect of psychological intervention on fear of cancer recurrence: a systematic review and meta-analysis. J Clin Oncol 2019 Nov 1;37(31):2899-2915. [doi: 10.1200/JCO.19.00572] [Medline: 31532725]
165. Hall DL, Wagner LI, Lebel S, Smith AB, Bergerot CD, Park ER. Guidelines needed for the management of fear of cancer recurrence in adult survivors of cancer in the United States: a consensus statement. Cancer 2024 Aug 15;130(16):2739-2742. [doi: 10.1002/cncr.35326] [Medline: 38630904]
166. Hofmann SG, Asmundson GJG. Acceptance and mindfulness-based therapy: new wave or old hat? Clin Psychol Rev 2008 Jan;28(1):1-16. [doi: 10.1016/j.cpr.2007.09.003] [Medline: 17904260]
167. Deuning-Smit E. Off the shelf: toward implementation of an evidence-based psychological intervention for fear of cancer recurrence [Dissertation]. : Radboud University; 2025 URL: https://hdl.handle.net/2066/317679 [accessed 2026-04-07]
168. Lebel S, Ozakinci G, Humphris G, et al. Current state and future prospects of research on fear of cancer recurrence. Psychooncology 2017 Apr;26(4):424-427. [doi: 10.1002/pon.4103] [Medline: 26891602]
169. Zomerdijk N, Jongenelis M, Short CE, Smith A, Turner J, Huntley K. Prevalence and correlates of psychological distress, unmet supportive care needs, and fear of cancer recurrence among haematological cancer patients during the COVID-19 pandemic. Support Care Cancer 2021 Dec;29(12):7755-7764. [doi: 10.1007/s00520-021-06369-5] [Medline: 34164740]
170. Sharpe L, Thewes B, Butow P. Current directions in research and treatment of fear of cancer recurrence. Curr Opin Support Palliat Care 2017 Sep;11(3):191-196. [doi: 10.1097/SPC.0000000000000288] [Medline: 28665813]
171. Levit LA, Byatt L, Lyss AP, et al. Closing the rural cancer care gap: three institutional approaches. JCO Oncol Pract 2020 Jul;16(7):422-430. [doi: 10.1200/OP.20.00174] [Medline: 32574128]
172. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. J Med Internet Res 2017 Aug 24;19(8):e295. [doi: 10.2196/jmir.7740] [Medline: 28838887]
173. Dobson R, Carter K, Cutfield R, et al. Diabetes text-message self-management support program (SMS4BG): a pilot study. JMIR Mhealth Uhealth 2015 Mar 25;3(1):e32. [doi: 10.2196/mhealth.3988] [Medline: 25830952]
174. Chen J, Sun Y, Shao Y, et al. Effects of psychological interventions on fear of cancer recurrence: a systematic review and network meta‐analysis. Precis Med Sci 2024;13(2):84-98. [doi: 10.1002/prm2.12131]
175. Nair AS. Publication bias - Importance of studies with negative results!. Indian J Anaesth 2019 Jun;63(6):505-507. [doi: 10.4103/ija.IJA_142_19] [Medline: 31263309]
176. Sterling TD, Rosenbaum WL, Weinkam JJ. Publication decisions revisited: the effect of the outcome of statistical tests on the decision to publish and vice versa. Am Stat 1995 Feb;49(1):108-112. [doi: 10.1080/00031305.1995.10476125]
177. Hinkel J, Heneghan C, Bankhead C. Selective outcome reporting in cancer studies: a scoping review. medRxiv. Preprint posted online on Jul 3, 2024. [doi: 10.1101/2024.07.02.24309826]
178. Higgins JPT, Thomas J, editors. Cochrane Handbook for Systematic Reviews of Interventions: Wiley; 2019. [doi: 10.1002/9781119536604]

XSL•FO
RenderX

179. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]

Abbreviations

ACT: acceptance and commitment therapy
CBT: cognitive behavioral therapy
CENTRAL: Cochrane Central Register of Controlled Trials
FCR: fear of cancer recurrence
FoP: fear of progression
MBI: mindfulness-based intervention
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
RCT: randomized controlled trial

Edited by M Balcarras, N Cahill; submitted 31.Jul.2025; peer-reviewed by E Bantum, M Pritchard; revised version received 11.Dec.2025; accepted 05.Jan.2026; published 22.Apr.2026.

Please cite as:
Li N, Wensley C, Reynolds L, Baraza W, Leyton TO, Jull A
Interventions to Reduce Fear of Cancer Recurrence Among People With Cancer: Scoping Review
JMIR Cancer 2026;12:e81579
URL: https://cancer.jmir.org/2026/1/e81579
doi:10.2196/81579

© Niu Li, Cynthia Wensley, Lisa Reynolds, Wal Baraza, Tatiana Osorio Leyton, Andrew Jull. Originally published in JMIR Cancer (https://cancer.jmir.org), 22.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Advanced Clinical-Based Technologies for Monitoring Physical Function in Breast Cancer Survivors: Scoping Review

Mahtab Azhdar[1], MSc; Adalberto Loyola-Sanchez[2], MD; Amber Wardrop[1], MSc; Martin Ferguson-Pell[3], Prof Dr

[1]Rehabilitation Sciences, Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada

[2]Division of Physical Medicine and Rehabilitation, Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

[3]Department of Physical Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Martin Ferguson-Pell, Prof Dr
Department of Physical Therapy
Faculty of Rehabilitation Medicine
University of Alberta
11405 87 Avenue NW
Edmonton, AB, T6G 1C9
Canada
Phone: 1 7809384368
Email: fe4@ualberta.ca

Abstract

Background: People surviving breast cancer often face long-term impairments in physical function, significantly impacting their quality of life. In recent years, a variety of technologies have been developed to monitor and assess these functions; however, there is no consolidated synthesis linking specific technologies to targeted functional domains and real-world clinical contexts, limiting comparability and translation into practice.

Objective: This scoping review aimed to systematically explore and map the use of advanced clinic-based technologies for assessing and monitoring key physical functions, such as balance, muscle strength, and range of motion, among individuals surviving breast cancer. The purpose of this review was not only to identify which technologies have been applied but also to clarify how they are being used, the clinical settings, target physical functions, assessment protocols, and types of outcomes measured. It further summarized the current patterns of use to inform and enhance clinical assessment practices.

Methods: A comprehensive literature search was conducted across MEDLINE, Scopus, CINAHL, and Web of Science databases, with no publication date restrictions. Eligible studies included adults with breast cancer assessed using advanced clinic-based technologies to monitor physical function. Screening and selection followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The data extraction captured study characteristics, participant demographics, technologies applied, and related outcomes. The extracted data were organized in Covidence and synthesized descriptively to map the types of technologies, assessed functional domains, and application settings across studies.

Results: Across the 17 included studies, the participants (N=719; age range between 30 and 75 years) were predominantly female and largely drawn from stage 0 to III breast cancer cohorts; 1 (5.9%) study reported a single male participant, and 2 (11.8%) studies did not specify participant sex. Among the 17 included studies, 11 (64.7%) were published from 2017 onward. Technologies spanned balance platforms (force plates, Technobody-PK 200 WL, Sensory Organization Test; 5/17, 29.4%), isokinetic dynamometry (Biodex systems; 4/17, 23.5%), and range of motion assessment via motion capture (3/17, 17.6%) or digital inclinometers (5/17, 29.4%). Sample sizes per study ranged from 20 to 100 participants (median 43), and follow-up durations varied from 1 session to 6 months.

Conclusions: Advanced clinic-based technologies for assessing balance, muscle strength, and range of motion in breast cancer survivors were identified across the literature, including balance platforms, isokinetic dynamometry, digital inclinometers, and markerless motion capture systems. Considerable heterogeneity in devices, outcome reporting, and study designs limited direct comparison across studies and prevented definitive conclusions about the superiority or clinical readiness of any single technology.

(JMIR Cancer 2026;12:e77894) doi:10.2196/77894

XSL•FO
RenderX

KEYWORDS

breast cancer; technology; functional assessment; markerless motion capture; isokinetic dynamometry

Introduction

Breast cancer remains the most prevalent malignancy among women worldwide, with increasing survival rates due to advances in early detection and treatment [1]. Although breast cancer mortality has declined in recent decades, as survival rates improve, attention has shifted from survival alone to long-term recovery, quality of life, and function, with rehabilitation emerging as a central component of survivorship care [2-4]. Physical function, the ability to perform physical tasks that enable daily activities and participation, is an important factor in cancer survivorship and rehabilitation and is framed within the International Classification of Functioning, Disability, and Health (ICF) [5]. Impairments in physical function after breast cancer treatment are commonly reported and include deficits in balance, muscle strength, and range of motion (ROM) [6-8]. Women undergoing chemotherapy may experience up to a 25% loss in strength and joint dysfunction [8]. Moreover, individuals surviving cancer have highlighted difficulties with balance and walking as the most common functional issues, with prevalence rates of 19% and 24%, respectively [9]. These functional limitations have been associated with poorer health-related quality of life and reduced mobility in survivorship cohorts [10,11]. At the same time, impairments in gait and balance control have been documented as potential contributors to fall risk and reduced independence in everyday life among breast cancer survivors [6]. According to the literature on cancer survivorship, over half of individuals who have undergone cancer treatments encounter physical function impairments [12].

The American Physical Therapy Association (APTA) Oncology Evaluation Database to Guide Effectiveness (EDGE) Task Force has provided evidence-based recommendations for standardized outcome measures in oncology rehabilitation. For balance assessment, the Task Force strongly supports the use of low-cost, performance-based tools such as the Fullerton Advanced Balance Scale, gait speed, Timed Up and Go, Five Times Sit-to-Stand, and the Balance Evaluation Systems Test, all rated as reliable and clinically feasible for cancer survivors [13]. In contrast, computerized balance systems such as force plates and the Sensory Organization Test (SOT) have been less recommended due to limited clinical utility and high cost, despite growing evidence supporting their sensitivity in detecting subtle postural sway and vestibular deficits in individuals with cancer and chemotherapy-induced peripheral neuropathy [14]. For shoulder ROM, the Task Force rated passive goniometry (score 4) as a recommended tool, while for muscle strength, handheld dynamometers (HHDs) (score 3) and manual muscle testing (score 2B) were endorsed as appropriate clinical measures [15]. Although these conventional methods remain the clinical standard, their limited sensitivity and responsiveness underscore the need for more objective, automated technologies capable of quantifying subtle changes in function [16,17].

Furthermore, a variety of technologies such as wearable sensors (accelerometers/pedometers), fitness trackers, smartphone apps, and advanced motion-capture systems are increasingly used to quantify physical function in breast cancer survivorship [18-22], providing objective, high-resolution data [21,22]. However, most reported applications remain in research or specialized settings with limited protocol and outcome standardization [19,22].

Accordingly, this review aims to map and characterize the use of such advanced, clinically based technologies in assessing key physical functions, including balance, muscle strength, and ROM, among individuals surviving breast cancer. For the purposes of this review, “advanced” is defined as instruments that provide automated, objective outputs beyond unaided observation or analogue readouts. “Clinic-based” denotes systems that can be operated in clinical rooms or clinically configured spaces by routine clinical staff with minimal specialist engineering support, for monitoring and assessing physical functions. These terms were selected based on their relevance within the ICF framework and their frequent association with impairment among breast cancer survivors [6,23,24]. By examining how these technologies are currently integrated into clinical practice, this review seeks to identify gaps in the existing literature and highlight areas where further research is needed.

Methods

Search Strategy

A comprehensive literature search was conducted across the MEDLINE, Scopus, CINAHL, and Web of Science databases. The initial search strategy was developed in collaboration with an experienced academic health sciences librarian (Liz Dennett) to ensure methodological rigor and comprehensive coverage. The selection of the search terms was informed by preliminary scoping searches and key indexing terms from prior reviews in oncology rehabilitation and motion analysis. The MEDLINE strategy served as the base and was adapted for syntax variations across databases. The search was conducted in multiple stages, starting with the initial search on February 2, 2024, followed by updates on June 24, 2024. Each search was repeated across the 4 databases to ensure comprehensive coverage of the relevant literature. The final search included combinations of controlled vocabulary, such as MeSH (Medical Subject Headings) terms and free-text terms related to breast cancer, physical function, and measurement technologies. The complete search strategy and related keywords are provided in Multimedia Appendix 1. In addition to peer-reviewed databases, gray-literature searches were limited to PhD dissertations. All identified articles were imported into Covidence software (Covidence Ltd) for screening. Duplicate records were automatically identified and removed by Covidence’s built-in algorithm, followed by manual verification by the reviewers to ensure accuracy.

Inclusion and Exclusion Criteria

Studies were included if they were published in English, had full-text availability online, involved adults diagnosed with breast cancer, and utilized technologies to objectively measure

XSL•FO
RenderX

one or more physical functions (balance, strength, and ROM) as primary or secondary outcomes. No publication date restrictions were applied. All study designs were considered for inclusion, excluding nonoriginal articles such as study protocols, reviews, conference abstracts, books, or editorials. Additionally, studies were excluded if they involved technologies that did not meet the criteria for being both advanced and clinically based. These technologies are intended to directly assess and monitor physical functions in patients, facilitating clinical decision-making and supporting rehabilitation processes. Technologies like Vicon, for instance, were excluded if they were not designed for clinical use. To be effective in clinical settings, technologies must be easily integrable into workflows without requiring complex setups or specialized technical expertise.

Screening and Selection Process

An initial search strategy was developed in collaboration with the academic librarian to capture studies using any technology to assess physical function in adults with breast cancer. During the title and abstract screening process, it became evident that this comprehensive approach would yield several hundred eligible studies, predominantly due to the extensive literature on accelerometers and pedometers, as well as the frequent application of laboratory-based marker-based motion-capture systems. To preserve methodological rigor and avoid redundancy with existing systematic reviews and meta-analyses focused on accelerometer and pedometer studies, while ensuring a clinically actionable synthesis, an additional consultation with the librarian was conducted to refine the eligibility criteria prior to initiating full-text review and data extraction. The inclusion criteria were subsequently refined to prioritize advanced, clinically deployable technologies for point-of-care assessment of 3 core domains of physical functions. As a result, articles that used traditional technologies, such as dynamometers or goniometers, which do not meet the established definition of advanced measurement technology for the purposes of this review; nonclinic–based accelerometers and pedometers; and some lab-based motion capture systems requiring reflective markers, multicamera stereophotogrammetry, calibration routines, and dedicated laboratory infrastructure, were excluded unless the authors explicitly described clinical deployment or adaptation for real-world clinical settings. The study selection process was conducted in several stages following established scoping review methodology. Initially, two reviewers (authors MA and AW) independently screened titles and abstracts of all retrieved references against the eligibility criteria. Discrepancies between the two reviewers at this stage were flagged within Covidence for subsequent resolution. Studies deemed potentially eligible by either reviewer advanced to full-text screening, which was similarly conducted independently by the same two reviewers (MA and AW). Following both screening phases, conflicts were systematically resolved through a structured consensus process. For any study where the two primary reviewers disagreed on inclusion or exclusion, the full text and relevant eligibility criteria were reviewed collaboratively by all three reviewers (MA, AW, and MFP) in a consensus meeting. During these discussions, each reviewer presented their rationale, and discrepancies were resolved through deliberation until unanimous agreement was reached. If consensus could not be achieved through discussion, the third reviewer (MFP) served as the final arbiter. All final inclusion decisions and data extraction were completed through this consensus process, with regular meetings held among the three reviewers to ensure consistency in the interpretation of eligibility criteria and data extraction procedures throughout the review.

Data Extraction

Data extraction was collaboratively performed by the three reviewers (MA, AW, and MFP). For each included study, details regarding the study title, year of publication, country, aim, and design were recorded. Participant characteristics, including demographic information, inclusion and exclusion criteria, and recruitment methods and settings, were extracted. Information on sample size, cancer treatment details, and the instruments employed, including technological characteristics, was collected. We also recorded methodological features relevant to implementation and key findings reported by the original authors in relation to the use or outcomes of the measurement technology. The extracted data were organized and managed in Covidence for transparency and synthesis.

Data Synthesis

The extracted data from relevant studies were synthesized narratively and summarized in tabular format to provide a comprehensive overview of advanced clinic-based technologies used to monitor and assess functional outcomes in individuals surviving breast cancer. Studies were first categorized by the component of physical function assessed (balance, strength, and ROM) and then organized according to the specific measurement technology employed. For each technology identified, the corresponding data analysis methods used in the original studies were extracted and documented. This synthesis sought to deliver an analysis of the current knowledge and methodologies used to assess the included physical functions. The findings were analyzed to highlight common methodologies, instruments, and outcomes reported in the literature. This process defines how functional outcomes are commonly evaluated in individuals with breast cancer using advanced and clinic-based technologies and identifies potential gaps or inconsistencies in current practices, including variations in measurement protocols, differences in data normalization methods, and the range of technologies applied to assess similar functional outcomes.

Results

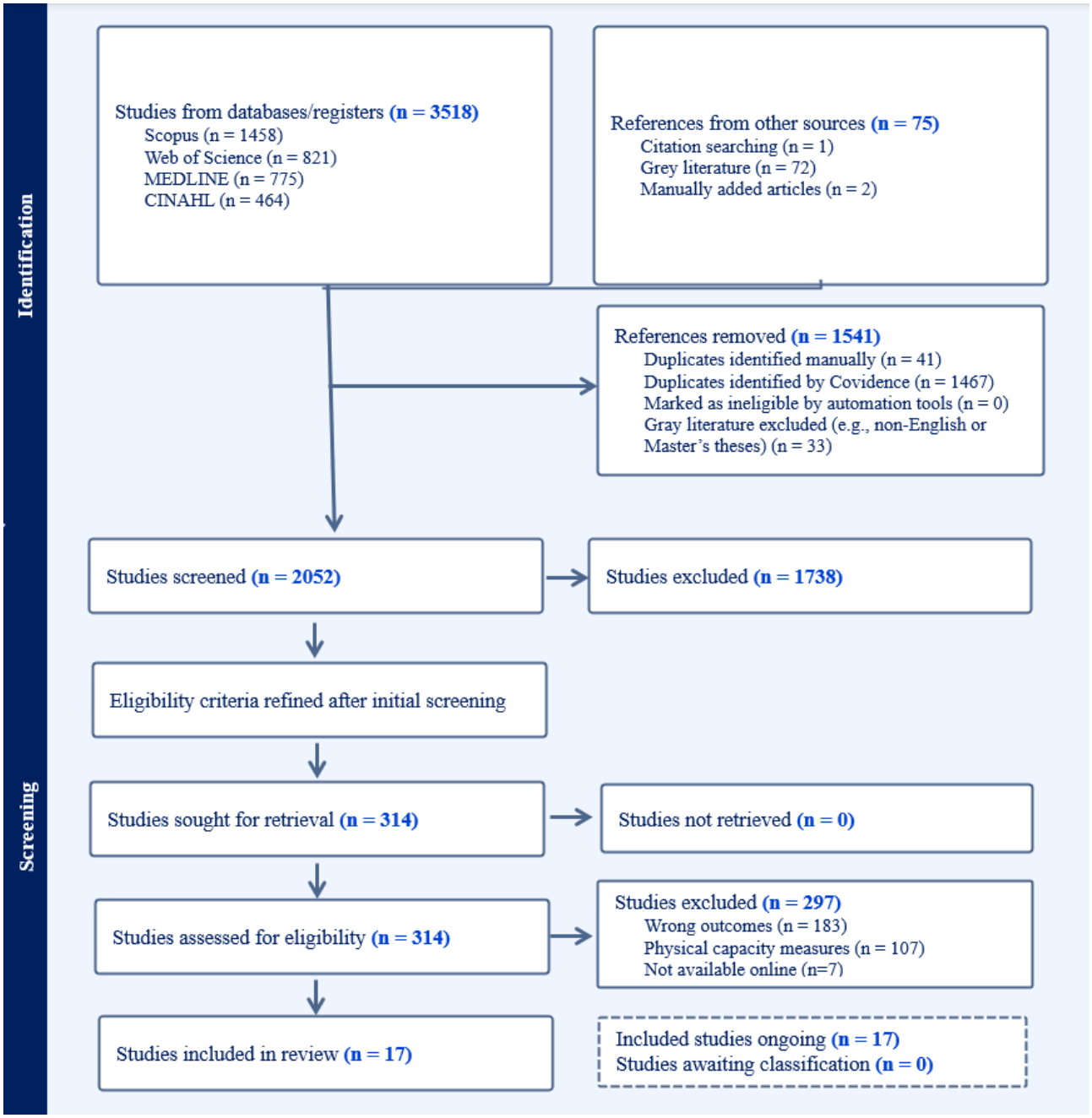
Study Identification

Initially, 3593 articles were retrieved from Medline, Scopus, Web of Science, CINAHL, and gray literature. After removing duplicates and nonrelevant gray literature (n=1508), a total of 2052 records were screened for eligibility. Following title and abstract screening, 314 articles were eligible for full-text review. Ultimately, 17 studies, each assigned a unique study ID (ranging from 001 to 017), were included for data extraction. These 17 studies, involving a total of 719 participants, were included in the review. The number of articles and methodologies utilized for reviewing, selecting, and verifying them at each stage of the

XSL•FO
RenderX

process is presented in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 1) [25]. While gray literature, including PhD dissertations, was considered, none met the inclusion criteria. Furthermore, the specific characteristics of the studies included in this study are comprehensively detailed in Table 1, providing an in-depth overview of their methodologies, sample populations, and key findings. Additionally, no formal quality assessment of the studies was conducted as part of this scoping review.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram for study selection.



XSL•FO
RenderX

Table 1. Characteristics of the included studies.

ID	Author (s)	Pub. Year	Title	Study setting and design	Sample size and participant characteristics	Physical function	Instruments	Key findings
001	(Bertoli et al) [26]	2023	Mat Pilates improves lower and upper body strength and flexibility in breast cancer survivors undergoing hormone therapy: a randomized controlled trial (HAPiMat study).	Hospital and randomized controlled trial	n=43; stage 0-III; ≥40 years; undergoing hormone therapy; sex not reported	Hip flexor and extensor muscle strength	Biodex Medical system 4 (Shirley)	Biodex assessments showed that the Pilates group significantly improved their lower body strength, with increases in isometric flexor-extensor peak torque and enhanced concentric and eccentric flexor peak torque and mechanical work.
002	(Rao and Pattanshetty) [27]	2022	Effect of myofascial release, stretching, and strengthening on upper torso posture, spinal curvatures, range of motion, strength, shoulder pain and disability, and quality of life in breast cancer survivors.	Hospital and pre - post experimental study	n=22; female, 35-70 years; newly diagnosed; treated with surgery, radiation, chemotherapy, or combination	Shoulder and cervical range of motion	Cervical range of motion was evaluated using a digital inclinometer	Significant improvements were observed for cervical movements.
003	(Højvig et al) [28]	2022	Donor-site morbidity following breast reconstruction with a latissimus dorsi flap - A prospective study.	Hospital and prospective observational study	n=20; female, 32-70 years; undergoing delayed breast reconstruction with latissimus dorsi flap	Shoulder Strength	Biodex System4 Pro-dynamometer	The study revealed that breast reconstruction using the LD[a] flap led to significant decreases in isometric shoulder strength for adduction and extension due to LD muscle removal, supporting its impact on shoulder girdle strength; however, isokinetic strength remained largely unchanged 12 weeks after surgery.
004	(Zabi˙t Özdemir and İyi˙gün) [29]	2022	Is there a difference in balance functions between breast cancer survivor women and healthy women?	Eastern Mediterranean University's Healthy Living Center and A prospective study	n=66; female, 35-70 years; 33 postmastectomy survivors vs 33 healthy controls; no chemo/radiation	Dynamic balance	Computer-based Dynamic Balance Platform (Technobody-PK 200 WL)	No differences were observed between the groups in subparameter dynamic balance measurements, computer-based dynamic balance platform assessments, or Y-Balance Test results.

XSL•FO
RenderX

ID	Author (s)	Pub. Year	Title	Study setting and design	Sample size and participant characteristics	Physical function	Instruments	Key findings
005	(Wechsler et al) [30]	2022	Persistent cancer-related fatigue predicts static and dynamic balance in women with a history of breast cancer	A breast cancer center at a large urban hospital and cross-sectional study	n=43; female, 30-85 years; postchemotherapy ± radiation; stage I-III	Static and dynamic balance	Force plates (Bertec Corporation model 4060NC)	CRF[b] independently and significantly impairs both static and dynamic balance individuals surviving cancer, leading to compensatory stabilization strategies and highlighting CRF's critical role in increased postural sway and fall risk even years after treatment.
006	(Artese et al) [31]	2021	Effect of functional impact training on body composition, bone mineral density, and strength in breast cancer survivors	Laboratory and randomized controlled trial	n=44; sedentary postmenopausal women, 52-68 years; stage 0-III; ≥3 months postchemotherapy/radiation.	Isokinetic concentric knee extension and flexion	Biodex Medical System 3 (Shirley)	Both groups demonstrated significant improvements over time in lower body strength, specifically in isokinetic knee extension and flexion at all tested speeds. However, after adjusting for baseline differences and time since diagnosis, there were no significant differences in posttraining knee strength measures between the two groups.
007	(Evans et al) [32]	2021	Examination of clinical and laboratory measures of static and dynamic balance in breast cancer survivors	In the Department of Rehabilitation Services at Alamance Regional Medical Center and a cross-sectional study	n=43 (20 breast cancer survivors + 23 controls); female, 40-70 years; stage 0-III; completed chemo/radiation ≤ 5 years	Static balance	SOT[c] was conducted using the NeuroCom SMART Balance Master (Natus Medical, Pleasanton)	NeuroCom SOT equilibrium assessments revealed that individuals surviving breast cancer generally maintained similar postural stability to the healthy control group. However, they showed significantly impaired balance in Conditions 2 and 3, which rely on proprioceptive and vestibular systems for maintaining balance.

XSL•FO
RenderX

ID	Author (s)	Pub. Year	Title	Study setting and design	Sample size and participant characteristics	Physical function	Instruments	Key findings
008	(Uhm et al) [33]	2020	Usefulness of Kinect sensor-based reachable workspace system for assessing upper extremity dysfunction in breast cancer patients	Konkuk University Medical Center and a cross-sectional study	n=20; age 46-62 years; unilateral breast cancer; sex not specified	Upper extremity active range of motion	The Kinect sensor-based reachable workspace analysis system	The reachable workspace was divided into four shoulder-centered quadrants: upper medial (Q1), lower medial (Q2), upper lateral (Q3), and lower lateral (Q4). The analysis revealed that the upper quadrants (1 and 3) on the affected side had significantly smaller reachable workspace areas compared to the unaffected side, while the lower quadrants showed no differences.
009	(Ribeiro et al) [34]	2019	Three-dimensional scapular kinematics, shoulder outcome measures and quality of life following treatment for breast cancer - a case control study	An outpatient breast cancer surgery (part of a hospital) and case control study	n=42; female, 40-60 years; 21 presurgery vs 21 controls	Shoulder range of motion	Shoulder motion measured by a digital inclinometer	The breast cancer surgery group exhibited reduced range of motion compared to healthy controls.
010	(Monfort et al) [35]	2017	Gait, balance, and patient-reported outcomes during taxane-based chemotherapy in early-stage breast cancer patients	An outpatient oncology clinic setting and longitudinal study	n=33; 32 female, 1 male; 36-59 years; Stage I-3	Static balance	Balance plate (Bertec Corp)	Cumulative exposure to taxane therapy was associated with notable declines in patients' balance, indicating a detrimental effect on their physical functionality.
011	(De Groef et al) [36]	2017	Effect of myofascial techniques for treatment of upper limb dysfunctions in breast cancer survivors: randomized controlled trial	Department of Physical Medicine and Rehabilitation of the University Hospitals Leuven and randomized controlled trial	n=48; female, 38-70 years; primary breast cancer; 23 intervention vs 25 control	Active shoulder range of motion	Inclinometer	The study found no significant differences between the group receiving myofascial therapy combined with physical therapy and the group receiving physical therapy alone in terms of shoulder range of motion.
012	(De Groef et al) [37]	2016	Arm lymphoedema and upper limb impairments in sentinel node-negative breast cancer patients: a one-year follow-up study	Multidisciplinary Breast Centre of University Hospitals Leuven and longitudinal study	n=100; female, 50-70 y	Shoulder range of motion	Gravity inclinometer	One year after the sentinel lymph node biopsy, 30% of patients experienced reduced shoulder range of motion.

XSL•FO
RenderX

ID	Author (s)	Pub. Year	Title	Study setting and design	Sample size and participant characteristics	Physical function	Instruments	Key findings
013	(Moreira et al) [38]	2015	A Kinect-based system for upper-body function assessment in breast cancer patients	Not specified and an observational study	n=48; female; 24 with lymphedema; age not reported	Upper-body joint range of motion includes shoulder, elbow, and wrist flexion/extension and shoulder abduction/adduction	Kinect-based system	The study validated a Kinect-based system for assessing upper-body function, demonstrating high accuracy in classifying normal vs. impaired function. The system, which uses kinematic data for machine learning classification, shows potential for remote monitoring and early detection of functional impairments during rehabilitation.
014	(Gritsenko et al) [39]	2015	Feasibility of using low-cost motion capture for automated screening of shoulder motion limitation after breast cancer surgery	Academic cancer center oncology clinic and descriptive study	n=20; women, 51-69 years; stage 0-III	Active and passive shoulder range of motion	Motion capture by Kinect	The study found that the low-cost Kinect motion capture system effectively identified moderate to severe shoulder motion impairments in individuals surviving breast cancer, with strong correlations to goniometric measurements for active movements.
015	(Winters-Stone et al) [40]	2011	Identifying factors associated with falls in postmenopausal breast cancer survivors: a multidisciplinary approach	Comprehensive cancer center and case-control plus prospective observation	n=59; female, 49-68 years; stage 0-III	Dynamic balance	The SOT, used in computerized dynamic posturography	Individuals surviving breast cancer had higher fall rates due to vestibular balance deficits from chemotherapy, particularly affecting dynamic balance, while static balance remained similar between fallers and non-fallers.
016	(Harrington et al) [41]	2011	Comparison of shoulder flexibility, strength, and function between breast cancer survivors and healthy participants	The Neuromuscular Research Laboratory and case-control study	n=48; female, 40-60 years; stage 0-III	Active and passive shoulder range of motion	Digital inclinometer	The study found that breast cancer survivors had reduced shoulder range of motion compared to healthy controls, particularly in flexion and external rotation.
017	(Waltman et al) [42]	2003	Testing an intervention for preventing osteoporosis in postmenopausal breast cancer survivors	In participants' homes or at convenient sites and pilot intervention study	n=21; female, 40-65 years; stage I-II	Muscle strength of the knee, hip, and wrist (flexion and extension)	Biodex System 2 multijoint testing	The study found significant improvements in muscle strength for hip flexion, hip extension, and knee flexion over 12 months.

[a]LD: latissimus dorsi.

[b]CRF: cancer-related fatigue.

[c]SOT: Sensory Organization Test.

XSL•FO
RenderX

Characteristics of the Included Studies

A review of 17 studies focusing on individuals surviving breast cancer revealed key insights into research locations, designs, timelines, participant stages, and objectives. The United States emerged as the most common location, hosting 7 (41.2%) of the included studies. The remaining research was conducted across a diverse set of countries, including Finland, Turkey, Portugal, Belgium, Brazil, South Korea, India, and Denmark. The review included 5 (29.4%) experimental studies, among which were 3 (17.6%) randomized controlled trials [26,31,36], 1 (5.9%) pretest-posttest experimental study [27], and 1 (5.9%) pilot intervention, [42] and 12 (70.6%) observational studies, including 4/17 (23.5%) cross-sectional studies [30,32,33,38], 3 (17.6%) case-control studies [34,40,41], 2 (11.8%) prospective observational studies [28,29], 2 (11.8%) longitudinal cohort studies [35,37], and 1 (5.9%) descriptive study [39]. The publication timeline indicated a growing interest in this area, with 11 (64.7%) studies published in 2017 or later, while the remaining 6 (35.3%) studies were published between 2003 and 2016 [37-42]. The studies aimed to assess different physical functions, with a particular emphasis on balance, strength, and flexibility related to breast cancer treatments. The objectives of the studies were varied, ranging from evaluating the effectiveness of specific interventions, such as mat Pilates [26] and myofascial techniques [36], to assessing the impact of cancer-related fatigue [30] on balance and postural control. Some studies also investigated the feasibility and usefulness of motion capture technologies for screening and assessing functional impairments [33,38,39].

Furthermore, the studies included in the review were conducted in multiple settings, predominantly within specialized medical facilities. Many studies took place in hospitals or hospital-affiliated centers. Isokinetic dynamometry using Biodex was reported mainly in hospital settings [26,28], with 1 (5.9%) study conducted in a laboratory environment [31]. Digital inclinometers were used across hospitals, rehabilitation services, and outpatient oncology clinics [34,36,37,41]. Kinect-based markerless motion capture (MMC) was implemented within oncology services spanning a medical center and an outpatient oncology clinic [33,38,39]. One (5.9%) study even extended its setting to participants' homes or convenient community sites [42].

In terms of recruitment strategies, many studies enrolled participants through state and hospital cancer registries, clinician referrals, and ongoing treatment facilities within large urban hospitals. Researchers also leveraged community outreach by utilizing social media platforms, distributing flyers in regional cancer hospitals, and engaging with breast cancer support groups. Local media outlets, such as newspapers and radio announcements, were employed to broaden their reach. Some studies relied on word-of-mouth, emails, and telephone contacts.

The studies included in this review primarily focused on individuals surviving breast cancer, with participants aged between 30 and 75 years, and an average age range of 50 to 60 years. Among the 15 (88.2%) studies that reported participant sex, all included only female participants except for 1 (6.7%) study, which included a single male participant [35]. A total of 2 (11.8%) studies did not report any information about participant sex [26,33]. Nine (53%) of the included studies explicitly reported participants' cancer stage; all 9 enrolled individuals with stage 0 to III disease [26,30-32,35,39-42]. Among the 17 studies, 5 (29.4%) studies involved participants undergoing or having recently completed chemotherapy [27,30-32,41], and 1 (5.9%) study specifically enrolled women receiving hormone therapy [26]. Moreover, 1 (5.9%) study focused on surgical reconstruction (latissimus dorsi flap) [28]. Two (11.8%) studies included postmastectomy survivors compared to healthy controls [29,32], and 3 (17.6%) enrolled mixed-treatment cohorts [27,30,40]. In terms of sample sizes, these varied considerably throughout the studies, ranging from as few as 20 participants to as many as 100. People with stage IV cancer and cognitive impairments were generally excluded. Additional exclusion criteria, such as uncontrolled cardiovascular or musculoskeletal conditions, severe neuropathy, or inability to provide informed consent, were applied to ensure participant safety and maintain the validity of assessment results.

Reported Instruments and Key Findings on Physical Function

This scoping review found that most studies assessed physical function using a variety of outcome measures, and ROM was reported in 8 (47.1%) of the included studies [27,33,34,36-39,41]. Balance outcomes were reported in 5 (29.4%) studies [29,30,32,35,40], and muscle strength outcomes were reported in 4 (23.5%) studies [26,28,31,42]. Of these, 3 (17.6%) studies employed motion capture technologies [33,38,39], Kinect-based systems to assess upper extremity kinematics and ROM in individuals surviving breast cancer. Inclinometers were used in 5 (29.4%) studies to further quantify shoulder movements [27,34,36,37,41]. Balance assessments were another key focus: 2 (11.8%) studies employed force plates to measure postural sway in both static and dynamic conditions [30,35], while the SOT was used in another 2 (11.8%) studies to assess balance under varying sensory inputs [32,40]. In addition, 1 (5.9%) prospective study evaluated dynamic balance using a computer-based dynamic balance platform (Technobody PK 200 WL) [29].

Moreover, strength assessments were conducted using different models of the Biodex System (2, 3, 4, and 4 Pro) to evaluate isometric and isokinetic strength in muscle groups, including the shoulder, hip, knee, and wrist [26,28,31,42]. Measurements included isometric strength normalized to body weight, as well as peak torque during isokinetic contractions at varying angles and velocities. Further details on the specific tools and data analysis methods used in each study are provided in Table 2.

XSL•FO
RenderX

Table 2. Overview of technologies and data analysis methods.

Category and technology used	Data analysis method	Author
Strength		
Biodex System 2 multijoint testing	The highest peak torque measure, obtained during 4 repetitions, was recorded in Newton meters and used for data analysis.	Waltman et al [42]
Biodex Medical System 3	The highest peak torque value from the 3 repetitions was recorded for each speed.	Artese et al [31]
Biodex Medical Systems 4	Peak torque and mechanical work, normalized to body mass, were analyzed.	Bertoli et al [26]
Biodex System 4 pro	The data were analyzed by averaging 3 trials of isometric and isokinetic strength (in Nm/kg).	Højvig et al [28]
Digital inclinometer	Active and passive shoulder flexion/extension and external/internal rotation range of motion were measured with 3 trials averaged for data analysis.	Harrington et al [41]; Rao and Pattanshetty [27]; Ribeiro et al [34]
Gravity inclinometer	The analysis assessed the prevalence of impaired shoulder range of motion, defined as an interlimb difference of 15 degrees or more between the affected and unaffected arms. Additionally, it evaluated the number of patients who experienced a decrease in shoulder ROM[a] greater than 15 degrees from their baseline measurements, indicating a significant loss of mobility after treatment or surgery.	De Groef et al [36,37]
Range of motion		
Kinect-based system	The study used a Kinect-based system to capture 3D motion data, extracting features like range of motion, hand height, elbow flexion, and movement acceleration. These were analyzed using machine learning algorithms to classify patients as having normal or impaired upper-body function.	Moreira et al [38]
Automated motion analysis system using Microsoft Kinect	The Microsoft Kinect sensor was used to assess shoulder motion limitations. The technology tracked body landmarks and converted the captured data into joint angles using a custom algorithm.	Gritsenko et al [39]
Microsoft Kinect Sensor	The Microsoft Kinect sensor captured participants' upper extremity motion trajectories during standardized seated arm movements. These data allowed for the reconstruction of each participant's reachable workspace envelope, divided into 4 quadrants relative to the shoulder joint. The reachable surface areas for each quadrant and the total workspace were calculated and then normalized by individual arm length to account for differences among participants.	Uhm et al [33]
Force plate model 4060NC (Bertec Corp) and recorded through Motion Monitor Software (Innsport Training, Inc)	Balance was assessed using a force plate to measure postural sway in the medial-lateral and anterior-posterior planes under both static and dynamic conditions. For static balance evaluation, participants stood still for 30 seconds before and after moderate-intensity exercise. Dynamic balance was measured during the rising phase of an Instrumented Sit-to-Stand test, where participants quickly stood up from a seated position, and sway was recorded throughout this transition.	Wechsler et al [30]
Balance		
Balance plate (Bertec Corp)	Data collected from the CoP[b] included measurements of its location and displacement during standing trials, specifically focusing on the root mean squared excursion in the medial-lateral direction to evaluate postural stability and the risk of falling.	Monfort et al [35]
Technobody-PK 200 WL, a computer-based dynamic balance platform	Participants performed the “Equilibrium Assessment” and “Sleight Assessment” tests on this device, which measured parameters such as anterior/posterior and medial/lateral sleight, balance assessments, number of targets reached, perimeter, and average pace. Each test was repeated 3 times, and the best score was recorded for analysis.	Zabi˙t Özdemi˙r and İyi˙gün [29]
Computerized dynamic posturography with the SOT	Participants' sway responses were recorded under 6 sensory conditions that manipulate visual and somatosensory information.	Winters-Stone et al [40]
NeuroCom SOT[c]	Equilibrium scores from the NeuroCom SOT under 6 conditions varying platform stability and visual input, with and without the serial sevens cognitive task, were analyzed to assess static balance and the impact of cognitive load.	Evans et al [32]

[a]ROM: range of motion.

[b]CoP: center of pressure.

[c]SOT: Sensory Organization Test.

XSL•FO
RenderX

Discussion

Principal Findings

This scoping review mapped 17 studies deploying advanced clinic-based technologies to assess physical function in breast cancer survivors. The review identified technologies used to assess three domains, namely ROM, balance, and muscle strength, and documented variation in instrument type. These technologies were applied in varied clinical environments, primarily hospital-based and outpatient oncology settings, to measure specific aspects of physical function through structured protocols.

The studies used advanced methods to determine how people maintain their balance. These detailed assessments consistently revealed information that simpler, one-time tests often overlook. Studies using advanced balance equipment in clinical settings measured different aspects of postural control and consistently revealed information that simpler, one-time tests often overlook. SOT protocols showed balance problems when proprioceptive and vestibular systems were challenged, even though performance was nearly normal under easier conditions. This suggests that balance deficits in breast cancer survivors depend on the situation rather than affecting all balance tasks equally [32,40]. This pattern supports the clinical value of multi-condition batteries that probe sensory reweighting, rather than relying solely on single-task screens. Force plates quantify quiet-stance stability by deriving center of pressure (CoP) signals and summarizing them with standard sway metrics, typically medial-lateral (ML) and anterior-posterior (AP) excursion (range or root mean square (RMS)), total path length, mean sway velocity, and planar sway area (often a 95% confidence ellipse). These metrics capture both the magnitude (eg, excursion, area) and temporal dynamics (eg, velocity) of postural control. Notably, key ML CoP metrics, including mean velocity, mean amplitude, and RMS displacement, have been linked to higher prospective fall risk [43]. In the included longitudinal study of taxane-based chemotherapy, Monfort et al [35] used a laboratory force plate to track changes in CoP behavior during treatment; cumulative taxane exposure was associated with deteriorations in balance control, reflected by increased sway and concurrent gait alterations (shorter step length, slower walking speed). Complementing this, Wechsler et al [30] showed that cancer-related fatigue independently predicted poorer postural stability on force-plate measurements. Fatigued individuals showed compensatory stabilization strategies and greater body sway in both standing still and moving conditions. Together, these findings suggest that when clinically relevant stressors are present (eg, neurotoxic chemotherapy and fatigue), force-plate CoP metrics could be sensitive to subtle balance changes.

For measuring ROM, the review describes a spectrum of technologies. At one end are digital inclinometers, which are widely accessible but depend on the operator. At the other end are automated MMC systems, which provide less biased 3D assessments and show promising evidence of clinical usefulness. The successful application of MMC for upper extremity ROM assessment in breast cancer survivors was demonstrated across 3 studies in diverse settings, including oncology clinics and outpatient centers. This suggests these technologies have overcome initial feasibility barriers and may be ready for broader clinical implementation [33,38,39]. The strong correlation between Kinect-derived reachable workspace metrics and QuickDASH, a validated and widely used tool for assessing upper extremity disability in breast cancer survivors [44], and the system's ability to detect side-to-side differences in functionally relevant movement zones (overhead reaching) demonstrate practical utility for monitoring recovery [33]. However, there is significant variation in how different studies measure ROM. Studies differ in whether they assess active versus passive motion, which planes or tasks they evaluate, and whether they use absolute measurements versus side-to-side comparisons. This inconsistency makes it difficult to compare studies and develop clear, evidence-based clinical guidelines. Taken together, within the bounds of the included studies, MMC appears to have promising potential for assessing upper-extremity function in breast cancer survivors.

Muscle strength was quantified with isokinetic dynamometry in controlled clinical or laboratory settings [26,28,31,42]. Systems such as the Biodex (models 2-4/4 Pro) are seen as the top choice for testing dynamic muscle strength. They control precise, preset angular velocities and capture full torque-angle-velocity relationships together with work and power outputs. These multidimensional profiles can reveal deficits in force production and velocity-dependent behavior (eg, concentric versus eccentric weakness) that are not apparent from single-value manual grades, and their test-retest reliability in musculoskeletal [45,46]. The trade-off is practical; true isokinetic testing requires dedicated equipment, space, regular calibration, and staff skilled in positioning, stabilization, familiarization, and protocol standardization, all of which limit routine deployment outside well-resourced centers. Given these constraints, many rehabilitation services rely on portable HHDs for strength assessment. Professional guidance from the APTA Oncology EDGE Task Force recognizes HHDs as an appropriate option for cancer populations, provided that clinicians use standardized patient and tester positioning, consistent lever arms, and repeated trials to improve reliability [15].

This scoping review had some limitations. First, the inclusion of solely English-language studies raised the possibility of language bias and the exclusion of relevant work published in other languages. Additionally, the review focused solely on 3 physical functions, potentially overlooking other important aspects of physical function in individuals surviving breast cancer. While multiple databases were searched, some relevant studies may not have been captured if they were published in databases that were not included in the search strategy, which can limit the comprehensiveness of the review. Finally, the variability of the included studies in design, clinical setting, treatment phase, and outcome definitions (eg, active versus passive ROM, task/plane selection, and absolute versus interlimb metrics) precluded quantitative synthesis and limits cross-study comparability. Specifically, individuals with stage IV cancer were typically excluded from most of the included studies to focus on nonmetastatic cases.

XSL•FO
RenderX

Comparison With Prior Work

A recent scoping review by Amarelo et al (2024) [47] examined the application of technological resources in cancer rehabilitation, identifying wearable devices, web-based platforms, mobile health (mHealth) apps, virtual reality, and exergaming as commonly employed tools. Their review emphasized the wide diversity of technologies and underscored the need for further research to assess their long-term effectiveness, cost-efficiency, and successful integration into clinical practice. In contrast, our scoping review narrows the focus to advanced clinic-based technologies specifically used to assess and monitor physical functions, balance, muscle strength, and ROM in individuals surviving breast cancer. By concentrating on this particular population and key physical functions, our review offers a detailed analysis of the clinical applicability, measurement properties, and limitations of these technologies. This focused approach addresses the gap identified by the previous work [47] regarding the need for more targeted research on the integration of technological tools into cancer rehabilitation practices. Moreover, our findings corroborate a previous review that recognizes force plates as effective tools for detecting subtle balance impairments [6] and Biodex systems as the gold standard for muscle strength assessment [48].

Conclusion

This scoping review synthesized current evidence on advanced clinic-based technologies used to assess and monitor 3 key physical functions, namely balance, muscle strength, and ROM, in individuals surviving breast cancer. The review found that balance assessments predominantly utilized force plates and SOT, muscle strength was assessed using various models of the Biodex isokinetic dynamometer, and range of motion was measured using digital inclinometers and MMC systems. These studies were conducted primarily in hospitals and specialized medical facilities, with participant samples spanning various treatment phases and survivorship stages. The review documented considerable heterogeneity in measurement approaches, clinical settings, and reporting practices across the included studies. These findings underline that a diverse toolkit is currently applied to measure physical function in breast cancer survivorship, but that variability in methods limits cross-study comparability. Mapping this landscape can help us prioritize targeted validation and implementation studies and inform the development of pragmatic guidance for selecting feasible, clinically useful technologies in rehabilitation practice.

Acknowledgments

The authors acknowledge the support provided by the Mathematics of Information Technology and Complex Systems (MITACS Accelerate Program) and the OKAKI Health Intelligence team.

We thank Liz Dennett, Public Service Librarian at the University of Alberta, for her expert assistance in developing and refining the database search strategies. Text editing and language polishing were assisted using ChatGPT by OpenAI. All content was reviewed and approved by the authors.

Data Availability

All data supporting the findings of this scoping review are included within the article and its supplementary materials. As this study synthesizes information from previously published literature, no new primary data were generated.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Complete search strategy and related keywords.
[DOCX File , 26 KB - cancer_v12i1e77894_app1.docx]

Multimedia Appendix 2
PRISMA-ScR (referred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews).
[DOCX File , 86 KB - cancer_v12i1e77894_app2.docx]

Multimedia Appendix 3
Responses to Reviewers comments.
[DOCX File , 94 KB - cancer_v12i1e77894_app3.docx]

References

1. Miller KD, Nogueira L, Devasia T, Mariotto AB, Yabroff KR, Jemal A, et al. Cancer treatment and survivorship statistics, 2022. CA Cancer J Clin 2022 Sep 23;72(5):409-436 [FREE Full text] [doi: 10.3322/caac.21731] [Medline: 35736631]

XSL•FO
RenderX

2. Devoogdt N, De Groef A. Physiotherapy management of breast cancer treatment-related sequelae. J Physiother 2024 Apr;70(2):90-105 [FREE Full text] [doi: 10.1016/j.jphys.2024.02.020] [Medline: 38519340]
3. Stout NL, Santa Mina D, Lyons KD, Robb K, Silver JK. A systematic review of rehabilitation and exercise recommendations in oncology guidelines. CA Cancer J Clin 2021 Mar 27;71(2):149-175 [FREE Full text] [doi: 10.3322/caac.21639] [Medline: 33107982]
4. Schapira L, Duffy CM. Cancer treatment and survivorship statistics, 2025: An urgent call to optimize health after cancer. CA Cancer J Clin 2025 May 30;75(4):277-279 [FREE Full text] [doi: 10.3322/caac.70017] [Medline: 40445135]
5. Painter P, Stewart AL, Carey S. Physical functioning: definitions, measurement, and expectations. Adv Ren Replace Ther 1999 Apr;6(2):110-123. [doi: 10.1016/s1073-4449(99)70028-2] [Medline: 10230878]
6. Hsieh KL, Wood TA, An R, Trinh L, Sosnoff JJ. Gait and balance impairments in breast cancer survivors: a systematic review and meta-analysis of observational studies. Arch Rehabil Res Clin Transl 2019 Jun;1(1-2):100001 [FREE Full text] [doi: 10.1016/j.arrct.2018.12.001] [Medline: 33570512]
7. Rasmussen GHF, Madeleine P, Arroyo-Morales M, Voigt M, Kristiansen M. Pain sensitivity and shoulder function among breast cancer survivors compared to matched controls: a case-control study. J Cancer Surviv 2023 Feb 26;17(1):150-159 [FREE Full text] [doi: 10.1007/s11764-021-00995-y] [Medline: 33495914]
8. Klassen O, Schmidt ME, Ulrich CM, Schneeweiss A, Potthoff K, Steindorf K, et al. Muscle strength in breast cancer patients receiving different treatment regimes. J Cachexia Sarcopenia Muscle 2017 Apr 28;8(2):305-316 [FREE Full text] [doi: 10.1002/jcsm.12165] [Medline: 27896952]
9. Cheville AL, Beck LA, Petersen TL, Marks RS, Gamble GL. The detection and treatment of cancer-related functional problems in an outpatient setting. Support Care Cancer 2008 May 17;17(1):61-67. [doi: 10.1007/s00520-008-0461-x] [Medline: 18478275]
10. Morishita S, Kasahara R, Yamamoto Y, Jinbo R, Takano A, Yasuda M, et al. Differences in the Relationships Between Muscle Strength, Muscle Mass, Balance Function, and Quality of Life for Middle-Aged and Older Breast Cancer Survivors. Integr Cancer Ther 2022 Dec 13;21:15347354221138574 [FREE Full text] [doi: 10.1177/15347354221138574] [Medline: 36511322]
11. Campos e Silva AC, Bergmann A, Araujo C, Montenegro AK, Tenorio ADS, Dantas D. Association of Handgrip Strength with Quality of Life in Breast Cancer Survivors: A Systematic Review and Meta-Analysis. Asian Pac J Cancer Prev 2022 Oct 01;23(10):3237-3245 [FREE Full text] [doi: 10.31557/apjcp.2022.23.10.3237] [Medline: 36308344]
12. Ness KK, Wall MM, Oakes JM, Robison LL, Gurney JG. Physical performance limitations and participation restrictions among cancer survivors: a population-based study. Ann Epidemiol 2006 Mar;16(3):197-205. [doi: 10.1016/j.annepidem.2005.01.009] [Medline: 16137893]
13. Huang M, Hile E, Croarkin E, Wampler-Kuhn M, Blackwood J, Colon G, et al. Academy of Oncologic Physical Therapy EDGE Task Force: A systematic review of measures of balance in adult cancer survivors. Rehabil Oncol 2019 Jul 01;37(3):92-103 [FREE Full text] [doi: 10.1097/01.reo.0000000000000177]
14. Croarkin E, Zampieri C. On the EDGE of Task Force recommendations: computerized balance assessment. Rehabil Oncol 2021 Jan;39(1):64-67 [FREE Full text] [doi: 10.1097/01.reo.0000000000000246] [Medline: 38404645]
15. Fisher MI, Lee J, Davies CC, Geyer H, Colon G, Pfalzer L. Oncology Section EDGE Task Force on breast cancer outcomes: a systematic review of outcome measures for functional mobility. Rehabilitation Oncology 2015;33(3):19-31. [doi: 10.1097/01893697-201533030-00004]
16. Sibley K, Straus S, Inness E, Salbach N, Jaglal S. Balance assessment practices and use of standardized balance measures among Ontario physical therapists. Phys Ther 2011 Nov;91(11):1583-1591 [FREE Full text] [doi: 10.2522/ptj.20110063] [Medline: 21868613]
17. Sibley KM, Straus SE, Inness EL, Salbach NM, Jaglal SB. Clinical balance assessment: perceptions of commonly-used standardized measures and current practices among physiotherapists in Ontario, Canada. Implement Sci 2013 Mar 20;8(1):33 [FREE Full text] [doi: 10.1186/1748-5908-8-33] [Medline: 23510277]
18. Samuel SR, Gandhi AR, Kumar KV, Saxena PP. Pedometer-based exercise interventions for patients with breast cancer receiving chemotherapy - a systematic review. Indian J Palliat Care 2020;26(1):105-109 [FREE Full text] [doi: 10.4103/IJPC.IJPC_175_19] [Medline: 32132793]
19. Blount DS, McDonough DJ, Gao Z. Effect of wearable technology-based physical activity interventions on breast cancer survivors' physiological, cognitive, and emotional outcomes: a systematic review. J Clin Med 2021 May 08;10(9):2015 [FREE Full text] [doi: 10.3390/jcm10092015] [Medline: 34066752]
20. Pan M, Wu K, Zhao J, Hou X, Chen P, Wang B. Effects of wearable physical activity tracking for breast cancer survivors: A systematic review and meta-analysis. Int J Nurs Knowl 2024 Apr 26;35(2):117-129. [doi: 10.1111/2047-3095.12418] [Medline: 36843066]
21. Lam WWT, Tang YM, Fong KNK. A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation. J Neuroeng Rehabil 2023 May 02;20(1):57 [FREE Full text] [doi: 10.1186/s12984-023-01186-9] [Medline: 37131238]

XSL•FO
RenderX

22. Gresham G, Hendifar AE, Spiegel B, Neeman E, Tuli R, Rimel BJ, et al. Wearable activity monitors to assess performance status and predict clinical outcomes in advanced cancer patients. NPJ Digit Med 2018 Jul 5;1(1):27 [FREE Full text] [doi: 10.1038/s41746-018-0032-6] [Medline: 31304309]
23. Hidding JT, Beurskens CHG, van der Wees PJ, van Laarhoven HWM, Nijhuis-van der Sanden MWG. Treatment related impairments in arm and shoulder in patients with breast cancer: a systematic review. PLoS One 2014 May 9;9(5):e96748 [FREE Full text] [doi: 10.1371/journal.pone.0096748] [Medline: 24816774]
24. Roberto M, Barchiesi G, Resuli B, Verrico M, Speranza I, Cristofani L, et al. Sarcopenia in Breast Cancer Patients: A Systematic Review and Meta-Analysis. Cancers (Basel) 2024 Jan 31;16(3):596 [FREE Full text] [doi: 10.3390/cancers16030596] [Medline: 38339347]
25. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
26. Bertoli J, Bezerra EDS, Winters-Stone KM, Alberto Gobbo L, Freitas IF. Mat Pilates improves lower and upper body strength and flexibility in breast cancer survivors undergoing hormone therapy: a randomized controlled trial (HAPiMat study). Disabil Rehabil 2023 Feb 02;45(3):494-503. [doi: 10.1080/09638288.2022.2032410] [Medline: 35107399]
27. Rao MS, Pattanshetty RB. Effect of myofascial release, stretching, and strengthening on upper torso posture, spinal curvatures, range of motion, strength, shoulder pain and disability, and quality of life in breast cancer survivors. Physiother Res Int 2022 Apr 19;27(2):e1939. [doi: 10.1002/pri.1939] [Medline: 35044712]
28. Højvig JH, Henriksen M, Bartholdy CR, Bonde CT. Donor-site morbidity following breast reconstruction with a latissimus dorsi flap - A prospective study. J Plast Reconstr Aesthet Surg 2022 Jul;75(7):2205-2210 [FREE Full text] [doi: 10.1016/j.bjps.2022.01.036] [Medline: 35183466]
29. Zabit Özdemir F, İyigün G. Is there a difference in balance functions between breast cancer survivor women and healthy women? Turk J Physiother Rehabil 2022 Dec 20;33(3):203-209. [doi: 10.21653/tjpr.1024868]
30. Wechsler S, Kneiss J, Adams B, Wood Magee LJ. Persistent cancer-related fatigue after breast cancer treatment predicts postural sway and post-exertional changes in sit-to-stand strategy. Rehabil Oncol 2022 Oct;40(4):162-171 [FREE Full text] [doi: 10.1097/01.reo.0000000000000308] [Medline: 36212795]
31. Artese A, Hunt R, Ormsbee M, Kim J, Arjmandi B, Panton L. Effect of functional impact training on body composition, bone mineral density, and strength in breast cancer survivors. Med Sci Sports Exerc 2021 Jan 01;53(1):90-101. [doi: 10.1249/mss.0000000000002438] [Medline: 32694366]
32. Evans ES, Ketcham CJ, Hibberd JC, Cullen ME, Basiliere JG, Murphy DL. Examination of clinical and laboratory measures of static and dynamic balance in breast cancer survivors. Physiother Theory Pract 2021 Nov 19;37(11):1199-1209. [doi: 10.1080/09593985.2019.1692391] [Medline: 31742463]
33. Uhm KE, Lee S, Kurillo G, Han JJ, Yang J, Yoo YB, et al. Usefulness of Kinect sensor-based reachable workspace system for assessing upper extremity dysfunction in breast cancer patients. Support Care Cancer 2020 Feb 29;28(2):779-786. [doi: 10.1007/s00520-019-04874-2] [Medline: 31144172]
34. Ribeiro IL, Camargo PR, Alburquerque-Sendín F, Ferrari AV, Arrais CL, Salvini TF. Three-dimensional scapular kinematics, shoulder outcome measures and quality of life following treatment for breast cancer - A case control study. Musculoskelet Sci Pract 2019 Apr;40:72-79. [doi: 10.1016/j.msksp.2019.01.012] [Medline: 30711913]
35. Monfort SM, Pan X, Patrick R, Ramaswamy B, Wesolowski R, Naughton MJ, et al. Gait, balance, and patient-reported outcomes during taxane-based chemotherapy in early-stage breast cancer patients. Breast Cancer Res Treat 2017 Jul;164(1):69-77 [FREE Full text] [doi: 10.1007/s10549-017-4230-8] [Medline: 28374323]
36. De Groef A, Van Kampen M, Verlvoesem N, Dieltjens E, Vos L, De Vrieze T, et al. Effect of myofascial techniques for treatment of upper limb dysfunctions in breast cancer survivors: randomized controlled trial. Support Care Cancer 2017 Jul 14;25(7):2119-2127. [doi: 10.1007/s00520-017-3616-9] [Medline: 28197849]
37. De Groef A, Van Kampen M, Tieto E, Schönweger P, Christiaens M, Neven P, et al. Arm lymphoedema and upper limb impairments in sentinel node-negative breast cancer patients: A one year follow-up study. Breast 2016 Oct;29:102-108. [doi: 10.1016/j.breast.2016.07.021] [Medline: 27479040]
38. Moreira R, Magalhães A, Oliveira H. A Kinect-based system for upper-body function assessment in breast cancer patients. J Imaging 2015 Nov 05;1(1):134-155. [doi: 10.3390/jimaging1010134]
39. Gritsenko V, Dailey E, Kyle N, Taylor M, Whittacre S, Swisher AK. Feasibility of using low-cost motion capture for automated screening of shoulder motion limitation after breast cancer surgery. PLoS One 2015 Jun 15;10(6):e0128809 [FREE Full text] [doi: 10.1371/journal.pone.0128809] [Medline: 26076031]
40. Winters-Stone KM, Torgrimson B, Horak F, Eisner A, Nail L, Leo MC, et al. Identifying factors associated with falls in postmenopausal breast cancer survivors: a multi-disciplinary approach. Arch Phys Med Rehabil 2011 Apr;92(4):646-652 [FREE Full text] [doi: 10.1016/j.apmr.2010.10.039] [Medline: 21367394]
41. Harrington S, Padua D, Battaglini C, Michener LA, Giuliani C, Myers J, et al. Comparison of shoulder flexibility, strength, and function between breast cancer survivors and healthy participants. J Cancer Surviv 2011 Jun 12;5(2):167-174. [doi: 10.1007/s11764-010-0168-0] [Medline: 21225372]

XSL•FO
RenderX

42. Waltman NL, Twiss JJ, Ott CD, Gross GJ, Lindsey AM, Moore TE, et al. Testing an intervention for preventing osteoporosis in postmenopausal breast cancer survivors. J Nurs Scholarsh 2003 Apr 23;35(4):333-338. [doi: 10.1111/j.1547-5069.2003.00333.x] [Medline: 14735675]
43. Piirtola M, Era P. Force platform measurements as predictors of falls among older people - a review. Gerontology 2006 Jan 27;52(1):1-16. [doi: 10.1159/000089820] [Medline: 16439819]
44. LeBlanc M, Stineman M, DeMichele A, Stricker C, Mao JJ. Validation of QuickDASH outcome measure in breast cancer survivors for upper extremity disability. Arch Phys Med Rehabil 2014 Mar;95(3):493-498 [FREE Full text] [doi: 10.1016/j.apmr.2013.09.016] [Medline: 24095658]
45. Tuominen J, Leppänen M, Jarske H, Pasanen K, Vasankari T, Parkkari J. Test-retest reliability of isokinetic ankle, knee and hip strength in physically active adults using Biodex System 4 Pro. Methods Protoc 2023 Mar 09;6(2) [FREE Full text] [doi: 10.3390/mps6020026] [Medline: 36961046]
46. Nugent EP, Snodgrass SJ, Callister R. The effect of velocity and familiarisation on the reproducibility of isokinetic dynamometry. Isokinet Exerc Sci 2015 Aug 01;23(3):205-214. [doi: 10.3233/ies-150582]
47. Amarelo A, Mota M, Amarelo B, Ferreira MC, Fernandes CS. Technological resources for physical rehabilitation in cancer patients undergoing chemotherapy: a scoping review. Cancers (Basel) 2024 Nov 25;16(23):3949 [FREE Full text] [doi: 10.3390/cancers16233949] [Medline: 39682137]
48. Drouin JM, Valovich-mcLeod TC, Shultz SJ, Gansneder BM, Perrin DH. Reliability and validity of the Biodex system 3 pro isokinetic dynamometer velocity, torque and position measurements. Eur J Appl Physiol 2004 Jan 1;91(1):22-29. [doi: 10.1007/s00421-003-0933-0] [Medline: 14508689]

Abbreviations

AP: anterior-posterior
APTA: American Physical Therapy Association
CoP: center of pressure
EDGE: Evaluation Database to Guide Effectiveness
HHD: handheld dynamometer
ICF: International Classification of Functioning, Disability and Health
MeSH: Medical Subject Headings
ML: medial-lateral
MMC: markerless motion capture
mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
ROM: range of motion
RMS: root mean square
SOT: Sensory Organization Test

Edited by N Cahill; submitted 21.May.2025; peer-reviewed by W Bloch, M Benton; comments to author 21.Oct.2025; accepted 16.Dec.2025; published 14.Jan.2026.

Please cite as:
Azhdar M, Loyola-Sanchez A, Wardrop A, Ferguson-Pell M
Advanced Clinical-Based Technologies for Monitoring Physical Function in Breast Cancer Survivors: Scoping Review
JMIR Cancer 2026;12:e77894
URL: https://cancer.jmir.org/2026/1/e77894
doi:10.2196/77894
PMID:

©Mahtab Azhdar, Adalberto Loyola-Sanchez, Amber Wardrop, Martin Ferguson-Pell. Originally published in JMIR Cancer (https://cancer.jmir.org), 14.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Measurements and Digital Technology Solutions to Monitor Physical Activity in Patients With Pediatric Cancer: Scoping Review

Greta Franceska Jermolenko[1*], Mg.Sc.Sal.; Guna Semjonova[2,3], PhD; Aija Klavina[1,4,5*], PhD; Evita Dubinina[2*], Mg.Sc.Sal.; Keita Augstkalne[2*], BA; Klavs Balamovskis-Kalnins[2*], BA; Alina Cesuna[6*]; Emil Syundyukov[2*], M.Sc. Comp.; Dace Bertule[3,7*], PhD; Madara Blumberga[7,8*], Mg.sc.sal.; Ilze Kundzina[9*], MDes

[1]Laboratory of Sports and Nutrition Research, Riga Stradins University, Riga, Latvia

[2]Latvian Academy of Sport Education, Riga Stradiņš University, Riga, Latvia

[3]Department of Rehabilitation, Rīga Stradiņš University, Riga, Latvia

[4]Department of Health Promotion and Rehabilitation, Lithuanian Sports University, Riga, Latvia

[5]Klaipėdos Valstybinė Kolegija/ Higher Education Institution, Klaipeda, Lithuania

[6]Faculty of Medicine and Life Sciences, University of Latvia, Riga, Latvia

[7]Children's Clinical University Hospital, Riga, Latvia

[8]Medical Education Technology Center, Rīga Stradiņš University, Riga, Latvia

[9]Joint study program Service Design Strategies and Innovation, Art Academy of Latvia, Riga, Latvia

[*]these authors contributed equally

Corresponding Author:
Guna Semjonova, PhD
Department of Rehabilitation
Rīga Stradiņš University
Dzirciema street 16
Riga, LV-1007
Latvia
Phone: 1 371 67 409 219
Email: guna.semjonova@rsu.lv

Abstract

Background: Patients with pediatric cancer often experience reduced physical activity (PA) due to treatment-related fatigue, functional limitations, and lack of structured exercise programs. Digital health solutions, including wearable sensors and augmented reality (AR)-based interventions, may offer new possibilities for monitoring and improving PA in this population.

Objective: This scoping review aims to address existing research gaps by identifying the instruments—both conventional and digital—used to monitor PA in patients with pediatric cancer during treatment. In addition, this study examines PA monitoring methods, identifies the variables collected, and explores the applicability of digital health solutions in facilitating PA engagement among patients with pediatric cancer.

Methods: In accordance with the Joanna Briggs Institute methodology, a systematic search was conducted across 8 scientific databases—ProQuest, Web of Science, EBSCO Complete, Google Scholar, ScienceDirect, Scopus, MEDLINE (PubMed), and Cochrane—on April 18 and 19, 2024. Studies were screened using the Rayyan AI-assisted review tool based on predefined inclusion criteria targeting children aged 7-19 years who were undergoing cancer treatment or were within 2 years posttreatment. Eligible studies included clinical trials and observational studies that examined objective (eg, wearable sensors) and subjective (eg, questionnaires and self-reports) approaches to PA monitoring. Keywords and controlled vocabulary (eg, MeSH [Medical Subject Headings] terms) were identified through a review of relevant literature. Data were extracted systematically to capture study characteristics, intervention types, and outcome measures. Extracted data were charted and synthesized narratively to identify patterns, technological applications, and research gaps in PA monitoring among patients with pediatric cancer.

Results: Twelve studies met the inclusion criteria and employed a range of PA monitoring tools. Digital health solutions, including Actical and Garmin VivoFit 3 devices, were used in 5 studies to assess step counts, gait cycles, and movement intensity. Self-reported measures were identified in 11 studies, most commonly the Activities Scale for Kids and the Pediatric Quality of

XSL•FO
RenderX

Life Inventory-Multidimensional Fatigue Scale, which provided insights into mobility and fatigue. Despite their feasibility, subjective assessments were limited by recall bias and motivational factors. Although digital health solutions—such as wearable sensors, gamification, and mobile applications—showed potential to improve PA adherence, their application remains underutilized, and evidence regarding their integration in pediatric oncology is limited.

Conclusions: Existing objective and subjective methods for monitoring PA provide valuable insights; however, gaps remain in the use of interactive digital health solutions, such as AR-based interventions, for PA monitoring and engagement. Future research should focus on integrating digital tools that not only track PA but also actively engage patients, enhance motivation, and support rehabilitation across both clinical and home settings.

(JMIR Cancer 2026;12:e73889) doi:10.2196/73889

KEYWORDS

pediatric oncology; cancer; physical activity monitoring; augmented reality; digital health solutions; gamification

Introduction

Every year, approximately 400,000 children and adolescents aged 0-19 are diagnosed with cancer worldwide [1]. Advances in childhood cancer treatment have significantly improved survival rates, with more than 80% of patients receiving modern cancer therapy surviving at least 5 years beyond diagnosis, and many achieving full remission [2,3]. However, despite these improvements in medical care, childhood cancer survivors remain at risk of recurrence, secondary malignancies, chronic conditions, and functional impairments [4]. These late effects of cancer and its treatment contribute to early mortality among survivors, making ongoing health management a critical concern [5]. Physical activity (PA) is increasingly recognized as a key component in mitigating the adverse effects of cancer and its treatment [6]. Nevertheless, published data indicate that fewer than 50% of patients with pediatric cancer meet PA guidelines, highlighting a serious challenge in maintaining adequate activity levels [7]. The World Health Organization (WHO) defines PA as any bodily movement produced by skeletal muscles that requires energy expenditure, including activities performed during leisure time, transportation, and daily routines—not just structured exercise [1]. Cancer-related fatigue is a major barrier to PA, contributing to sedentary behavior and reduced mobility [8]. In addition, some patients with pediatric cancer experience physical limitations that affect their ability to perform daily activities, as observed in conditions such as clear cell sarcoma, which often affects the legs, feet, arms, hands, and torso [9]. Given these challenges, implementing interventions to counteract cancer-related physical side effects is a major priority in pediatric oncology. There is growing evidence that PA is not only safe during both the acute treatment phase and survivorship [10] but also effective in reducing cancer treatment–related side effects [11,12]. Encouraging PA in patients with pediatric cancer is therefore essential; however, existing approaches face significant barriers. Currently, hospital-based PA interventions are the most common approach, led by physiotherapists, occupational therapists, and nurses who encourage children to remain active during therapy sessions [13]. However, pediatric oncology rehabilitation services are often underutilized, with few children referred for specialized rehabilitation. Medical professionals have identified a shortage of dedicated pediatric oncology rehabilitation services, while rehabilitation specialists have highlighted the lack of appropriate environments and equipment [14]. These limitations exacerbate challenges related to organizing PA, particularly for patients who experience fatigue and mobility restrictions [7]. To address these gaps, digital technology is emerging as a valuable complement to conventional PA interventions. Technologies such as wearable sensors, augmented reality (AR), and the Internet of Things offer new approaches to engaging patients with pediatric cancer in PA and monitoring their progress. For instance, wearable devices can track step counts, gait cycles, and movement patterns in real time, providing immediate feedback on PA performance [15]. Mobile health (mHealth) apps that incorporate gamification techniques have proven effective in increasing PA by offering interactive challenges and motivational feedback, thereby helping to sustain engagement among patients with pediatric cancer [16,17]. In addition, virtual reality (VR)– and AR-based interactive digital health solutions can create immersive exercise environments, helping to overcome the lack of dedicated health care—particularly rehabilitation—spaces, while providing personalized and engaging PA experiences [18]. These technologies offer scalable, flexible, and patient-centered solutions that could enhance PA participation and adherence in pediatric oncology care [19,20]. Understanding the specific individual needs of patients with pediatric cancer is crucial for developing effective PA interventions that align with their physical capabilities and treatment constraints. This scoping review aimed to address existing research gaps by identifying conventional and digital health instruments used to monitor PA in patients with pediatric cancer during treatment. Furthermore, it examined PA monitoring methods, characterized the variables collected, and explored the potential of digital health solutions to facilitate PA engagement in patients with pediatric cancer.

The research questions guiding this review are as follows:

- What conventional and digital methods are used to monitor PA in children and adolescents with cancer?
- What variables collected by conventional and digital health instruments are used to monitor PA in children and adolescents with cancer?
- What is the applicability of different conventional and digital health instruments for monitoring PA levels in children and adolescents with cancer?
- How are digital health solutions and PA monitoring tools incorporated into interventions aimed at improving PA in children and adolescents with cancer?

XSL•FO
RenderX

In addressing these questions, the authors use the term "variables" specifically to refer to objective measurements, such as step count and gait cycles/minute, distinguishing them from subjective values obtained through self-reported questionnaires. By mapping the available PA monitoring tools and interventions, this review aims to contribute to a more comprehensive understanding of how digital solutions can support PA promotion in patients with pediatric cancer.

Methods

Protocol and Registration

This scoping review was conducted between April and September 2024. The review followed the Joanna Briggs Institute (JBI) methodology for scoping reviews [21,22] and was reported in accordance with PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews; Multimedia Appendix 1) guidelines [23]. The 6-stage framework proposed by Arksey and O'Malley [24] was applied. The protocol was registered with the Center for Open Science [25], and the review was conducted without deviations from the registered protocol.

Eligibility Criteria

Overview

Based on recommendations from the JBI [21,22] and the research questions formulated to refine the focus of the scoping review and develop an effective search strategy, we defined the eligibility criteria according to the Population, Concept, and Context framework [22].

Population: Children and Adolescents With Cancer

- Children and adolescents with cancer during treatment or up to 2 years after treatment. Based on existing studies, this 2-year period—including active treatment and early recovery—is used because it allows researchers to assess how PA supports the restoration of cardiorespiratory fitness and muscle strength while helping to alleviate lingering posttreatment fatigue and physical limitations [26].
- Studies including children and adolescents aged 7-19 years. This age range was selected because it corresponds to the pediatric cancer population. Younger children are generally unable to complete questionnaires or engage in PA independently, while older age groups are typically classified as adults with cancer. Children and adolescents in this age range can effectively engage in self-directed activities and structured exercise programs. Moreover, older adolescents face unique physical and psychological challenges during the transition to adulthood, making their inclusion essential for understanding the impact of PA during and after cancer treatment [6,13].

Concepts: Digital Health Solutions Such as Wearable Sensors, Gamification, and Mobile Apps

- Wearable sensors, AR, and gamification provide new approaches to engaging patients with pediatric cancer in PA while enabling continuous monitoring of their progress [16].
- The incorporation of AR and gamification in pediatric cancer care is justified by their demonstrated effectiveness in enhancing PA levels among children and adolescents [17,18,20].
- Digital health solutions enable objective, data-driven monitoring of PA patterns, allowing health care professionals to tailor interventions and evaluate treatment-related functional outcomes [16,20].

Context: Physical Activity

- PA, as defined by the WHO, encompasses bodily movement that enables individuals to perform daily activities effectively and engage in exercise [13].
- The term "physical activity" includes all forms of bodily movement performed in educational, recreational, home, or community settings, including but not limited to aerobic exercise, resistance training, flexibility, endurance activities, and stretching routines.
- Key parameters for evaluating PA include frequency, intensity, duration, type of activity, and distance covered. In pediatric cancer, assessing a child's engagement in daily PA can provide valuable insights into recovery and health outcomes after treatment.

Selection of Records of Evidence

Individual studies published between 2000 and March 2024 were included. The starting year of 2000 was selected to capture research conducted during the period when digital technologies began to be systematically introduced into health care and rehabilitation, marking the global transition toward digital health strategies and culminating in the WHO's Global Strategy on Digital Health 2020-2025 [27]. The selection of records for this review was guided by predefined inclusion and exclusion criteria to ensure a comprehensive and relevant assessment of digital technologies for monitoring PA in patients with pediatric cancer. Inclusion and exclusion criteria were defined according to the Population, Concept, and Context framework (Textbox 1).

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

1. Inclusion criteria

- Population: children and adolescents diagnosed with cancer
- Concept: digital health solutions related to physical activity (eg, augmented reality, gamification, digital tools, digital environments)
- Context: physical activity outcomes
- Setting: supervised clinical environment
- Study type: original studies with any design or data type (quantitative and qualitative)
- Publication status: published in a peer-reviewed journal
- Publication language: English
- Full-text available
- Included keywords: cancer, pediatric, oncology, exercise, monitoring, physical activity, fitness, movement, digital technologies, augmented reality, gamification, digital tools, digital environment.

2. Exclusion criteria

- Population: children and adolescents more than 2 years after cancer treatment
- The study does not include patients aged 7-19
- Concept: digital technologies not related to physical activity (eg, for health management)
- Context: nonphysical activity outcomes or monitoring; specific physical fitness measures
- Setting: nonclinical
- Study type: other study types (eg, protocols, narrative reviews, or systematic reviews)
- Publication status: published without peer review, dissertations, books, conference papers, letters, or editorials
- Publication language: written in a language other than English
- Full-text not available
- Excluded keywords: drug, in vitro, animal, mice, mouse, animals, bacteria, murine, rat, fish, canine, rodents, transgenic, rodent, piglets, rabbits
- Mental health outcomes are excluded unless they are mentioned in combination with a physical activity outcome.
- Duplicate data: if studies provide overlapping datasets or are part of the same project without additional insights, they should be excluded to avoid repetition.

Information Sources and Search Strategy

A comprehensive search was conducted in ProQuest, Web of Science, EBSCO Complete, Google Scholar, ScienceDirect, Scopus, MEDLINE (PubMed), and Cochrane on April 18-19, 2024. The search strategy (Multimedia Appendix 2) combined keywords and controlled vocabulary (eg, MeSH [Medical Subject Headings] terms) related to pediatric cancer, digital technologies, PA, and outcome measurement, using Boolean operators. The final search strategy was adapted for each database. Only articles published in English were included, and gray literature was excluded. Following the database searches, all records were imported into the Rayyan AI-assisted review tool, after which the study selection process was conducted in 3 phases. First, titles and abstracts were screened, followed by full-text review to identify articles relevant to the research questions of this scoping review. In the final phase, a data extraction template was created within the Rayyan AI-assisted review tool to facilitate the systematic collection of key information from the selected articles. Given the anticipated volume of articles, all 11 coauthors participated in the selection process at each stage, with each article independently reviewed by 2 team members (AK and DB). If the Rayyan AI-assisted review tool identified a conflict regarding inclusion or exclusion, a third reviewer (GS) resolved it. Data extraction was performed independently and cross-checked to enhance accuracy. In addition, the team met regularly online to address any issues arising during the article selection phases.

Data Charting Process

Extracted data from the reviewed articles were categorized and analyzed using a structured set of variables. Key elements of the analysis included the paper ID, first author (name and surname), journal name, and article title, providing a bibliographic foundation for the review. To contextualize the research geographically and demographically, the country of the study, study aim, and participant information (number, sex, age, diagnosis, comorbidities, treatment status, and setting) were documented. In addition, details regarding participants' diagnosis and treatment were recorded to further clarify the clinical context. The methodology of each study was assessed, including study type (eg, cross-sectional, longitudinal), applied methods (PA monitoring instruments, questionnaires, proxy reports, digital technologies), specific tools used for monitoring, and study outcomes (objectives, data type, inclusion or empowerment, intervention effectiveness, and potential effects

XSL•FO
RenderX

of monitoring methods on activity levels). In this review, the treatment period was defined as the active treatment phase, extending up to 2 years posttreatment.

Data Synthesis and Analyses

Following the final stages of the 6-stage framework for scoping reviews proposed by Arksey and O'Malley [24], data synthesis was conducted after identifying, selecting, and charting the relevant studies. Extracted information was first organized in a structured Excel (Microsoft Corporation) matrix, enabling systematic comparison across studies. The synthesis process involved 2 main phases:

- Descriptive numerical summary, where studies were grouped by publication year, country, cancer type, participant age, treatment phase, study design, and type of PA intervention or monitoring technology. Frequency counts and distributions were used to illustrate research trends and methodological characteristics.
- Qualitative thematic synthesis, where data related to intervention characteristics, PA measurement methods, feasibility, and outcomes were coded inductively by 2 reviewers (AK and DB). Codes were iteratively compared and refined through discussion among all coauthors until consensus was reached. Emerging categories were then grouped into overarching themes, including (1) feasibility and adherence of PA monitoring, (2) barriers and facilitators to engagement, and (3) identification of digital and exercise-based interventions.

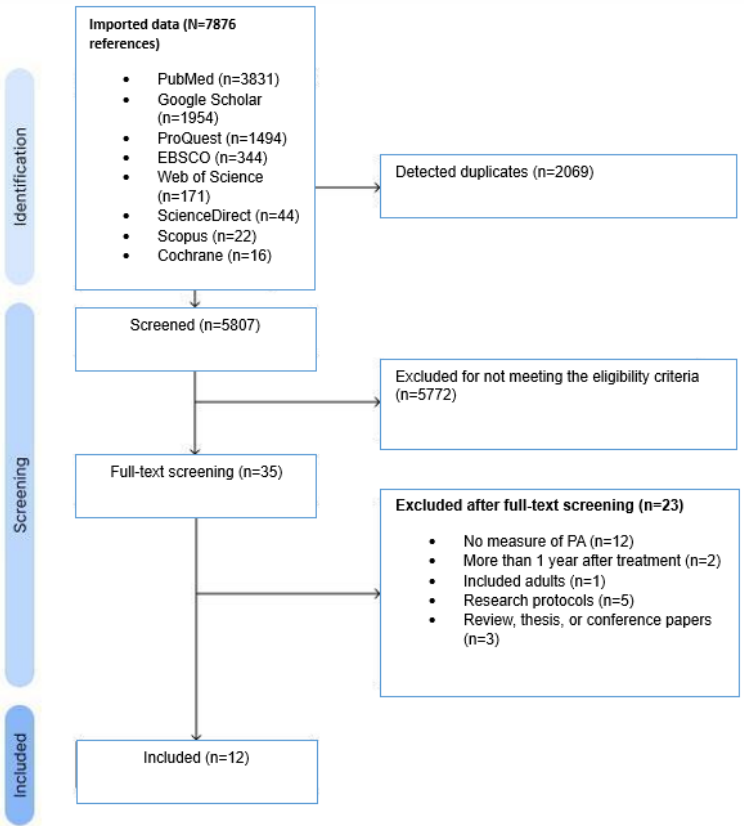
This structure enabled the systematic identification of recurring patterns, differences, and knowledge gaps across studies. Results were synthesized narratively and organized in tables to highlight key themes, relationships, and interconnections among the examined variables.

Results

Selection of Records of Evidence

Our search identified a total of 7876 records across 8 electronic databases. Of these, 2069 duplicates were removed, leaving 5807 records eligible for screening, which indicates the feasibility of our search strategy. Following the initial screening based on eligibility criteria, 5772 records were excluded, and 35 were sought for retrieval. Of these 35 records, 23 were excluded because they did not include a measure of PA, the research was conducted more than 2 years after treatment, the sample included adults, or the records were research protocols, reviews, theses, or conference papers. A total of 12 publications met all eligibility criteria and were included in this scoping review. The source and selection of evidence are presented in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the literature search and study selection process.



Characteristics of Included Studies

The characteristics of the 12 selected studies included in this review are summarized in Table 1. The studies encompassed various research designs, including 4 (33%) randomized controlled trials [28-31], 1 (8%) quasi-randomized trial [32], 4 (33%) cross-sectional studies [33-36], 1 (8%) multicenter cohort study [37], 1 (8%) pilot study [15], and 1 (8%) prospective

XSL•FO
RenderX

observational study [38]. The included studies investigated PA levels, motor performance, and quality of life among children and adolescents diagnosed with cancer, both during and after treatment. Studies were conducted in Germany [16,27,33,37], the Netherlands [15,29,30,33], the United States [37,38], Hong Kong [36], and South Korea [32]. The interventions employed objective PA monitoring tools, such as accelerometers (Actical, Move 3, Step Watch 3, Garmin vívofit 3) and subjective self-report measures, including the Activities Scale for Kids (ASK), Pediatric Quality of Life Inventory (PedsQL)-Multidimensional Fatigue Scale, and other health-related quality-of-life instruments. Across the 12 studies, a total of 1083 participants (ages 3-20 years) were assessed, with sample sizes ranging from 25 to 482. Most participants had leukemia (acute lymphoblastic leukemia or acute myeloid leukemia) or lymphoma (non-Hodgkin lymphoma or Hodgkin), while others had solid tumors (eg, bone tumors, brain tumors, neuro-oncology cases). Some studies included mixed cancer diagnoses without specific classification. Two studies examined participants posttreatment [32,35], whereas the remaining 10 focused on patients undergoing active treatment or within the first year after completion. For example, Kang et al [32] investigated a mobile game–based healthy lifestyle program for childhood cancer survivors in South Korea, while Stössel et al [35] compared PA behaviors before, during, and after cancer treatment in Germany.

XSL•FO
RenderX

Table 1. Characteristics of included studies for analysis.

Author	Country	Participants' age and setting	Disease	Research design	Objective instruments used to assess and monitor physical activity	Self-reported instruments used to assess and monitor physical activity
Gaser et al [28]	Germany	• 41 participants aged 4-18 years • During the treatment	• Acute lymphoblastic leukemia (n=25, 61%) • Non-Hodgkin lymphoma (n=12, 29%) • Acute myeloid leukemia (n=4, 10%)	Randomized controlled trial	• The accelerometer Move 3 • The Functional Activity of Daily Living Screen with everyday tasks • The Motor Performance in Paediatric Oncology test	• The Self-Reported Activities Scale for Kids
Gaser et al [31]	Germany	• 41 participants aged 4-18 years • During the treatment	• Acute lymphoblastic leukemia (n=25) • Acute myeloid leukemia (n=4) • Non-Hodgkin lymphoma (n=12) • Second primary cancer (n=2)	Randomized controlled trial	• The accelerometer Move 3 • The Functional Activity of Daily Living Screen with everyday tasks • The Motor Performance in Paediatric Oncology test	• The Self-Reported Activities Scale for Kids
Braam et al [33]	The Netherlands	• 60 participants aged 8-18 years • Treated with chemotherapy, radiotherapy, or both during or within the first year after cancer treatment	• Acute lymphoblastic leukemia (n=17) • Acute myeloid leukemia (n=8) • Brain tumor (n=8) • Hodgkin lymphoma (n=7) • Bone tumor (n=7) • Non-Hodgkin lymphoma (n=5) • Rhabdomyosarcoma (n=3) • Chronic myeloid leukemia (n=2) • Others (n=3)	Cross-sectional study	• Actical activity monitor	• PedsQL-MFS[a] • The Participation in Sports Before the Cancer Diagnosis Questionnaire; a subscale of the Self Perception Profile Questionnaire for children aged 8-11 years and for adolescents aged 12-18 years • Children's Depression Inventory

XSL•FO
RenderX

Author	Country	Participants' age and setting	Disease	Research design	Objective instruments used to assess and monitor physical activity	Self-reported instruments used to assess and monitor physical activity
Götte et al [34]	Germany	• 28 participants with a mean age of 13.8 (SD 2.8) years • During cancer treatment	• Leukemia (n=13) • Acute lymphoblastic leukemia (n=9) • Acute myeloid leukemia (n=4) • Bone tumor (n=9) • Ewing sarcoma (n=3) • Osteosarcoma (n=6) • Localized at the lower limb (n=3) • Localized at trunk/upper limb (n=6) • Lymphoma (n=2) • Other solid tumor (n=4)	Cross-sectional study design	• Step Watch 3	• Physical activity questionnaire from the German Health Interview and Examination Survey for Children and Adolescents of the Robert Koch Institute
Withycombe et al [38]	The United States	• 65 participants aged 8-17 years • Enrollment generally occurred during the first 6 months of cancer therapy, but at least 4 weeks after diagnosis and at least 3+ weeks after cancer definitive surgery (if applicable)	• Leukemia/lymphoma (n=38) • Solid tumor (n=16); • Neuro-oncology (n=11)	Cross-sectional study design	• The Garmin vívofit 3	• A 9-question ecological survey • PROMIS[b]
Rehorst-Kleinlugtenbelt et al [15]	The Netherlands	• 25 participants aged 3.1-17 years • Undergoing active cancer treatment	• Hematological malignancy (n=17) • Solid tumors (n=8)	Cross-sectional study design	• Accelerometry	• N/A[c]
Mack et al [37]	The United States		• First diagnosis of cancer of any type	Multicenter cohort study	• N/A	

XSL•FO
RenderX

Author	Country	Participants' age and setting	Disease	Research design	Objective instruments used to assess and monitor physical activity	Self-reported instruments used to assess and monitor physical activity
		• 482 caregivers • Patients were aged 7-18 years. Participants received upfront cancer treatment, including chemotherapy and radiotherapy.				• Child self-reports and caregiver proxy reports were collected for PROMIS pediatric domains, including mobility (physical functioning), pain interference, fatigue, depressive symptoms, anxiety, and psychological stress.
van Dijk-Lokkart et al [30]	The Netherlands	• 68 participants aged 7-18 years • Patients were still receiving treatment or were within the first year after cessation of treatment	• Diagnosed with any type of childhood cancer	Randomized controlled trial	• Actical activity monitor	• Cancer-related fatigue was assessed using both child self-report and parent proxy report versions of the PedsQL-MFS
Lam et al [36]	Hong Kong	• 76 participants aged 9-18 years • After treatment	• Leukemia (n=32, 42.1%), lymphoma (n=7, 9.2%), brain and spinal tumor (n=16, 21.1%), bone tumor (n=10, 13.2%), and others (n=11, 14.5%)	A cross-sectional study	• N/A	• The Chinese University of Hong Kong: Physical Activity Rating for Children and Youth • Physical Activity Self-Efficacy Questionnaire • The PedsQL cancer module version 3.0 • PedsQL
Kang et al [32]	South Korea	• 51 participants aged 6-13 years • All participants were childhood cancer survivors whose treatment was terminated at least 12 months prior	• Childhood cancer survivors	Quasi-randomized trial	• N/A	• The Child Healthy Lifestyle Profile adapted for children, parents, or guardians filling out the profile
Stössel et al [35]	Germany		• Diagnosed with any type of cancer	Cross-sectional, multicenter study	• N/A	• Physical activity questionnaire, which is in parts based on the German Health Interview and Examination Survey for Children and Adolescents (KiGGS)

XSL•FO
RenderX

Author	Country	Participants' age and setting	Disease	Research design	Objective instruments used to assess and monitor physical activity	Self-reported instruments used to assess and monitor physical activity
		• 114 patients with cancer and 37 healthy controls between 4 and 20 years of age • Completed intensive cancer treatment				
Braam et al [29]	Germany	• 66 participants aged 8-18 • During treatment or within 12 months after treatment	• N/A	Randomized controlled trial	• Cardiorespiratory fitness assessed by VO_{2peak}[d] (ml kg–1 min–1) using indirect calorimetry • Muscular strength was assessed using a hand-held dynamometer	• Overall fatigue was assessed using the child self-report version of the PedsQL-MFS (acute version) • General health-related quality of life was assessed using the Dutch self-report version of the PedsQL • Behavioral problems were assessed using the Youth Self-Report • Athletic competence and global self-worth were assessed using the corresponding subscales of the Self-Perception Profile • Depressive symptoms were assessed using the Children's Depression Inventory

[a]PedsQL-MFS: Pediatric Quality of Life Inventory-Multidimensional Fatigue Scale.

[b]PROMIS: Patient-Reported Outcome Measurement Information System.

[c]N/A: not applicable.

[d]VO_{2peak}: peak oxygen uptake.

PA Monitoring Methods and Obtained Variables in Patients With Pediatric Cancer

Both objective and self-reported methods were used to assess PA levels, motor performance, and fatigue in children and adolescents with cancer. Across the included studies, these 2 approaches complemented each other, providing quantitative and subjective perspectives on PA behavior and its relationship to treatment-related fatigue and recovery. Objective PA monitoring was employed in 5 studies [28,30,33,34,38], using accelerometers such as Actical, Move 3, Step Watch 3, and Garmin vívofit 3. van Dijk-Lokkart et al [30] used Actical accelerometers with a required wear time of at least 500 minutes/day to measure PA levels in children undergoing or recently completing cancer treatment. Similarly, Gaser et al [28] utilized the Move 3 accelerometer alongside motor performance tests to examine PA in patients with leukemia and lymphoma. Withycombe et al [38] used the Garmin vívofit 3 over a 4-day period to track step counts and correlate them with symptom reports. This illustrates that objective PA monitoring focused primarily on step counts and time spent in different intensity zones (light, moderate, and vigorous), reflecting a general shift toward using wearable accelerometers as a feasible and noninvasive method for tracking daily movement patterns during and after cancer treatment. In addition, subjective self-reported PA assessment methods were applied in 7 studies [28,30,33-37]. For example, the ASK, which includes 30 items across 7 subdomains (personal care, dressing, other skills, locomotion, play, standing up, and movement), and the PedsQL-Multidimensional Fatigue Scale were used. Gaser et al [28] combined ASK with motor performance testing to evaluate functional abilities such as mobility and locomotion, balance and coordination, strength and endurance, as well as fatigue and energy levels. Mack et al [37] employed caregiver-reported measures to assess mobility, pain interference, and fatigue, incorporating surveys that evaluated children's ability to perform daily tasks, their experience of pain, and levels of physical exhaustion. Götte et al [34] assessed

XSL•FO
RenderX

fatigue using the PedsQL-Multidimensional Fatigue Scale, which captures general, sleep-related, and cognitive fatigue. Additional self-report instruments, such as the Child Health Utility 9D used by Stössel et al [35], were applied to assess health-related quality of life, while structured questionnaires like the Pediatric Outcomes Data Collection Instrument, used by Braam et al [33], provided insights into mobility and participation in PAs. These self-reported instruments captured complementary domains—mobility, participation, fatigue, and quality of life—allowing researchers to contextualize accelerometer-derived data with patient-perceived outcomes and daily functional capacity. The integration of objective and self-reported measures provided a comprehensive understanding of patients' functional status across different stages of cancer care. For comparison, while some studies focused on the active treatment phase, others, such as Kang et al [32] and Stössel et al [35], specifically examined PA behaviors in posttreatment survivors. Kang et al [32] investigated the effects of a mobile game–based healthy lifestyle program on PA levels, sedentary behavior, and overall quality of life in childhood cancer survivors in South Korea. Stössel et al [35] analyzed PA behaviors before, during, and after cancer treatment, focusing on changes in daily activity levels, mobility patterns, and engagement in sports or recreational activities. Both studies provided valuable insights into long-term functional outcomes, highlighting the impact of cancer treatment on sustained PA and overall well-being in survivors. This comparison between active treatment and posttreatment groups revealed common patterns: patients typically exhibited reduced PA levels during therapy, with gradual improvement during recovery, although many survivors continued to experience limitations in endurance and persistent fatigue. In the included studies, PA monitoring primarily focused on assessing overall activity levels using accelerometers to calculate step counts and estimate time spent in light, moderate, and vigorous activity. None of the studies conducted detailed activity recognition beyond basic actions such as walking, running, or sitting, nor did they employ multisensor systems for body motion analysis or fatigue detection. More sophisticated approaches—such as recognizing complex daily activities, detecting nuanced movement patterns, or assessing early signs of fatigue—were absent, highlighting an important gap for future research.

Applicability of PA Monitoring Instruments in Pediatric Oncology

Overall, the studies included in this review demonstrated that PA monitoring instruments were feasible for use with patients with pediatric cancer. Feasibility was largely determined by patient adherence and device usability, both of which were reported as satisfactory in most studies. Withycombe et al [38] utilized the Garmin vívofit 3, worn on the wrist for 4 consecutive days, to measure step counts and activity patterns, observing that wearable technology effectively tracked PA, although results varied according to treatment stage and location (eg, fewer steps during hospitalization). van Dijk-Lokkart et al [30] reported good compliance with accelerometer-based PA monitoring using Actical accelerometers worn on the hip for at least 500 minutes/day over a 1-week period. Gaser et al [28] employed the Move 3 accelerometer, also positioned on the hip, with a minimum required wear time of 5 days. Some studies imposed restrictions on PA measurement, such as excluding nonambulatory patients or requiring a minimum number of valid wear days for data inclusion [34,37]. While feasibility and compliance rates were generally acceptable, the studies also highlighted important contextual and individual barriers affecting the accuracy and consistency of PA monitoring in this population. Mack et al [37] found that fatigue, treatment side effects, and hospitalization influenced step counts and overall compliance. Braam et al [29] identified psychological factors, including lack of motivation and fear of overexertion, as key barriers to PA engagement. Lam et al [36] cited time constraints and limited access to PA resources as additional challenges for participants. Despite these barriers, authors suggested that self-monitoring through wearable devices served as a motivational tool, encouraging participants to maintain PA levels. These findings suggest that while device-based PA monitoring is technically and behaviorally feasible, its effectiveness depends on addressing both physical and psychosocial barriers that influence patient participation during cancer treatment and recovery.

Interventions for Improving PA in Pediatric Oncology

PA interventions varied across the 12 selected studies, incorporating individualized exercise programs, supervised training, and digital health tools to enhance and monitor PA levels. Intervention durations ranged from 4 weeks to 12 months, depending on study design and patient condition. Most studies reported retention rates above 80% and adherence rates ranging from 70% to 95%, indicating strong engagement with the interventions [28-30,32]. For example, Braam et al [29] reported a 92% adherence rate to supervised PA sessions, which included aerobic training, resistance exercises, and motor performance activities 2-3 times/week at moderate-to-high intensity. Similarly, Kang et al [32] reported a 90% program completion rate and high user engagement metrics (eg, daily log-ins and activity tracking) with a mobile game–based lifestyle program for childhood cancer survivors, encouraging daily movement. van Dijk-Lokkart et al [30] implemented individually tailored programs with varying intensity and duration based on the child's physical condition, reporting over 85% adherence in most participants. Gaser et al [28] also employed individualized exercise protocols for patients with leukemia and lymphoma, focusing on moderate-intensity exercises, and reported a 100% retention rate and 88% session completion. Across these interventions, personalization of exercise type, frequency, and duration emerged as a key determinant of both adherence and safety, reinforcing the importance of individualized approaches in pediatric oncology rehabilitation.

Table 2 (also see Multimedia Appendix 3) summarizes the PA monitoring methods, variables, applicability, and interventions across the studies in pediatric oncology, providing detailed information on the tools and strategies used to assess and promote PA in this population. A summary addressing each research question, accompanied by easily interpretable tables, is presented in Multimedia Appendix 4, offering an overview of how the selected studies address the research questions.

XSL•FO
RenderX

Table 2. Summary of PA[a] monitoring methods, variables, applicability, and interventions in pediatric oncology.

Reference	1. What methods are used for PA monitoring in patients with pediatric cancer?	2. What variables are collected to monitor PA in patients with pediatric cancer?	3. What is the applicability of different instruments to facilitate the PA level (monitoring) in patients with pediatric cancer?	4. What interventions are used to improve PA in patients with pediatric cancer research?
Gaser et al [28]	• The Move 3 accelerometer (movisens GmbH) • Self-reported questionnaire	• Step count, amplitude of moderate-to-vigorous PA, body position, and wear time	• Participants wore the device on the right hip during the daytime and removed the device during nighttime sleep. • Records of ≥4 days of ≥8 hours/day of wear time were included. • The younger participants (n=5, ages 4-7 years) felt more disturbed by the sensor on the hip. As a result, all of them refused the measurement. Because of the accelerometer's algorithm, the PA could be calculated only for those participants aged over 7 years.	N/A[b]
Gaser et al [31]	• The Move 3 accelerometer (movisens GmbH) • Self-reported questionnaire	• Step count	• Participants wore the device on the right hip during the daytime and removed the device during nighttime sleep. • Records of ≥4 days of ≥8 hours/day of wear time were included. • Reasons for invalid measurements were the lack of compliance and unscheduled inpatient hospitalizations.	Exercise program–specific strength training combined with a standard care exercise program (2-3 exercise sessions/week).
Braam et al [33]	• Actical accelerometer (B series; Philips Respironics Actical MiniMitter) • Self-reported questionnaires	• Counts/minute; the acceleration signal is summed over a specific time interval (epoch). A 15-second epoch was used in the study.	• The activity monitor was attached to an elastic waist belt and worn on the left hip during daytime at waking hours (between 6:00 AM and 11:59 PM) for 4 consecutive days (Wednesday-Saturday). The device was removed while bathing. • When the device was worn for less than 500 minutes/day, the measurement was considered invalid. • The memory capacity of the accelerometer did not allow assessment of PA by 15-second epoch for 7 days; therefore, 4 days were used. Missing data on 3 days within the measurement week remain a limitation.	N/A
Götte et al [34]	• Step Watch 3 sealed uniaxial Activity Monitor (Orthocare Innovations) • Self-reported questionnaire	• The volume of activity/day (gross counts/day) and intensity of activity (gross counts/minute)	• The device was attached to the ankle with an elastic strap. • Participants wore the activity monitor for 7 consecutive days from the morning after waking up to bedtime. • Days with <8 hours of wear time were excluded. • The study concluded that objective measures should be preferentially used for the assessment of PA in children and adolescents with cancer to ensure accurate and reliable data. Self-reports can complement objective measures by capturing activities or sports that are not reflected in step counts, as well as individuals' expectations and attitudes toward exercise.	N/A

XSL•FO
RenderX

Reference	1. What methods are used for PA monitoring in patients with pediatric cancer?	2. What variables are collected to monitor PA in patients with pediatric cancer?	3. What is the applicability of different instruments to facilitate the PA level (monitoring) in patients with pediatric cancer?	4. What interventions are used to improve PA in patients with pediatric cancer research?
Withycombe et al [38]	• The Garmin vívofit 3 accelerometer • Self-reported questionnaire	• Step count	• Participants wore an accelerometer for 7 days. Data were included if available for at least 4 days during a defined 7-day period. • Eligible days included a minimum of 10 hours of wear time between 6 AM and 10 PM. • Step monitoring may serve as an objective indicator for overall symptom count, fatigue, PA, and physical function.	N/A
Rehorst-Kleinlugtenbelt et al [15]	• The Actical (Philips Respironics, Mini Mitter Co, Inc)	• Step count • Counts/minute • 15-second epoch	• The device was fastened to an elastic waist belt strap and worn on the right hip. A minimum wearing time of 8 hours/day was required, with a minimum of 4 valid days a week. • Parents or participants maintained a "wearing time" activity diary. • The study found that accelerometry is suitable for the objective assessment of PA in children with childhood cancer during their treatment. The data gave a presentation of their PA behavior during the day. Accelerometers provide an objective assessment of PA and can be used in different kinds of patients.	N/A
Mack et al [37]	• Self-reported questionnaire • Caregiver proxy report	• PROMIS[c] assessments of the child's physical function (mobility) and symptoms, including pain interference, fatigue, depressive symptoms, anxiety, and psychological stress; 5 response categories. Each question's recall period is the past 7 days.	• Our findings suggest that proxy reporting is influenced by the proxy's personal experience of symptoms and function as well as the child's experience. • Caregivers tended to overestimate symptoms and underestimate function relative to children themselves.	N/A
van Dijk-Lokkart et al [30]	• Actical accelerometer (B series; Philips Respironics Actical MiniMitter) • Self-reported questionnaire • Caregiver proxy report	• Counts/minute, 15-second epoch	• The accelerometer was worn on the hip during daytime at waking hours (between 6:00 AM and 11:59 PM) on 4 consecutive days (Wednesday-Saturday), at least 500 minutes/day over a 1-week period. • Although the PedsQL-MFS[d] has acceptable psychometric properties, including content validity, internal consistency, and responsiveness, there are inconsistent reports regarding known group validity in pediatric cancer.	Cardiorespiratory and muscle strength training twice a week for 12 weeks at a physical therapy sports center near the child's home.
Lam et al [36]	• Self-reported questionnaires		• N/A	N/A

XSL•FO
RenderX

Reference	1. What methods are used for PA monitoring in patients with pediatric cancer?	2. What variables are collected to monitor PA in patients with pediatric cancer?	3. What is the applicability of different instruments to facilitate the PA level (monitoring) in patients with pediatric cancer?	4. What interventions are used to improve PA in patients with pediatric cancer research?
		• The Chinese University of Hong Kong: Physical Activity Rating for Children and Youth—score ranges from no exercise at all (0) to vigorous exercise on most days (10). • PA self-efficacy score: self-confidence in PA participation, from "not sure," "a little sure," to "very sure." • The PedsQL cancer module version 3.0: How much of a problem was a task over the last month, from 0 to 4 (0=never, 1=almost never, 2=sometimes, 3=often, 4=almost always)?		
Kang et al [32]	• Self-reported questionnaire • Caregiver proxy report	• N/A	• The healthy lifestyle program based on a mobile serious game assessed the following subdimensions: health responsibility, PA, nutrition, positive life perspective, interpersonal relations, stress management, and spiritual health. • No significant effects were observed for any subdimension except PA.	A healthy lifestyle program based on a mobile serious game that promotes healthy behaviors through the completion of 26 quests, encompassing 7 subcomponents: nutrition, exercise, hygiene, interpersonal relationships, stress management, meaning of life, and health responsibility.
Stössel et al [35]	• Self-reported questionnaire	• Participants rated their overall level of PA on a visual analog scale, ranging from "not at all physically active" to "very physically active." They also reported their PA across different domains, including type of activity, minutes/day, and intensity level. For analysis, total PA was calculated in minutes/week and categorized by intensity as light, moderate, or vigorous.	• N/A	N/A
Braam et al [29]		• Mean counts/minute (15-second time interval)		

XSL•FO
RenderX

Reference	1. What methods are used for PA monitoring in patients with pediatric cancer?	2. What variables are collected to monitor PA in patients with pediatric cancer?	3. What is the applicability of different instruments to facilitate the PA level (monitoring) in patients with pediatric cancer?	4. What interventions are used to improve PA in patients with pediatric cancer research?
	• Actical accelerometer (B series; Philips Respironics Actical MiniMitter) • Self-reported questionnaires • Caregiver proxy report		• The accelerometer was attached to an elastic waist belt and worn on the left hip during waking hours (6:00 AM to 11:59 PM). Participants wore the device for 4 consecutive days (Wednesday-Saturday). After the monitoring period, the accelerometers were returned to the research team by postal mail. Compliance with accelerometer use was low during the final study measurement week. Nonuse was primarily attributed to discomfort associated with wearing the device on a hip-mounted belt, with complaints reported particularly by girls and children who were overweight.	The 12-week intervention comprised 24 individual physical exercise sessions, consisting of two 45-minute exercise sessions/week conducted at a local physical therapy practice, and one 60-minute psychosocial training session every 2 weeks delivered to the child at the treating pediatric oncology hospital.

[a]PA: physical activity.

[b]N/A: not applicable.

[c]PROMIS: Patient-Reported Outcome Measurement Information System.

[d]PedsQL-MFS: Pediatric Quality of Life Inventory-Multidimensional Fatigue Scale.

In addition to traditional exercise and supervised rehabilitation programs, several studies integrated digital components to enhance monitoring capabilities and sustain motivation beyond clinical settings. Digital health interventions, such as Kang et al's [32] mobile game–based program, wearable step-count monitoring [15,38], and other digital tracking tools, were used across multiple studies to assess PA levels and support motivation in patients with pediatric cancer. These programs were implemented in diverse settings, including hospitals [29,38], physical therapy centers [29,30], and home-based contexts [15,32], with supervision provided by physiotherapists, pediatric oncology specialists, or through digital applications. The inclusion of gamified and interactive elements was particularly effective in maintaining engagement among children and adolescents, consistent with evidence that motivation and enjoyment play critical roles in sustaining PA behavior during and after cancer treatment. Despite challenges such as fatigue and other treatment-related side effects, studies consistently reported that structured PA programs—including those incorporating digital elements—were associated with improved adherence and participation [30,32,38]. These digital and traditional interventions targeted a range of physical and psychosocial outcomes, with variations in duration and intensity influencing their overall effectiveness. PA intervention durations varied across studies examining effects on physical functioning [29,30,38], fitness [29,31], fatigue [30,37], quality of life [32,36], and psychological well-being [29,37]. For example, Kang et al [32] evaluated a mobile game–based intervention, which increased daily activity, improved health behaviors, and enhanced self-reported well-being. van Dijk-Lokkart et al [30] implemented a structured PA program that reduced fatigue and improved functional capacity over time. Rehorst-Kleinlugtenbelt et al [15] examined a structured physical exercise intervention delivered in both hospital and home settings, which led to increased PA levels and better adherence to exercise programs over time. By contrast, Braam et al [29] examined a structured exercise intervention delivered in a hospital and at a local physical therapy practice, which showed no significant beneficial effects on physical outcomes.

Discussion

Principal Findings

This scoping review identified various PA assessment instruments, including digital solutions, used to monitor PA in patients with pediatric cancer. Self-reported questionnaires remain the most frequently used instruments due to their accessibility and cost-effectiveness [28-30,32-39]. While these instruments provide valuable insights from the patient's perspective, their reliability is limited by recall bias and motivational factors, making them inherently subjective. Although widely utilized, self-reported questionnaires have been increasingly complemented by accelerometry, which provides empirical, quantifiable measurements of PA [38,39]. Self-reported surveys are often favored because they are inexpensive, easy to administer, and adaptable to different age groups, making them feasible even in larger study settings. Self-reported questionnaires also provide important patient-centered insights by capturing subjective experiences such as fatigue, pain interference, and broader quality-of-life outcomes (eg, Götte et al [34], Mack et al [37], Stössel et al [35], Braam et al [33]). However, their accuracy is limited by recall bias, motivational influences, and social desirability, which may distort reported activity levels. While self-reported surveys offer valuable perspectives on how children and adolescents perceive their PA and related symptoms, they are less reliable for accurately quantifying activity levels. This underscores the importance of complementing self-reported surveys with objective monitoring tools to achieve a more comprehensive assessment of PA in pediatric cancer.

XSL•FO
RenderX

Self-reported questionnaires remain widely used due to their low cost, ease of administration, and ability to capture subjective patient experiences that objective measures may not assess. These instruments are also feasible for large-scale studies and adaptable to different age groups. However, subjective methods are prone to bias, social desirability, and motivational influences, which can affect the accuracy of data outcomes. In summary, self-reported questionnaires demonstrate lower reliability in quantifying PA levels compared with objective instruments such as accelerometers.

Accelerometer-based approaches provided objective data such as step counts, wear-time compliance, and time spent at different activity intensity levels (eg, Actical in van Dijk-Lokkart et al [30], Garmin vívofit 3 in Withycombe et al [38], Move 3 in Gaser et al [28]). In some studies, accelerometry was combined with motor performance testing to assess balance, coordination, strength, and endurance [28]. Other studies extended the focus to participation in recreational or social activities, linking activity patterns to broader dimensions of daily life [33,34]. Overall, accelerometers appear to be practical and feasible tools for objectively monitoring PA in patients with pediatric cancer. Accelerometers provide reliable, quantifiable data and are generally well-tolerated by participants; however, their usefulness is limited by the need for compliance, inability to capture complex or context-specific activities, and the lack of standardized protocols across studies. When used alongside complementary instruments, accelerometers can be a valuable component of PA monitoring in both research and clinical practice. Despite growing interest in interventions using digital tools, this review found that only 1 study employed a mobile game–based approach for PA tracking in patients with pediatric cancer [32], which also functioned as a mobile engagement tool. In this study, children participated in a serious game requiring completion of 26 quests, including 7 subelements (nutrition, exercise, hygiene, interpersonal relationships, stress management, meaning of life, and health responsibility), to promote a healthy lifestyle [32]. This review demonstrated the potential of digital solutions to integrate PA monitoring with the promotion of healthy lifestyle habits during cancer treatment in clinical settings. The selected studies also highlighted the limited research outcomes available to date. While digital tools may be feasible and applicable, their broader implementation in pediatric oncology remains largely unexplored.

The integration of PA monitoring with patient engagement strategies for behavior change, such as gamification and AR, holds promise for enhancing both data accuracy and patient motivation [18,32]. This highlights a research gap where interactive digital solutions could play a pivotal role in optimizing PA interventions for patients with pediatric cancer. While prior studies have extensively examined the effectiveness of digital technology in promoting PA [15,31,32,38], only a few have explored how PA monitoring itself can facilitate PA participation and adherence [30,32,38]. Although most studies relied on accelerometers and questionnaires, a gap remains regarding the use of AR and gamification specifically for PA monitoring in pediatric oncology. Studies employed a variety of methods to promote PA in children and adolescents with cancer, including self-reported questionnaires, wearable accelerometers, and digital game–based tools. Self-reported questionnaires, such as the ASK used by Gaser et al [28] and the PedsQL-Multidimensional Fatigue Scale applied by Götte et al [34], Mack et al [37], and others [33,36], captured subjective experiences of mobility, fatigue, pain, and quality of life, providing valuable patient-centered insights, though they are limited by recall bias and motivational factors. Wearable devices, including Actical accelerometers in van Dijk-Lokkart et al [30], Move 3 in Gaser et al [28], and Garmin vívofit 3 in Withycombe et al [38], offered objective, continuous measures of step counts, activity intensity, and motor performance, enhancing accuracy and enabling longitudinal monitoring, but they may fail to capture complex or context-specific activities and depend on participant compliance. Digital interventions, such as Kang et al's [32] mobile game–based program, combined PA tracking with engagement and health education, supporting adherence and motivation; however, evidence remains limited, and these tools are largely unvalidated. Hybrid approaches integrating subjective and objective measures, as applied in several studies [32,38], show promise for providing a multidimensional understanding of PA and potentially improving adherence, though they require greater resources and supervision. Overall, these findings highlight the need to further explore innovative interventions that implement digital tools designed to be both engaging and user-friendly for patients with pediatric cancer. For instance, AR-based interactive games have been proposed as a promising approach, as they may enable real-time monitoring and provide personalized support for daily PA. Such technologies have the potential to offer health care professionals accurate information about a patient's daily activity level and enhance motivation to remain active during different treatment stages. Further research is required to evaluate the impact of digital tools on physical and mental health outcomes in patients with pediatric cancer, while also addressing the limitations of current methods highlighted in this review.

Implication for Practice

This review underscores the need for a standardized approach to PA monitoring in patients with pediatric cancer, as current practices are heterogeneous and often lack consistency. While innovative digital solutions, such as AR and gamification, may offer promising approaches to enhance motivation and adherence to PA interventions, their applicability and effectiveness in pediatric oncology remain largely unexplored and require further investigation.

Additionally, this review highlighted the need for a standardized approach to PA monitoring in patients with pediatric cancer. The diversity of accelerometers [28-34,38] and self-report questionnaires [28-38] underscores the necessity for unified guidelines to ensure consistency in data collection and interpretation. Developing standardized guidelines for PA assessment would enhance comparability across studies and support the integration of digital monitoring solutions into clinical practice.

Limitations

This review has several limitations. It included only articles published in English, which may have resulted in the exclusion of relevant research published in other languages. Additionally,

XSL•FO
RenderX

the focus on children and adolescents aged 7-19 years who were undergoing treatment or within 2 years posttreatment may have limited insights into long-term PA outcomes beyond this period. Furthermore, heterogeneity in study designs, PA measurement tools, and intervention methodologies posed challenges for direct comparisons. The aim of this review was not to evaluate the validity or reliability of the selected instruments but to map the available measurement methods and digital solutions used to monitor PA in patients with pediatric cancer. The limited adoption of interactive digital solutions highlights a critical research gap, underscoring the need for further studies to evaluate their feasibility, appropriateness, and long-term impact on PA engagement in this population.

Future Directions

Our findings indicate that future research should focus on exploring how PA monitoring can actively facilitate PA participation in patients with pediatric cancer. While numerous studies have examined PA monitoring tools, only a few have assessed their impact on promoting PA. Further investigation is also needed into the use of AR-based interventions and gamification for PA monitoring in pediatric cancer. Among the 12 included studies, none involved an AR-based intervention, and the only digital engagement tool identified was a gamified mobile game by Kang et al [32]. AR represents an unexplored but promising area for future research in pediatric oncology rehabilitation. Interactive digital interventions have the potential to both measure and enhance PA engagement by providing real-time feedback and motivation [18]. Evaluating patient acceptance, clinical integration, and long-term adherence to such technologies will be crucial to optimizing their effectiveness. Furthermore, the absence of standardized PA monitoring protocols highlights the need for unified guidelines for digital monitoring solutions tailored to patients with pediatric cancer. Establishing consistent assessment criteria would improve comparability across studies and facilitate the integration of digital monitoring solutions into routine clinical practice.

Conclusions

Despite the growing need to monitor and promote PA in patients with pediatric cancer, opportunities for interactive and engaging PA monitoring remain limited. The findings of this scoping review indicate an emerging body of literature on digital health technologies, including wearable sensors such as accelerometers, mHealth apps, and gamification, which are being explored for PA assessment and engagement. Future research should further investigate the purpose, scope, and integration of digital health technologies to facilitate interactive and personalized approaches for effectively monitoring and enhancing PA in both clinical and home-based settings for patients with pediatric cancer.

Acknowledgments

We extend our gratitude to the Riga Stradins University librarian and all collaborators involved for their support in making this research possible. The authors declare the use of generative artificial intelligence in the research and writing process. According to the GAIDeT (Generative AI Delegation Taxonomy) (2025), proofreading and editing were delegated to the generative artificial intelligence ChatGPT (OpenAI, GPT-5) under full human supervision. Responsibility for the final manuscript lies entirely with the authors. Generative artificial intelligence tools are not listed as authors and do not bear responsibility for the outcomes.

Funding

This project was funded by the Internal RSU and External RSU with the LSPA Consolidation Grant “Monitoring of physical activity habits of pediatric oncology patients using a digital approach with augmented reality-based interactive game elements” (RSU/LSPA-PA-2024/1-0002).

Authors' Contributions

Conceptualization: GS, AK, ED
Data curation: KA, KBK, ED, AC, GFJ
Formal analysis: AK, DB, GS, ES, KA, AC
Funding acquisition: GS, KA, ES
Investigation: MB, DB, AK, GS, KA, KBK, AC, GFJ, ED, IK
Methodology: DB, AK, GS, KA, KBK, AC, GFJ, ED, IK
Project administration: GS, KA
Resources: DB, AK, GS, KA, KBK, AC, GFJ, ED, IK
Supervision: GS, AK, DB
Validation: AK, GS, DB, ED
Visualization: KA, KBK
Writing – original draft: GFJ, KA, KBK, AC, IK, GS, ED
Writing – review & editing: GFJ, KA, KBK, AC, IK, GS, ED, AK, DB, MB

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
PRISMA-ScR checklist.
[DOCX File , 3892 KB - cancer_v12i1e73889_app1.docx]

Multimedia Appendix 2
Search strategy.
[DOCX File , 2313 KB - cancer_v12i1e73889_app2.docx]

Multimedia Appendix 3
Detailed summary of physical activity monitoring methods, variables, applicability, and interventions in pediatric oncology.
[DOCX File , 23 KB - cancer_v12i1e73889_app3.docx]

Multimedia Appendix 4
Research questions.
[PDF File (Adobe PDF File), 245 KB - cancer_v12i1e73889_app4.pdf]

References

1. World Health Organization (WHO). CureAll framework: WHO Global Initiative for Childhood Cancer: increasing access, advancing quality, saving lives. WHO. Geneva, Switzerland: World Health Organization; 2021. URL: https://iris.who.int/handle/10665/347370 [accessed 2025-10-13]
2. Armstrong GT, Liu Q, Yasui Y, Neglia JP, Leisenring W, Robison LL, et al. Late mortality among 5-year survivors of childhood cancer: a summary from the Childhood Cancer Survivor Study. J Clin Oncol 2009 May 10;27(14):2328-2338 [FREE Full text] [doi: 10.1200/JCO.2008.21.1425] [Medline: 19332714]
3. Landier W, Skinner R, Wallace WH, Hjorth L, Mulder RL, Wong FL, et al. Surveillance for late effects in childhood cancer survivors. J Clin Oncol 2018 Jul 20;36(21):2216-2222 [FREE Full text] [doi: 10.1200/JCO.2017.77.0180] [Medline: 29874139]
4. Fidler MM, Reulen RC, Winter DL, Kelly J, Jenkinson HC, Skinner R, British Childhood Cancer Survivor Study Steering Group. Long-term cause-specific mortality among 34,489 five-year survivors of childhood cancer in Great Britain: population-based cohort study. BMJ 2016 Sep 01;354:i4351 [FREE Full text] [doi: 10.1136/bmj.i4351] [Medline: 27586237]
5. Misiąg W, Piszczyk A, Szymańska-Chabowska A, Chabowski M. Physical activity and cancer care—a review. Cancers (Basel) 2022 Aug 27;14(17):14174154 [FREE Full text] [doi: 10.3390/cancers14174154] [Medline: 36077690]
6. Winter C, Müller C, Hoffmann C, Boos J, Rosenbaum D. Physical activity and childhood cancer. Pediatr Blood Cancer 2010 Apr;54(4):501-510. [doi: 10.1002/pbc.22271] [Medline: 19743298]
7. Larsen EH, Mellblom AV, Larsen MH, Ruud E, Thorsen L, Petersen NN, et al. Perceived barriers and facilitators to physical activity in childhood cancer survivors and their parents: a large-scale interview study from the International PACCS Study. Pediatr Blood Cancer 2023 Jan;70(1):e30056. [doi: 10.1002/pbc.30056] [Medline: 36251019]
8. Ness KK, Mertens AC, Hudson MM, Wall MM, Leisenring WM, Oeffinger KC, et al. Limitations on physical performance and daily activities among long-term survivors of childhood cancer. Ann Intern Med 2005 Nov 01;143(9):639-647 [FREE Full text] [doi: 10.7326/0003-4819-143-9-200511010-00007] [Medline: 16263886]
9. Gauß G, Beller R, Boos J, Däggelmann J, Stalf H, Wiskemann J, et al. Adverse events during supervised exercise interventions in pediatric oncology—a nationwide survey. Front Pediatr 2021;9:682496 [FREE Full text] [doi: 10.3389/fped.2021.682496] [Medline: 34490156]
10. Morales JS, Valenzuela PL, Rincón-Castanedo C, Takken T, Fiuza-Luces C, Santos-Lozano A, et al. Exercise training in childhood cancer: a systematic review and meta-analysis of randomized controlled trials. Cancer Treat Rev 2018 Nov;70:154-167. [doi: 10.1016/j.ctrv.2018.08.012] [Medline: 30218787]
11. Morales JS, Valenzuela PL, Herrera-Olivares AM, Baño-Rodrigo A, Castillo-García A, Rincón-Castanedo C, et al. Exercise interventions and cardiovascular health in childhood cancer: a meta-analysis. Int J Sports Med 2020 Mar;41(3):141-153. [doi: 10.1055/a-1073-8104] [Medline: 31935777]
12. World Health Organization (WHO). Physical activity: fact sheet. WHO. Geneva, Switzerland: WHO; 2024 Jun 26. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2025-10-13]
13. San Juan AF, Wolin K, Lucía A. Physical activity and pediatric cancer survivorship. Recent Results Cancer Res 2011;186:319-347. [doi: 10.1007/978-3-642-04231-7_14] [Medline: 21113771]
14. Geelen SJG, van Dijk-Huisman HC, de Bie RA, Veenhof C, Engelbert R, van der Schaaf M, et al. Barriers and enablers to physical activity in patients during hospital stay: a scoping review. Syst Rev 2021 Nov 04;10(1):293 [FREE Full text] [doi: 10.1186/s13643-021-01843-x] [Medline: 34736531]
15. Rehorst-Kleinlugtenbelt L, Bekkering W, van der Torre P, van der Net J, Takken T. Physical activity level objectively measured by accelerometery in children undergoing cancer treatment at home and in a hospital setting: a pilot study. Pediatric Hematology Oncology Journal 2019 Dec 04;4(4):82-88 [FREE Full text] [doi: 10.1016/j.phoj.2019.12.004]

XSL•FO
RenderX

16. Devine KA, Viola AS, Coups EJ, Wu YP. Digital health interventions for adolescent and young adult cancer survivors. JCO Clin Cancer Inform 2018 Dec;2:1-15 [FREE Full text] [doi: 10.1200/CCI.17.00138] [Medline: 30652583]
17. Sween J, Wallington SF, Sheppard V, Taylor T, Llanos AA, Adams-Campbell LL. The role of exergaming in improving physical activity: a review. J Phys Act Health 2014 May;11(4):864-870 [FREE Full text] [doi: 10.1123/jpah.2011-0425] [Medline: 25078529]
18. Rizzo AA, Kim GJ. A SWOT analysis of the field of virtual rehabilitation and therapy. Presence (Camb) 2005;14(2):119-146. [doi: 10.1162/1054746053967094]
19. Domin A, Spruijt-Metz D, Theisen D, Ouzzahra Y, Vögele C. Smartphone-based interventions for physical activity promotion: scoping review of the evidence over the last 10 years. JMIR Mhealth Uhealth 2021 Jul 21;9(7):e24308 [FREE Full text] [doi: 10.2196/24308] [Medline: 34287209]
20. Lee AM, Chavez S, Bian J, Thompson LA, Gurka MJ, Williamson VG, et al. Efficacy and effectiveness of mobile health technologies for facilitating physical activity in adolescents: scoping review. JMIR Mhealth Uhealth 2019 Feb 12;7(2):e11847. [doi: 10.2196/11847] [Medline: 30747716]
21. Peters MDJ, Godfrey C, McInerney P, Khalil H, Larsen P, Marnie C, et al. Best practice guidance and reporting items for the development of scoping review protocols. JBI Evid Synth 2022 Apr 01;20(4):953-968. [doi: 10.11124/JBIES-21-00242] [Medline: 35102103]
22. Peters MDJ, Marnie C, Tricco AC. Scoping reviews. In: JBI Manual for Evidence Synthesis. Adelaide, SA, Australia: JBI; 2024.
23. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
24. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb 02;8(1):19-32. [doi: 10.1080/1364557032000119616]
25. Jermoļenko GF, Augstkalne K, Semjonova G. Measurements and digital technology solutions to monitor physical activity in pediatric oncology patients: a scoping review protocol. Open Science Framework (OSF). 2025 Feb 10. URL: https://10.17605/OSF.IO/ZW48R [accessed 2026-01-08]
26. Hudson MM, Bhatia S, Casillas J, Landier W, Section on Hematology/Oncology, Children's Oncology Group, American Society of Pediatric Hematology/Oncology. Long-term follow-up care for childhood, adolescent, and young adult cancer survivors. Pediatrics 2021 Sep;148(3):e2021053127 [FREE Full text] [doi: 10.1542/peds.2021-053127] [Medline: 34462344]
27. World Health Organization (WHO). Global strategy on digital health 2020–2025. WHO. Geneva, Switzerland: WHO; 2021. URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2026-01-08]
28. Gaser D, Peters C, Götte M, Oberhoffer-Fritz R, Feuchtinger T, Schmid I, et al. Analysis of self-reported activities of daily living, motor performance and physical activity among children and adolescents with cancer: Baseline data from a randomised controlled trial assessed shortly after diagnosis of leukaemia or non-Hodgkin lymphoma. Eur J Cancer Care (Engl) 2022 Mar;31(2):e13559. [doi: 10.1111/ecc.13559] [Medline: 35150025]
29. Braam KI, van Dijk-Lokkart EM, Kaspers GJL, Takken T, Huisman J, Buffart LM, et al. Effects of a combined physical and psychosocial training for children with cancer: a randomized controlled trial. BMC Cancer 2018 Dec 27;18(1):1289 [FREE Full text] [doi: 10.1186/s12885-018-5181-0] [Medline: 30587148]
30. Van Dijk-Lokkart EM, Steur LMH, Braam KI, Veening MA, Huisman J, Takken T, et al. Longitudinal development of cancer-related fatigue and physical activity in childhood cancer patients. Pediatr Blood Cancer 2019 Dec;66(12):e27949. [doi: 10.1002/pbc.27949] [Medline: 31436372]
31. Gaser D, Peters C, Oberhoffer-Fritz R, Götte M, Feuchtinger T, Schmid I, et al. Effects of strength exercise interventions on activities of daily living, motor performance, and physical activity in children and adolescents with leukemia or non-Hodgkin lymphoma: results from the randomized controlled ActiveADL study. Front Pediatr 2022;10:982996 [FREE Full text] [doi: 10.3389/fped.2022.982996] [Medline: 36425395]
32. Kang KA, Kim HH, Kim SJ, Song IH, Lee MJ, Lee SY, et al. Effectiveness of a healthy lifestyle program based on a mobile serious game for childhood cancer survivors: a quasi-randomized trial. J Pediatr Nurs 2024;77:35-44. [doi: 10.1016/j.pedn.2024.03.006] [Medline: 38479061]
33. Braam KI, van Dijk-Lokkart EM, Kaspers GJL, Takken T, Huisman J, Bierings MB, et al. Cardiorespiratory fitness and physical activity in children with cancer. Support Care Cancer 2016 May;24(5):2259-2268 [FREE Full text] [doi: 10.1007/s00520-015-2993-1] [Medline: 26581899]
34. Götte M, Seidel CC, Kesting SV, Rosenbaum D, Boos J. Objectively measured versus self-reported physical activity in children and adolescents with cancer. PLoS One 2017;12(2):e0172216 [FREE Full text] [doi: 10.1371/journal.pone.0172216] [Medline: 28207820]
35. Stössel S, Neu M, Oschwald V, Söntgerath R, Däggelmann J, Eckert K, et al. Physical activity behaviour in children and adolescents before, during and after cancer treatment. Sport Sci Health 2019 Dec 18;16(2):347-353. [doi: 10.1007/s11332-019-00612-7]

XSL•FO
RenderX

36. Lam KKW, Li WHC, Chiu SY, Chan GCF. The impact of cancer and its treatment on physical activity levels and quality of life among young Hong Kong Chinese cancer patients. Eur J Oncol Nurs 2016 Apr;21:83-89. [doi: 10.1016/j.ejon.2016.01.007] [Medline: 26952682]
37. Mack JW, McFatrich M, Withycombe JS, Maurer SH, Jacobs SS, Lin L, et al. Agreement between child self-report and caregiver-proxy report for symptoms and functioning of children undergoing cancer treatment. JAMA Pediatr 2020 Nov 01;174(11):e202861 [FREE Full text] [doi: 10.1001/jamapediatrics.2020.2861] [Medline: 32832975]
38. Withycombe JS, McFatrich M, Hinds PS, Bennett A, Lin L, Maurer SH, et al. Can steps per day reflect symptoms in children and adolescents undergoing cancer treatment? Cancer Nurs 2022;45(5):345-353 [FREE Full text] [doi: 10.1097/NCC.0000000000001062] [Medline: 35131975]
39. Au WW, Recchia F, Fong DY, Wong SHS, Chan DKC, Capio CM, et al. Effect of wearable activity trackers on physical activity in children and adolescents: a systematic review and meta-analysis. Lancet Digit Health 2024 Sep;6(9):e625-e639 [FREE Full text] [doi: 10.1016/S2589-7500(24)00139-0] [Medline: 39112110]

Abbreviations

AR: augmented reality
ASK: Activities Scale for Kids
JBI: Joanna Briggs Institute
MeSH: Medical Subject Headings
mHealth: mobile health
PA: physical activity
PedsQL: Pediatric Quality of Life Inventory
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
VR: virtual reality
WHO: World Health Organization

Edited by N Cahill; submitted 13.Mar.2025; peer-reviewed by A Pouple, Y Xiao; comments to author 24.Jul.2025; revised version received 04.Nov.2025; accepted 17.Nov.2025; published 29.Jan.2026.

Please cite as:
Jermolenko GF, Semjonova G, Klavina A, Dubinina E, Augstkalne K, Balamovskis-Kalnins K, Cesuna A, Syundyukov E, Bertule D, Blumberga M, Kundzina I
Measurements and Digital Technology Solutions to Monitor Physical Activity in Patients With Pediatric Cancer: Scoping Review
JMIR Cancer 2026;12:e73889
URL: https://cancer.jmir.org/2026/1/e73889
doi:10.2196/73889
PMID:

©Greta Franceska Jermolenko, Guna Semjonova, Aija Klavina, Evita Dubinina, Keita Augstkalne, Klavs Balamovskis-Kalnins, Alina Cesuna, Emil Syundyukov, Dace Bertule, Madara Blumberga, Ilze Kundzina. Originally published in JMIR Cancer (https://cancer.jmir.org), 29.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Internet of Things–Enhanced Mathematical Oncology: Conceptual Framework for Adaptive Cancer Care Modeling

Adeniyi Onasanya[1], MSc, MBA; Ramona Kyabaggu[2], MSc, MHIS; Daryl Hepting[1], PhD

[1]Department of Computer Science, University of Regina, 3737 Wascana Parkway, Regina, SK, Canada

[2]Johnson Shoyama Graduate School of Public Policy, University of Regina, Regina, SK, Canada

Corresponding Author:
Adeniyi Onasanya, MSc, MBA
Department of Computer Science, University of Regina, 3737 Wascana Parkway, Regina, SK, Canada

Abstract

The Internet of Things (IoT) is transforming various industries, including health care. IoT-based systems are increasingly prevalent in consumer health applications, while intelligent or smart devices equipped with sophisticated sensors are gaining recognition for their potential to improve clinical care practice and decision-making. Cancer care is a particularly promising area for IoT applications, enabling real-time and personalized interventions. However, empirical research on the effects of IoT in this field is limited due to the complexities inherent in cancer as a dynamic disease and the paucity of IoT-generated data available for research. This presents an opportunity to apply mathematical modeling to understand the effects of IoT under various scenarios. These analytical and "in silico" mathematical approaches are instrumental with limited data. Such models support the analysis of treatment uncertainty and patient response while balancing patient preferences, clinical outcomes, and system-level constraints. Grounded in mathematical oncology and health informatics, this paper proposes a conceptual framework that integrates real-time IoT data as dynamic inputs into adaptive mathematical models to simulate cancer dynamics. By exploring applications across multiple levels of analysis, the study demonstrates how IoT-enhanced mathematical models could inform implementation and optimize oncology services, addressing a critical gap in current research.

(JMIR Cancer 2026;12:e73997) doi:10.2196/73997

KEYWORDS

business process modeling; cancer care services; health care; health care informatics; health care systems; internet of things; IoT-based health care; mathematical modeling; mathematical oncology

Introduction and Motivation

The Internet of Things (IoT) represents a paradigm shift in technological advancement, enabling interconnected, intelligent, and responsive systems through continuous data collection, communication, and analysis. As a key driver of Industry 4.0, IoT facilitates the convergence of information technology and operational technology by interconnecting physical objects in internet-aware environments. Its rapid adoption across industries, governments, and society underscores its value in enabling interoperable, data-driven services beyond traditional machine-to-machine communication [1].

In health care, numerous IoT applications have been proposed to support care delivery, access, and patient safety [2-9]. However, despite growing interest, few studies empirically evaluate or mathematically model the effects of IoT interventions, particularly in cancer care. While mathematical oncology and IoT-based cancer care systems have mainly evolved in parallel, research that unifies these approaches remains limited. Integrating IoT with mathematical oncology offers an opportunity to link mechanistic disease dynamics with real-time, patient-specific data, enabling more adaptive and patient-centered cancer care.

The primary goal of this work is to investigate how integrating IoT technologies with mathematical oncology can enhance the modeling, analysis, and optimization of cancer care through real-time, data-driven, and adaptive approaches to improve treatment efficacy and patient outcomes. This study specifically builds on an established melanoma immunotherapy model [10] to develop a conceptual framework that integrates real-time IoT data into adaptive mathematical models of cancer dynamics and behavior. This framework extends treatment strategies and supports business process modeling by embedding IoT-enabled devices within business process modeling elements—tasks, events, activities, and resources—allowing continuous monitoring and evaluation of cancer conditions and characteristics in IoT-enabled care systems.

The paper is organized as follows. The *Overview and Literature* section reviews cancer, common treatment strategies, and existing mathematical models, highlighting their limitations. The *Case Study* section on immunotherapy for melanoma presents a conceptual framework for adaptive cancer care modeling using real-time IoT data as dynamic parameter inputs.

XSL•FO
RenderX

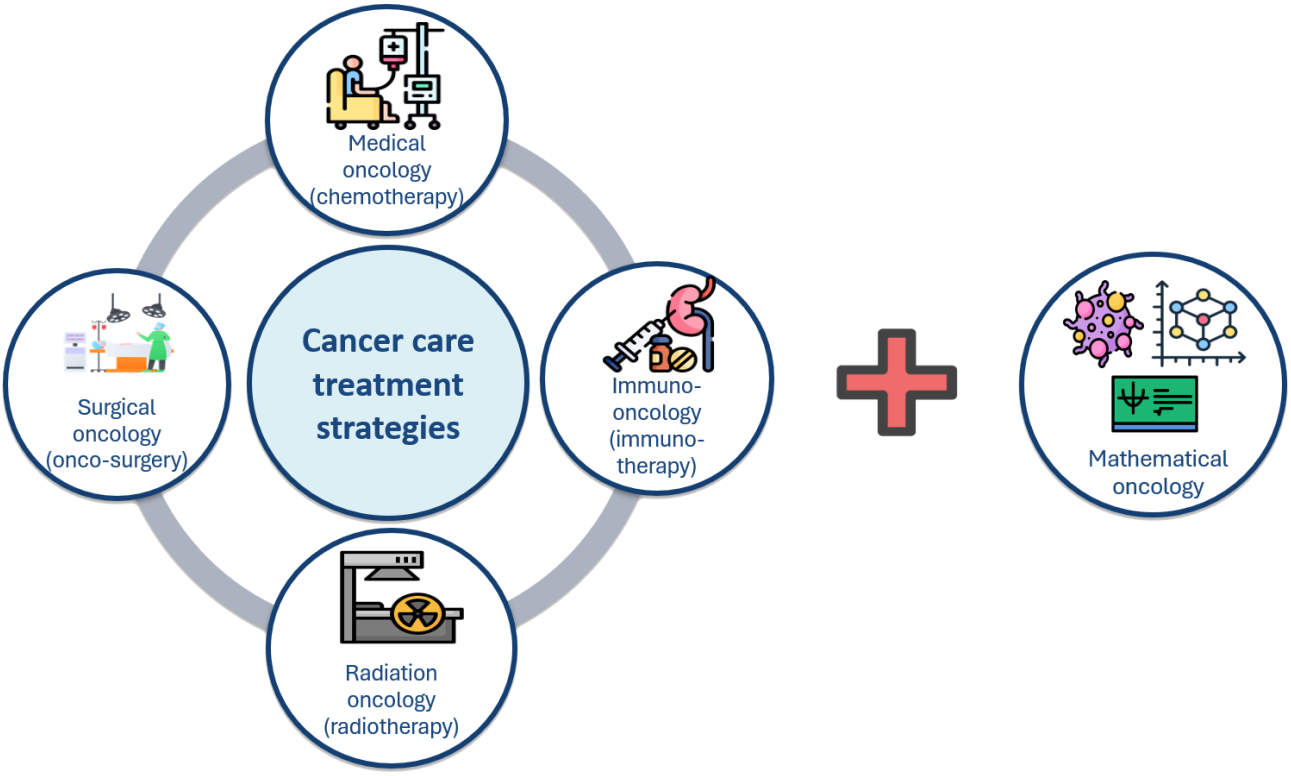
The *Innovative Modeling Approach* section examines IoT-based health care applications in cancer care and integrates IoT with business process management to develop an end-to-end cancer care model. The *Implementation and Challenges* section discusses operational and security considerations for IoT-enabled cancer care systems. The paper concludes with key findings and directions for future research.

Overview and Literature

Cancer and Treatment Strategies

Oncology is the medical discipline dedicated to the prevention, diagnosis, and treatment of cancer. Cancer results from genetic mutations that disrupt normal cellular regulation, leading to uncontrolled proliferation, tumor formation, tissue invasion, and potential metastasis [11]. Four primary clinical treatment modalities are widely used: onco-surgery, chemotherapy, radiotherapy, and immunotherapy, which are often applied individually or in combination, as illustrated in Figure 1.

Figure 1. Cancer care treatment strategies and mathematical oncology.



Mathematical oncology is an interdisciplinary field that applies mathematical methods to study cancer dynamics and treatment outcomes, traditionally focusing on the responses of normal and cancerous cells to therapeutic inputs [6,12-16]. These models emphasize biological and physiological parameters—such as cell growth rates, population dynamics, invasion, and progression—to predict tumor growth, treatment response, and metastasis, thereby informing clinical decision-making.

With the rapid advancement of IoT in health care, integrating smart, connected devices into patient care enhances the practical relevance of these models. IoT-enabled mathematical modeling leverages real-time data from connected sensors to support adaptive, data-driven oncological practices beyond traditional mathematical oncology.

Mathematical Modeling of Cancer Treatment

Overview

Cancer has a substantial global impact, with treatment outcomes varying widely due to pronounced patient and tumor heterogeneity, which complicates the prediction of disease progression. Cancer arises from genetic alterations driven by inherited predispositions, environmental exposures such as radiation and chemical carcinogens, and viral agents (deoxyribonucleic acid or ribonucleic acid viruses), leading to dysregulated cell cycles through oncogene activation and inactivation of tumor suppressors [17,18]. Mathematical oncology offers quantitative frameworks to improve prediction and support more effective treatment strategies.

A wide range of mathematical models has been developed to characterize tumor growth, treatment dynamics, and control mechanisms, particularly for chemotherapy, radiotherapy, and immunotherapy, though onco-surgery remains largely unaddressed. Growing evidence supports the use of mathematical modeling in cancer treatment, as patient-specific responses and tumor size dynamics challenge population-level, “one-size-fits-all” approaches [12,13,18-20]. Models of untreated tumor growth include linear, exponential, logistic, and sigmoidal forms such as Bertalanffy and Gompertz functions [20,21], while additional formulations describe immune-tumor

XSL•FO
RenderX

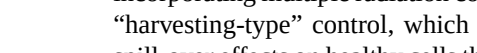
interactions under immunotherapy, chemotherapy, and vaccine-based treatments [4,22].

Radiotherapy or Radiation Oncology

Liu and Yang [23] examined the dynamical behavior of normal cells under periodic radiation using mathematical models, deriving conditions for the persistence or extinction of normal and irradiated cells. Belostotski et al [13] proposed a Lotka-Volterra–based competition model capturing interactions between healthy and cancer cells under radiotherapy, incorporating multiple radiation control mechanisms, including "harvesting-type" control, which has been shown to produce spill-over effects on healthy cells that can be quantified to refine treatment strategies [13,19]. Weerasinghe et al [24] highlighted the importance of quantitative models in studying cancer cell plasticity, using them to inform therapeutic design and optimization across tumor progression stages such as avascular and vascular growth, angiogenesis, invasion, and metastasis within the tumor microenvironment.

Chemotherapy or Medical Oncology

Isaeva et al [16] use differential equation–based chemotherapy models to examine interactions among tumor, healthy, and immune cells, linking theoretical analysis with empirical evidence. Similarly, Song et al [10] propose a model that incorporates three cell populations and drug concentration to study immune-mediated tumor control and the side effects of chemotherapy on both tumor and effector cells. The dynamics of tumor and immune cell populations under chemotherapy are described by the following model [10]:



with initial conditions:

.N(0)=N0>0, L(0)=L0>0, T(0)=T0>0, and u(0)=u0>0

The variables $T(t)$ and $N(t)$ represent the population sizes of cancer cells and Natural Killer *NK* cells, respectively, at any given time t; $L(t)$ for the cytotoxic tumor cell (CTL) population; and $\mu(t)$ for the amount of drug at the tumor site. The parameters in equation (1), along with their descriptions, are as depicted in Table 1.

Table . Description of the parameters in equation (1).

Parameter	Description	Parameter	Description
	Growth rate of NK[a] cells	β_1	CTL[b] death rate
	Inverse of NK cells capacity	β_2	Rate of CTL-induced tumor death
c	Growth rate of tumor	ν	Influx of drug
d	Inverse of tumor capacity	ω	Drug decay rate
r	Activation rate of CTLs	k_N, k_L	Immune cells killed by drug[c]
μ	CTL death rate	k_T	Tumor cell killed by drug[c]
α_1	NK cell death rate	—[d]	—
α_2	Tumor death rate of NK	—	—

[a]NK: natural killer cell.

[b]CTL: cytotoxic tumor cell.

[c]Note that the 3 different response coefficients are denoted by k_N, k_L, and k_T because chemotherapy affects all 3 cell populations and the model assumes side effects on both tumor and effector cells [10].

[d]Not applicable.

Immunotherapy or Immuno-Oncology

In immunotherapy, the effect of vaccine therapy is considered a parametric perturbation of the model commonly used in chemotherapy, in which simulations show that the efficiency of vaccine therapy depends on tumor size, the condition of the immune system, and the response of the organism to vaccination. The goal of immunotherapy is to enhance the antitumor resistance of an organism and improve the immune system's condition [12]. There are 3 known main categories of immunotherapy: immune response modifiers (cytokines), monoclonal antibodies, and vaccines. Such immune response modifiers as interleukin-2 (IL-2), interferon-α (IFN-α), and tumor necrosis-α are widely used in cancer immunotherapy [10,12,23]. The system of equations alluded to in [10] for tumor-immune dynamics is based on the presence of treatment components consisting of 5 populations: tumor cells (T), cytotoxic tumor cell population (L), cytokine IL-2 population (I_2), chemotherapeutic drug (C), and IFN-α population (I).

(2)dTdt=−aTln (bTa)−c(I) TL−MT(I2)(1−e−C)T

(3)dLdt=d+eLI2−fL−ML(I2)(1−e−C)L

(4)dI2dt=VI2(t)+gTT+l−jLI2−kTI2

(5)dCdt=VC(t)−pC

(6)dIdt=VI(t)−qI

The tumor growth is described by Gompertzian law (the first term in equation 1). The destruction of tumor cells by *CTL* is presented by the second term in equation (1) in which the destruction rate is proportional to the number of tumor cells and *CTL* cell populations. In equation (3), d characterizes the steady inflow of *CTL* into the tumor site. The second and third terms in equation (2) describe CTL proliferation in response to IL-2 action and *CTL* death rate. In equations (3-5), ($I = I_2$, C, I)

XSL•FO
RenderX

describes the external influxes of IL-2, a chemotherapeutic drug, and IFN-α, respectively. Since therapy is assigned to a certain schedule, these influxes are taken to be time-dependent. IL-2 production in equation (3) is described by hyperbola (the second term), which allows us to consider a limitation in the stimulation of the immune system by the growing tumor. At small T (ie, $T <=$), the growth rate is nearly linear in tumor size, while for big tumor (ie, $T>>$), it tends to a maximum constant value g. The parameter influences the IL-2 production rate. The smaller the value of , the quicker the IL-2 production rate achieves its maximum value g.

Radiotherapy and Mathematical Oncology

Belostotski and Freedman [19] introduced a mathematical model for radiotherapy treatment strategy based on the Lotka-Volterra competition system, which was later extended by Liu and Yang [23]. The model assumes that cancerous and healthy cells coexist within the same region of the organism and that radiation preferentially eliminates cancer cells while exposing healthy cells to lower doses. Radiotherapy is thus modeled as a harvesting-type control mechanism that regulates the rate of change in both cancer and healthy cell populations, as discussed in Multimedia Appendix 1.

Limitations of Mathematical Models

While mathematical models offer valuable frameworks for understanding cancer dynamics, several challenges limit their direct clinical applicability. Ghaffari et al [12] highlight a key gap between theoretical mathematical oncology and clinical practice, largely due to the limited availability of high-quality, patient-specific clinical data required for robust model validation. Song et al [10] similarly note that difficulties in accurately predicting treatment outcomes hinder practical implementation. Additional challenges include patient-level variability in parameter estimation, scarcity of comprehensive clinical datasets, reliance on simplifying assumptions that may not capture complex biological behaviors, and the difficulty of modeling treatment resistance. As noted by Yin et al [20], these factors introduce significant uncertainty into clinical use. In particular, critical parameters often derived from in vitro or in vivo experiments are difficult or impossible to measure directly, necessitating indirect estimation and inferential approaches. Collectively, these limitations constrain the routine clinical deployment of mathematical cancer models [25].

The study of complex IoT interventions in cancer care is constrained by the limited availability of IoT data, heterogeneity in care processes and interventions, and the intrinsic variability of cancer characteristics. Despite these challenges, mathematical models stand to benefit substantially from IoT integration, while modeling and simulation within IoT environments offer opportunities to mitigate many existing limitations. Incorporating additional data sources, such as nanotechnology-enabled sensing, and enabling real-time adaptation to diverse data streams can enhance predictive accuracy and clinical relevance. Although it has been suggested that the complete eradication of cancer cells is unattainable regardless of treatment type or timing [21], emerging evidence indicates that wireless sensor networks and IoT technologies can significantly improve treatment personalization and effectiveness, as demonstrated in our prior work [7]. Addressing data-related barriers—particularly IoT data governance, interoperability, quality, privacy, and security across the data lifecycle—is, therefore, essential for effective integration.

Overall, Yin et al [20] provide a comprehensive overview of mathematical models for tumor dynamics in cancer treatment, including empirical models of therapeutic effects, as well as algebraic, partial differential, stochastic, and deterministic formulations such as game-theoretic and integral-differential approaches. Building on this foundation, the models discussed in the preceding section—ranging from ordinary differential equations and tumor heterogeneity frameworks to biologically integrated tumor process models—inform the adaptation of these mathematical approaches to an IoT-enabled treatment context. This adaptation aims to strengthen computational and simulation capabilities, addressing the limitations of purely analytical models, as discussed in the following section.

Patient and Clinician Adoption of IoT Interventions in Cancer Care

The adoption of the proposed approach in cancer care requires a supportive environment for change, including adequate resources and process redesign. Key resources include human, physical, network, and information assets, with patients and clinicians central to implementation and evaluation. Patient adoption is likely to be influenced by concerns about data privacy and security; targeted education, transparent data governance, user-centered design, and ongoing support can enhance trust, engagement, and adherence. Clinician adoption, on the other hand, depends on early involvement in system design to ensure alignment with clinical workflows and usability requirements. Although workload and technical complexity may pose barriers, the potential to streamline workflows, improve task integration, and optimize resource utilization can support clinician acceptance and sustained use.

The IoT intervention is intended to serve as a mathematical control mechanism in which IoT-aware parameters actively modify the system dynamics of the mathematical models. While the tumor-immune interaction equations are adapted from established mathematical oncology models, the proposed IoT-enhanced immunotherapy framework introduces a fundamental shift by enabling IoT-driven parameter adaptation. Unlike traditional IL-2 tumor-immune models that assume static or population-averaged parameters, the IoT-enhanced formulation incorporates time-varying, patient-specific parameter modulation informed by continuous physiological and biomarker sensing. This enables dynamic responses to toxicity, treatment tolerance, and disease progression, improving predictive stability and clinical relevance. The additional term introduced in the revised equations represents an IoT-driven feedback mechanism that adaptively modulates immune dynamics based on real-time patient data, extending classical fixed-parameter models into responsive, patient-specific systems. This formulation explicitly illustrates and justifies the role of IoT-driven modifications, as discussed in subsequent sections.

XSL•FO
RenderX

Case Study (Immunotherapy for Melanoma)

Overview

In demonstrating the impact of IoT integration in cancer care services, this study presents a case study based on an adapted immunotherapy mathematical model (equations 2-6) introduced in the *Overview of Cancer and Treatment Approaches* section. Immunotherapy was selected because such models are widely applied in melanoma treatment and leverage immune system stimulation to target cancer cells. Compared to other treatment modalities, immunotherapy requires continuous monitoring and dynamic management to mitigate adverse effects, making IoT-enabled devices particularly valuable for real-time patient monitoring and adaptive, data-driven care.

Background

Melanoma patients often receive immunotherapy, which includes treatments with cytokines and immune checkpoint inhibitors. In this treatment option, the continuous monitoring and precise control of drug administration can significantly enhance the efficacy of such treatment therapies through the interfacing and integrating smart and intelligent connected devices.

IoT Integration

1. Enabled (wearable) sensors: Patients wear IoT-enabled devices that monitor physiological parameters indicative of immune responses, such as body temperature, heart rate, and biomarker levels in sweat or blood.
2. Embedded (implantable) devices: Smart connected sensors are implanted near the tumor site to continuously measure tumor size and local immune markers directly.

This demonstrates the IoT integration for cancer treatment, which allows the enabled and embedded sensors and devices to be deployed in cancer care services by forming autonomous wireless ad hoc networks that interact with patients to collect information from patients and their environments. Through these ad hoc networks, the sensor nodes collaborate to perform the required task or tasks, such as processing, transmission, and aggregation of data from various sources, which are used to determine the nature of the illness by those in the circle of care to determine the nature of treatments and diagnoses to offer. As observed, data and information from sensors are transmitted or routed from various patient sites within the network to the data center through cloud services, facilitating the exchange of data and information from patient locations to the main hospital data center.

During implementation, various components are integrated to ensure the functioning of IoT-based systems or environments. Although these elements are broadly grouped at the component level, together they form an IoT-based component ecosystem. In practice, modeling an IoT-enhanced immunotherapy system for melanoma requires the integration, interconnectivity, and coordination of multiple components, including device, communication criteria, access technologies, communication network, data management, and application layers. Each component serves a distinct role in enabling effective operation and service delivery within the system.

The communication network component enables data sharing and traffic transmission from the point of observation to the data management platform and onward to various system applications. It operates through the access technologies component and supports communication across the network using routing protocols, different topologies and architectures, transmission channels, geographical locations, wireless sensor networks, wireless access points, and internet connectivity from the internet service provider, while ensuring security and encryption.

Therefore, adopting an appropriate routing strategy in wireless sensor networks is essential to balance optimality and efficiency while maintaining computation and communication performance [26]. This is especially important as health care resources become more interconnected and the number of sensor motes continues to grow. These motes collect sensory data by measuring, monitoring, and recording physical conditions in the health care environment and are connected through the internet to a central network infrastructure that enables data transmission and exchange, ultimately improving operational processes and service delivery.

Proposed IoT-Enhanced Mathematical Framework

Mathematical model adaptation: The existing model equations can be adapted by introducing terms (ie, new variables) representing data inputs from IoT devices and adjusting treatment protocols based on real-time data.

The IoT-enhanced equations are derived from the baseline model by introducing IoT-driven input functions that adapt system dynamics in real time. For each added IoT function parameter, we provide corresponding biological and computational rationales to justify its inclusion in the mathematical framework.

Formula 1. Formula for tumor cell dynamics (T)

$$dTdt=-aTln\,(bTa)-c(I)\,TL-MT(I2)(1-e-C)T+ST(t)$$

As shown in Formula 1, $S_T(t)$ represents the real-time tumor size adjustment based on data from the implantable or embedded sensors for its effects on immunotherapy.

- Biological rationale: Tumor burden evolves heterogeneously and may change rapidly in response to immunotherapy. Implantable or embedded sensors (eg, tumor microenvironment probes, imaging-derived biomarkers) can provide near-real-time indicators of tumor growth, hypoxia, or metabolic activity that are not captured by fixed growth or killing rates.
- Computational rationale: The term $S_T(t)$ introduces a corrective adjustment to tumor dynamics based on observed tumor-state deviations. Computationally, this term allows the tumor growth trajectory to be continuously recalibrated using sensor-derived feedback, improving model responsiveness and reducing reliance on static growth assumptions.

Formula 2. Formula for CTL dynamics (L):

$$dLdt=d+eLI2-fL-ML(I2)(1-e-C)L+SL(t)$$

XSL•FO
RenderX

As shown in Formula 2, $S_T(t)$ represents the CTL modulation influenced by wearable sensor data for its effects on immunotherapy.

- Biological rationale: CTL activity is sensitive to systemic stress, fatigue, and immune suppression, which can be indirectly inferred from wearable sensor data such as heart rate variability, temperature, or activity levels. These physiological signals correlate with immune competence and treatment tolerance.
- Computational rationale: The term $S_T(t)$ enables time-varying modulation of CTL population dynamics in response to the patient's state. This prevents the unrealistic persistence of immune efficacy under adverse conditions and allows the model to reflect immune exhaustion or recovery observed clinically.

Formula 3. Formula for IL-2 Dynamics (I_2):

$$dI2dt=VI2(t)+gTT+l-jLI2-kTI2+SI2(t)$$

As shown in Formula 3, SI2(t) represents an adjustment in IL-2 based on continuous biomarker monitoring for its effects on immunotherapy that is, a function of a sensor-derived toxicity or inflammation signal $z(t)$.

To illustrate the qualitative effect of IoT integration on immunotherapy dynamics, we introduce a normalized sensor-derived control signal $z(t) \in [0,1]$ representing a composite patient toxicity and immune-activation index. This signal, computed from continuously monitored physiological and biomarker data (eg, body temperature, heart rate variability, oxygen saturation, and inflammation-related markers), drives the IoT feedback term SI2(t;z(t)) in the IL-2 dynamics. The term SI2(t;z(t)) enables real-time, patient-specific modulation of immune response, transforming the model from a static, open-loop formulation into an adaptive, feedback-driven system. This enhances predictive stability and clinical plausibility by reflecting treatment adjustments informed by continuous monitoring rather than fixed schedules.

- Biological rationale: IL-2 levels are tightly linked to immune activation and toxicity, with excessive IL-2 associated with severe adverse effects. Continuous biomarker monitoring (eg, cytokine levels, inflammatory markers) provides early signals of immune overactivation or suppression.
- Computational rationale: The IoT-driven SI2(t) acts as an adaptive feedback mechanism that adjusts IL-2 dynamics in real time. This allows the model to represent dose adjustment, treatment delay, or attenuation strategies triggered by early biomarker changes, improving stability and clinical plausibility.

Formula 4. Formula for drug dynamics (C):

$$dCdt=VC(t)-pC+SC(t)$$

As shown in Formula 4, $S_C(t)$ represents the chemotherapeutic drug infusion rates controlled by IoT devices to maintain optimal drug levels for its effects on immunotherapy.

- Biological rationale: Chemotherapeutic and immunotherapeutic drug infusion rates are increasingly controlled by smart pumps and IoT-enabled delivery systems that respond to patient-specific tolerance and toxicity signals.
- Computational rationale: The term $S_C(t)$ converts drug administration from a predefined input into a sensor-driven control variable. This enables dynamic regulation of drug concentration to maintain therapeutic efficacy while minimizing toxicity, reflecting real-world adaptive dosing practices.

Formula 5. Formula for IFN-α dynamics (I):

$$dIdt=VI(t)-qI+SI(t)$$

As shown in Formula 5, $S_I(t)$ represents the IFN-α level adjustments based on sensor data for its effects on immunotherapy.

- Biological rationale: IFN-α plays a central role in immune activation but is also associated with systemic side effects. Sensor-derived indicators such as fever, fatigue, and inflammatory response provide indirect measures of IFN-α activity and tolerance.
- Computational rationale: The IoT term $S_I(t)$ allows IFN-α dynamics to be adaptively adjusted based on patient response, preventing sustained overactivation and enabling more realistic immune regulation within the model.

Implementation

- *Real-time data processing*: IoT devices transmit data to a central system where they are processed in real time. The system uses algorithms to predict and adjust the influx rates (VI2(t), VC(t), VI(t)) and introduces corrections (ST(t), SL(t), SI2(t), SC(t) and SI(t)) based on the patient's current state and treatment response.
- *Feedback system*: The treatment regimen is adjusted dynamically, allowing for personalized therapy schedules that optimize immune response and tumor suppression based on continuous feedback.

Conclusion

Integrating IoT in immunotherapy allows for dynamic, personalized treatment adjustments that can potentially lead to better patient outcomes by ensuring optimal drug levels and immune responses at all times. This approach leverages the power of real-time data to make informed decisions on the complex dynamics of cancer treatment.

Definitions of Parameters

The mathematical modeling of oncology treatment pathways and real-time patient monitoring relies on a specialized set of variables, which are categorized into biological parameters, IoT input functions, and time-dependent external influxes

1. Biological parameters:
   - Tumor growth parameters (a and b): They reflect the natural growth rate and carrying capacity of the tumor.
   - Immune response modifiers ($c(I)$, $M_T(I_2)$, and $M_L(I_2)$): They depend on the specific immune response elicited by different cytokines or drugs.

XSL•FO
RenderX

- Drug effectiveness (e^{-c}): It represents the effectiveness of the chemotherapeutic drug, typically modeled as an exponential decay.
- *CTL* dynamics (*d*, *e*, and *f*): They represent the rate of *CTL* inflow to the tumor, *CTL* proliferation rate, and death rate, respectively.
- Production and decay rates (*g*, *j*, *k*, *p*, and *q*): They control the dynamics of *IL-2*, *CTL*, and drug concentrations.
- I: It affects the production rate of IL-2 by the tumor.

2. IoT input function parameters:
   - Sensor data adjustments (ST(t), SL(t), SI2(t), SC(t), and SI(t)): They are model functions that adjust parameters based on real-time data from IoT devices.
3. External influxes:
   - Time-dependent influxes (VI2(t), VC(t), VI(t)): They define how external treatments are administered over time.

Innovative Modeling Approach to Cancer Care Through IoT

Overview

Cancer research is increasingly moving toward intelligent, IoT-enabled health care delivery for cancer care [1,5-7,22]. Integrating IoT technologies with mathematical modeling enhances the precision of treatment dynamics by enabling the adaptive control of therapy parameters and patient-specific responses. This convergence gives rise to a mathematically modeled IoT-based cancer care system that embeds control mechanisms, parameters, and behaviors within IoT-driven clinical workflows.

Simulation and computational approaches can effectively bridge the gap between the theoretical frameworks of mathematical oncology and clinical practice. Central to these models are mathematical equations, especially differential equations, which quantify the behaviors and relationships within biological systems [12,16,17]. These models use specific biological parameters, such as growth rates or drug responses, enabling the simulation of diverse scenarios. Real-time data integration allows for further dynamic adjustments to parameters in response to individual-level factors or conditions. Overall, the responsiveness and dynamism of computational and simulation-enhanced mathematical models support the rapid validation of IoT-enabled cancer treatment effects under different scenarios.

The complex modeling required to validate the effects of IoT-enabled cancer treatment underscores the need for approaches that are both responsive and informed by underlying mechanisms. These models transcend the limitations of data reliance typical of population-level statistical methods, including some computational machine-learning techniques. Instead, they use mathematical models that clarify dynamic causal mechanisms between individuals, supplemented by the best available data, thereby enhancing predictive accuracy for personalized medicine.

In oncology, mathematical modeling approaches such as agent-based models and cellular automata are increasingly used [12]. These models are pivotal in understanding the biological complexities of cancer and developing effective therapies. When considering the application and implementation of IoT technologies, methods such as process simulation and digital models offer another dimension through which mathematical oncology can inform and optimize health system operations. These approaches provide advanced computational techniques to enhance the efficiency and effectiveness of service delivery and overall management within health care systems, ensuring a comprehensive and integrated approach to both individual patient care and broader health system settings.

By integrating IoT and related technologies, we can address many challenges, such as resistance and assumptions, that have hindered the effectiveness of various models in clinical trials and routine patient care. This integration enhances personalized approaches, helping to pinpoint the most effective treatment options and therapeutic regimens. Additionally, the use of advanced sensors, such as nanodevices implanted in patients, allows for the real-time monitoring of cancer's nature and progression. Coupled with intelligent medical image analysis and the application of 3D computer-interactive technologies in virtual reality, augmented reality, and mixed reality contexts, these innovations offer significant potential to improve cancer treatment and the patient's health care experience more effectively and efficiently [3].

Data sharing in health care environments, particularly with IoT-generated big data, presents a challenge due to the diversity of smart connected sensors, objects, devices, data, services, and other supporting applications resulting from the real-time flow and volume of data between systems. The adoption of health care standards and protocols for data interoperability, such as Health Level 7 connectivity, Digital Imaging and Communications in Medicine, and Fast Healthcare Interoperability Resources, may be critical to facilitating appropriately scaled and seamless data exchange between systems. In contrast, conformance to privacy and security standards and policies, in addition to these technical standards, can help protect highly sensitive personal health information from privacy risks and incidents.

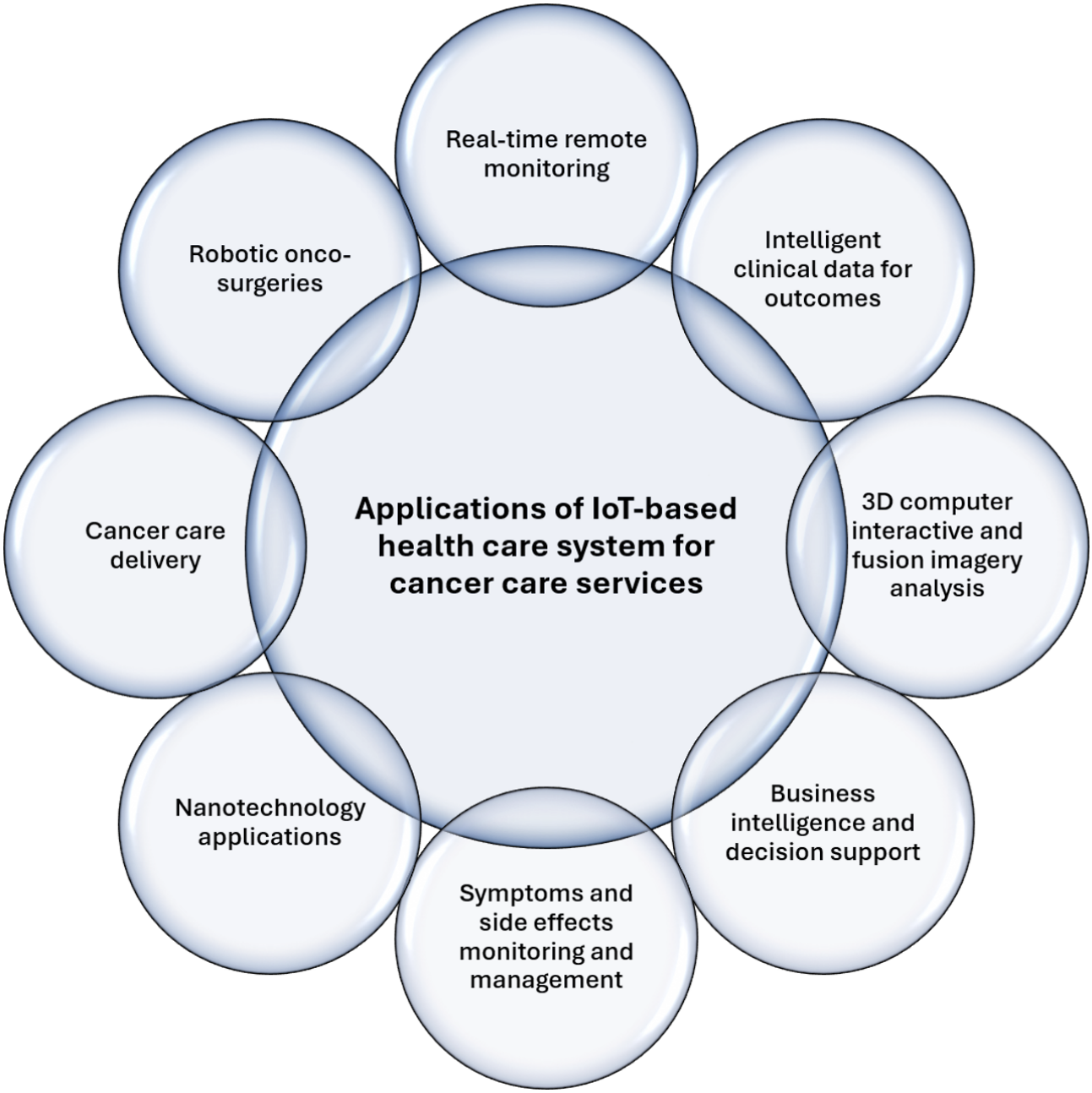
Applications of IoT-Based Health Care System for Cancer Care

Overview

Mathematical biological parameters describing cancer growth, cell populations, invasion, and progression are integrated into IoT-based cancer care services to enhance the design, simulation, and deployment of integrated health care solutions. The convergence of IoT and related technologies enables novel, data-driven approaches to cancer care. Figure 2 illustrates key IoT-based health care applications for cancer care.

XSL•FO
RenderX

Figure 2. Applications of Internet of Things (IoT)–based health care system for cancer care.



Real-Time and Remote Monitoring

IoT devices are seamlessly integrated with cancer treatment protocols to enable real-time and remote monitoring of patients, including the capture and exchange of vital signs, patient-reported measures, activity, and biometric data across various geographical locations, enabled by IoT-connected technologies.

Intelligent Clinical Data for Outcomes

Integrating IoT-related technologies into cancer care facilitates a significant shift from traditional to digital health care. The electronic health record is anticipated to help manage this data stream, characterized by an expanded volume and variety of quality data for routine use in value-based learning health systems. These capabilities, not offered by IoT modeling approaches alone, support comprehensive patient management, including communication, monitoring, alerting, diagnosis, prognosis, and treatment. However, streams of personal data collected from IoT-connected devices carry ethical implications. Ethical guidelines and regulations for cancer patient data must be established to safeguard patient data against hacking, compromise, misuse, misdiagnosis, unauthorized access, attacks, and breaches, while improving transparency and accountability.

3D Computer Interactive and Fusion Imagery Analysis

Intelligent medical image analysis using 3D computer interactive and fusion imagery technologies will significantly improve cancer treatment, especially in personalized health care experiences, making treatment more effective and efficient.

Business Intelligence and Decision Support

IoT-based cancer care provides a framework for analyzing patient data from IoT-enabled devices through business

XSL•FO
RenderX

analytics. This assists health care providers and clinical researchers in converting data streams about patients' health conditions into actionable insights and evidence-based decision-making at all levels.

Symptoms and Side Effects Monitoring and Management

Systemic cancer treatment may lead to significant side effects and complications such as allergies, side effects, and other adverse effects. Current mathematical modeling approaches do not account for these variances, but IoT platforms provide an interface to collect data from various devices and analyze them to detect anomalies. This capability is essential for rapid diagnosis and treatment adjustments, thereby effectively improving the quality of life of patients with cancer by managing these issues.

Nanotechnology Applications

The integration of IoT with nanotechnology (such as nanomaterials, nanodevices, nanostructures, nanoparticles, nanoprocessors, and nanonetworks) and communication engineering has led to the development of the Internet of Nano-Things and the Internet of Bio-Nano Things. These platforms facilitate interconnectivity, automation, and data exchange by embedding biosensors and computing nanodevices in biological systems, thereby enhancing cancer treatment through precise monitoring and targeted therapy delivery.

Cancer Care Delivery

Integrating IoT technologies with oncological treatment strategies changes how cancer care is administered by providing more responsive, personalized, and accessible treatment experiences that promote patient survival and quality of life. IoT devices, including smart sensors, enable the continuous monitoring of cancer evolution, optimizing treatment plans, and minimizing exposure of healthy tissues to harmful effects. In addition, IoT collaborative delivery models facilitate real-time communication and coordination among health care providers and empower patients by providing access to their health data for shared decision-making.

Robotic Onco-Surgeries

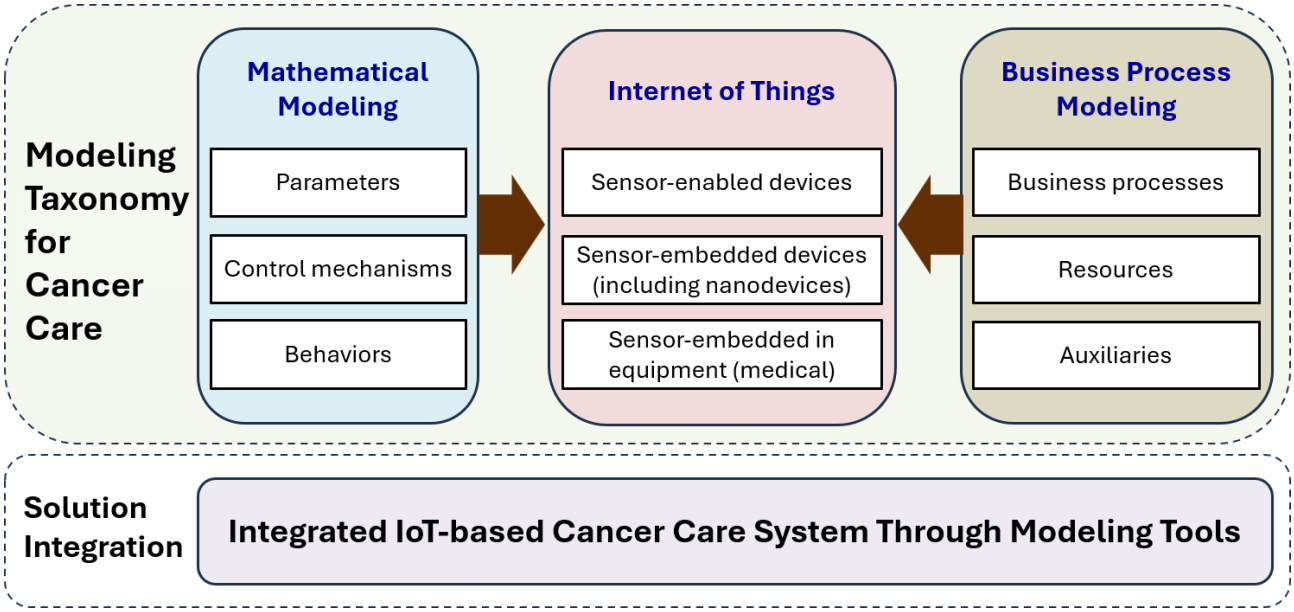
Integrating IoT with surgical robotics enables precise onco-surgical procedures by supporting high-accuracy, time-intensive interventions under continuous monitoring. IoT-enabled robots, augmented with embedded artificial intelligence, enhance surgical precision in targeting and excising locoregional tumors, including lymphadenectomy, while preserving healthy tissue within clinical margins. Real-time IoT monitoring during and after surgery supports patient stability, surgical condition awareness, and coordinated management of drains and tubes, thereby improving patient safety and postoperative recovery.

Integrated IoT-Based Cancer Care System Through Modeling Tools

Cancer care services stand to benefit substantially from IoT-enabled technologies; however, as IoT technologies enable more personalized oncology treatment strategies, mathematical frameworks must evolve to incorporate the advancements made possible by IoT. Mathematical oncology provides the foundation for IoT-adapted models that support drug delivery optimization, predictive patient monitoring, treatment scheduling, and interpretation of complex sensor-generated data streams. This integration ensures that IoT applications in oncology are clinically relevant and operationally feasible.

With continued advances, integrating IoT, mathematical modeling, and complementary analytical tools supports more automated and adaptive cancer care services, informing workflow redesign, clinical decision support, and improved care delivery. Figure 3 illustrates the taxonomy and interrelationships of cancer care modeling approaches examined in this study and future work, showing how IoT-enabled devices are embedded within business process modeling artifacts—tasks, events, activities, and resources—to enable the continuous monitoring of cancer conditions and characteristics.

Figure 3. Taxonomy of Modeling tools of an integrated Internet of Things (IoT)–based cancer care system.



XSL•FO
RenderX

Business process management (BPM) tools are widely used in health care due to their service-oriented nature [27,28]. In cancer care, BPM provides a structured approach to evaluate IoT-enabled services, tasks, and workflows, while IoT enhances efficiency and service delivery. BPM modeling helps identify redundancies, inefficiencies, and gaps in current processes, guiding the design of improved or future processes or workflows that better support the circle of care. It also informs optimal resource allocation and effective IoT interventions to optimize processes and improve patient outcomes. IoT-enabled BPM enables real-time monitoring and analysis of patient conditions and treatment responses using on-body, in-body, and off-body devices, supporting proactive and adaptive care.

Incorporating IoT into health care systems for cancer care services, especially radiotherapy, has been revolutionizing. A recent report by Chen et al [2] suggests that radiotherapy has evolved as a frontier discipline in cancer care services, driven by innovative advancements in artificial intelligence, which extend human intelligence. This has resulted in the widespread implementation of “intelligent radiotherapy," encompassing various tools, namely, big data, deep learning, extended reality, digital twin, radionics, and the IoT. Recognizing that target contouring and treatment planning are at the heart of the entire radiotherapy process, it becomes clear that there is considerable potential for improvement in this area. Radiation oncologists can begin to enhance the radiotherapy process with intelligent radiotherapy that facilitates automation and intelligence, especially through business process modeling. This demonstrates that when evaluating effects, we cannot overlook the significance of IoT and related digital technologies in the modern health care system landscape.

All these applications and technologies, through smart and intelligent devices and equipment, are adopted and adapted to underpin smart cancer care service delivery, which supports the conclusion that cancer treatment under a combination of multiple treatment strategies will yield better outcomes. This extends even to the use of the mathematical treatment strategies or model-based approaches. Drawing from the assertion by Ghaffari et al [12] that “the field of mathematical oncology is somewhat disconnected from the clinical practice of oncology,” we cannot apply mathematical modeling with only traditional oncology treatments in mind; instead, we must adapt to model the complex interventions that are enhanced by technological advancements.

Implementation and Challenges

It is essential to address the technological, regulatory, operational, and security challenges associated with implementing IoT-enabled cancer care systems within the proposed modeling framework. Mathematical and business process simulations offer a practical means or mechanism to translate or operationalize compliance, security, interoperability, and reliability requirements into testable parameters, enabling their evaluation before system deployment.

From a business process perspective, cancer-care workflows (eg, monitoring, alert generation, clinical decision, and intervention) can be modeled using discrete-event or business process model and notation-based simulations in which network latency, data loss, authentication delays, interoperability failures, or device downtime are injected as stochastic disturbances. These perturbations propagate through the workflow as delayed alerts, missed detections, or increased clinician response times, allowing the system to quantify process-level degradation such as time-to-intervention, queue buildup, or alarm fatigue under varying operational conditions.

From a mathematical modeling perspective, these process-level delays can be directly coupled to treatment-response models (eg, tumor-immune interaction models). For example, latency in vital-sign transmission or intermittent sensor dropout can be represented as delayed or incomplete feedback signals, leading to mistimed dose adjustments or prolonged exposure during adverse events. Simulation enables the quantification of the resulting clinical impact, such as reduced dose intensity, increased toxicity risk, or diminished therapeutic response.

Together, this integrated simulation framework functions as a predeployment environment, enabling the identification of critical thresholds—such as maximum allowable latency, minimum data completeness, or minimum sampling rates under low-power constraints—beyond which safety or efficacy is compromised. These findings can be translated into evidence-based design requirements that guide decisions on edge analytics, redundancy, interoperability standards, and security architectures, thereby influencing the performance of IoT-enabled devices and sensors.

Discussion and Future Work

This research proposes an integrated framework that combines mathematical modeling and IoT-based interventions to enhance cancer care delivery. Drawing on concepts from mathematical oncology, the study explores selected models to examine treatment dynamics while acknowledging the broader landscape of existing cancer models and their associated limitations. These limitations—particularly data scarcity, uncertainty, and system complexity—motivate the incorporation of IoT technologies to balance the trade-offs among patient preferences, clinical outcomes, and health care system constraints. An adaptive, innovative IoT-enabled modeling approach has been developed and demonstrated through a melanoma immunotherapy case study, highlighting the potential of real-time sensing and adaptive decision support. Data interoperability is identified as a key challenge, necessitating the use of health care data standards and secure protocols to enable real-time data exchange while preserving privacy and security.

Future research will advance this work along two complementary directions: mathematical modeling and business process modeling. From a mathematical perspective, future studies will incorporate additional parameters and empirically derived patient data to refine and validate treatment-response models. From a business process perspective, IoT-enabled cancer-care workflows will be simulated to assess performance, scalability, and regulatory compliance across different levels of oncology service delivery.

XSL•FO
RenderX

The integration of these perspectives offers a pathway toward digitally transformed cancer care, enabling the systematic evaluation of both clinical and operational capabilities across the continuum of care—from screening and diagnosis to treatment administration, rehabilitation, and follow-up. Thus, this disrupts current health care practices by replacing traditional health care with a frontier of digital health care. Methodologically, this will involve structured business process modeling using BPMN 2.0 (by means of mathematical modeling of cancer care services), application of the BPM lifecycle to transition from as-is to to-be care pathways, and simulation of IoT-enabled future-state models. Together, mathematical and process-based simulations provide a rigorous foundation for improving care quality, supporting decision-making, and guiding the evolution of IoT-enabled cancer care systems.

Acknowledgments

The authors acknowledge Maher Elshakankiri from the University of Toronto for his early discussions related to this research. We would like to recognize our use of OpenAI's ChatGPT 5.0 for copy and line editing, which has helped improve the readability of this paper.

Funding

This research was partly supported by the University of Regina President's Publication Fund.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A system of radiation treatment models based on competitive Lotka-Volterra dynamics and Manley's asymmetrical division framework.
[DOCX File, 23 KB - cancer_v12i1e73997_app1.docx]

References

1. Onasanya A, Lakkis S, Elshakankiri M. Implementing IoT/WSN based smart Saskatchewan Healthcare System. Wireless Netw 2019 Oct;25(7):3999-4020. [doi: 10.1007/s11276-018-01931-2]
2. Chen G, Cui J, Qian J, et al. Rapid progress in intelligent radiotherapy and future implementation. Cancer Invest 2022 May;40(5):425-436. [doi: 10.1080/07357907.2022.2044842] [Medline: 35225723]
3. Dimitrov DV. Medical internet of things and big data in healthcare. Healthc Inform Res 2016 Jul;22(3):156-163. [doi: 10.4258/hir.2016.22.3.156] [Medline: 27525156]
4. Riazul Islam SM, Humaun Kabir M, Hossain M. The internet of things for health care: a comprehensive survey. IEEE Access 2015;3(10):678-708. [doi: 10.1109/ACCESS.2015.2437951]
5. Onasanya A, Elshakankiri M. IoT implementation for cancer care and business analytics/cloud services in healthcare systems. Presented at: Proceedings of the 10th IEEE/ACM International Conference on Utility and Cloud Computing (UCC 2017); Dec 5-8, 2017; Austin, Texas, USA p. 205-206. [doi: 10.1145/3147213.3149217]
6. Onasanya A, Elshakankiri M. Secured cancer care and cloud services in IoT/WSN based medical systems. Presented at: Second EAI International Conference on Smart Grid and Internet of Things (SGIoT 2018); Jul 11, 2018. [doi: 10.1007/978-3-030-05928-6_3]
7. Onasanya A, Elshakankiri M. Smart integrated IoT healthcare system for cancer care. Wireless Netw 2021 Aug;27(6):4297-4312. [doi: 10.1007/s11276-018-01932-1]
8. Yin Y, Zeng Y, Chen X, Fan Y. The internet of things in healthcare: an overview. J Ind Inf Integr 2016 Mar;1:3-13. [doi: 10.1016/j.jii.2016.03.004]
9. Chatbot App. URL: https://chatbotapp.ai [accessed 2024-04-19]
10. Song G, Liang G, Tian T, Zhang X. Mathematical modeling and analysis of tumor chemotherapy. Symmetry 2022;14(4):704. [doi: 10.3390/sym14040704]
11. Evan GI, Vousden KH. Proliferation, cell cycle and apoptosis in cancer. Nature 2001 May 17;411(6835):342-348. [doi: 10.1038/35077213] [Medline: 11357141]
12. Ghaffari Laleh N, Loeffler CML, Grajek J, et al. Classical mathematical models for prediction of response to chemotherapy and immunotherapy. PLoS Comput Biol 2022 Feb;18(2):e1009822. [doi: 10.1371/journal.pcbi.1009822] [Medline: 35120124]
13. Belostotski G. A control theory model for cancer treatment by radiotherapy [Master's thesis]. : University of Alberta; 2005 URL: https://ualberta.scholaris.ca/server/api/core/bitstreams/62ac2fde-2dd8-49a2-b4d4-e61efd8dac2f/content [accessed 2026-06-19]
14. Brady R, Enderling H. Mathematical models of cancer: when to predict novel therapies, and when not to. Bull Math Biol 2019 Oct;81(10):3722-3731. [doi: 10.1007/s11538-019-00640-x] [Medline: 31338741]

XSL•FO
RenderX

15. Enderling H, Alfonso JCL, Moros E, Caudell JJ, Harrison LB. Integrating mathematical modeling into the roadmap for personalized adaptive radiation therapy. Trends Cancer 2019 Aug;5(8):467-474. [doi: 10.1016/j.trecan.2019.06.006] [Medline: 31421904]
16. Isaeva OG, Osipov VA. Different strategies for cancer treatment: mathematical modelling. Comput Math Methods Med 2009 Jan;10(4):253-272. [doi: 10.1080/17486700802536054]
17. Kuang Y, Nagy JD, Eikenberry SE. Introduction to Mathematical Oncology: CRC Press; 2016.
18. Luo MC, Nikolopoulou E, Gevertz JL. From fitting the average to fitting the individual: a cautionary tale for mathematical modelers. Front Oncol 2022;12:793908. [doi: 10.3389/fonc.2022.793908] [Medline: 35574407]
19. Belostotski G, Freedman HI. A control theory model for cancer treatment by radiotherapy. Int J Appl Mech 2005;25(4):447-480 [FREE Full text]
20. Yin A, Moes D, van Hasselt JGC, Swen JJ, Guchelaar HJ. A review of mathematical models for tumor dynamics and treatment resistance evolution of solid tumors. CPT Pharmacometrics Syst Pharmacol 2019 Oct;8(10):720-737. [doi: 10.1002/psp4.12450] [Medline: 31250989]
21. Manley O. A mathematical model of cancer networks with radiation therapy. J Young Investig 2014;27(6) [FREE Full text]
22. 5 IoT applications in healthcare field you must know. DataFlair. URL: https://data-flair.training/blogs/iot-applications-in-healthcare [accessed 2024-09-15]
23. Liu Z, Yang C. A mathematical model of cancer treatment by radiotherapy. Comput Math Methods Med 2014;2014:172923. [doi: 10.1155/2014/172923] [Medline: 25478002]
24. Weerasinghe HN, Burrage PM, Burrage K, Nicolau DV. Mathematical models of cancer cell plasticity. J Oncol 2019;2019:2403483. [doi: 10.1155/2019/2403483] [Medline: 31814825]
25. Vieira LC, Costa RS, Valério D. An overview of mathematical modelling in cancer research: fractional calculus as modelling tool. Fractal Fract 2023;7(8):595. [doi: 10.3390/fractalfract7080595]
26. Sohraby K, Minoli D, Znati T. Wireless Sensor Networks Technology, Protocols, and Applications: Wiley; 2007.
27. Rolón E, Chavira G, Orozco J, Soto JP. Towards a framework for evaluating usability of business process models with BPMN in health sector. Procedia Manuf 2015;3:5603-5610. [doi: 10.1016/j.promfg.2015.07.748]
28. Sari MP, Pratama NR, Dachyar M. Application Internet of Things (IoT) in healthcare to optimize waiting time using the business process reengineering (BPR) approach. Presented at: Proceedings of the International Conference on Industrial Engineering and Operations Management; Mar 7-10, 2022 URL: https://ieomsociety.org/proceedings/2022istanbul/820.pdf [accessed 2026-06-19]

Abbreviations

BPM: business process management
CTL: cytotoxic tumor cell
IFN-α: interferon-α
IL-2: interleukin-2
IoT: Internet of Things

Edited by J Bender; submitted 15.Mar.2025; peer-reviewed by FS Gharehchopogh, GK Gupta, L Raymond Guo; revised version received 22.Mar.2026; accepted 17.May.2026; published 13.Jul.2026.

Please cite as:
Onasanya A, Kyabaggu R, Hepting D
Internet of Things–Enhanced Mathematical Oncology: Conceptual Framework for Adaptive Cancer Care Modeling
JMIR Cancer 2026;12:e73997
URL: https://cancer.jmir.org/2026/1/e73997
doi:10.2196/73997

© Adeniyi Onasanya, Ramona Kyabaggu, Daryl Hepting. Originally published in JMIR Cancer (https://cancer.jmir.org), 13.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

AI Agents for Multimodal Oncology Diagnosis: Toward Transparent and Traceable Clinical Decision Support

Liuyang Yang[1*], PhD; Liyu Shan[2,3*], PhD; Xiangmei Yao[1*], PhD; Renbin Zhao[1], MD; Zengzheng Li[1], MD; Shuai Feng[1], MD; Yajie Wang[1], PhD

[1]Department of Hematology, The First People's Hospital of Yunnan Province, Affiliated Hospital of Kunming University of Science and Technology, Kunming, Yunnan, China

[2]The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Peking University Cancer Hospital Yunnan, Kunming, Yunnan, China

[3]The Second Affiliated Hospital of Kunming Medical University, Kunming, Yunnan, China

[*]these authors contributed equally

Corresponding Author:
Yajie Wang, PhD
Department of Hematology
The First People's Hospital of Yunnan Province, Affiliated Hospital of Kunming University of Science and Technology
No.157 Jinbi road
Kunming, Yunnan, 650032
China
Phone: 86 871 63639921
Email: kbb165wyj@sina.com

Abstract

Cancer diagnosis depends on data from radiology, digital pathology, molecular profiling, laboratory testing, and longitudinal clinical records. AI performs well in selected tasks, but most systems remain narrow and disconnected from the iterative reasoning required in oncology. This Viewpoint defines an AI agent as a feedback-driven system that maintains task state, selects among governed tools, observes results, and revises its plan under explicit safety constraints. This definition separates agents from multimodal foundation models, retrieval-augmented generation, and fixed workflow automation. We organize the discussion across multimodal data collection, preprocessing, fusion and representation learning, and diagnostic decision support. We distinguished agent-level evidence, component- or infrastructure-level evidence, and prospective propositions throughout. Clinical translation will require resilient failure handling, guideline version control, prospective evaluation, computational and workflow feasibility, and clinician authority over final decisions. The near-term opportunity is therefore transparent and traceable clinical decision support rather than autonomous cancer diagnosis.

(JMIR Cancer 2026;12:e103545) doi:10.2196/103545

KEYWORDS

artificial intelligence agent; multimodal oncology diagnosis; clinical decision support; large language model; human–artificial intelligence collaboration; diagnostic workflow

Introduction

Cancer remains a leading cause of morbidity and mortality. Global Cancer Observatory (GLOBOCAN) 2022 estimated that there were nearly 20 million new cancer cases and 9.7 million cancer-related deaths worldwide in that year [1]. Treatment delays are associated with higher mortality across several cancer indications [2]. Timely diagnostic pathways are therefore important, but they must also account for genomic, histopathological, microenvironmental, and temporal heterogeneity [3].

Modern oncology diagnosis is intrinsically multimodal. Clinicians combine imaging, pathology, laboratory results, molecular assays, and longitudinal electronic records, often across separate specialties and information systems. This fragmentation creates cognitive and coordination burdens that motivate integrated workflows [4,5]. Task-specific AI has advanced lesion detection, pathology classification, molecular modeling, and risk prediction [4,6-8]. However, most deployed or evaluated models address a single bounded input-output task rather than the full diagnostic process.

XSL•FO
RenderX

Large language models (LLMs), multimodal foundation models, and tool-use frameworks create a technical basis for coordinating these tasks [9,10]. However, a model that accepts several modalities is not automatically an agent. Similarly, retrieval-augmented generation (RAG) adds external evidence to generation, and fixed workflow orchestration links predetermined modules, but neither property alone establishes runtime planning or feedback-driven control.

This Viewpoint examines AI agents as auditable orchestration layers for multimodal oncology diagnostics. We organized the field around 4 connected stages: data collection, preprocessing, multimodal fusion and representation learning, and diagnostic decision support. Our central position is deliberately bounded: agentic orchestration may improve coordination and auditability, but clinical benefit and safety remain deployment-specific questions.

From Task-Specific AI to Agentic Oncology Diagnostics

We used an operational definition of agentic behavior. In a conventional fixed pipeline, an output is generated through a predetermined composition, $y=f_n(...f_1(x))$, and the sequence of functions does not change at runtime. An AI agent instead maintains a state s_t and selects an action $a_t = \pi(g, s_t, A_t)$ from an allowed tool set A_t. For a clinical goal g, it observes each tool's result and updates its state before choosing the next action. The practical distinction therefore lies in conditional planning, tool selection, observation, and revision within a governed loop.

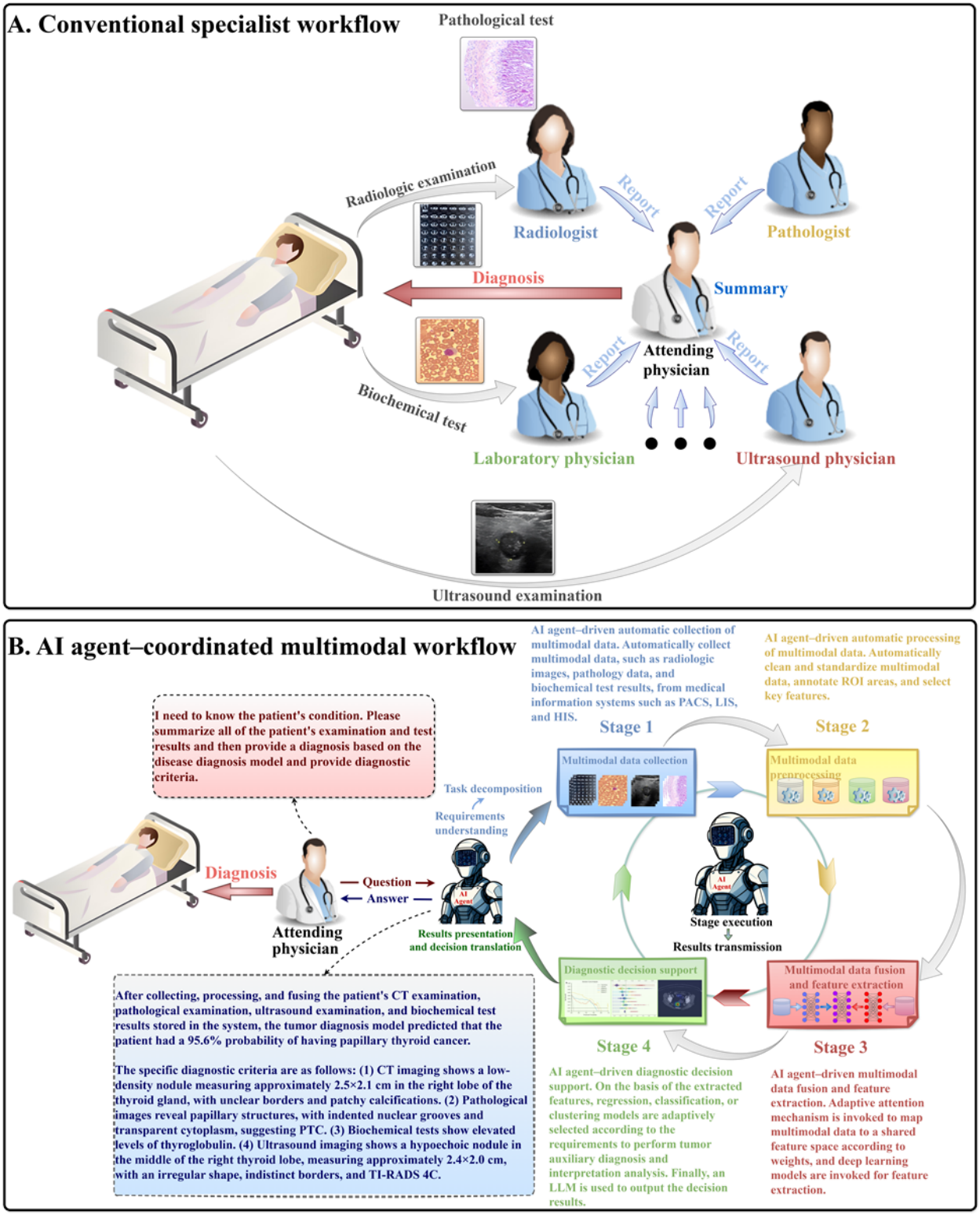
Four adjacent concepts should remain separate. A multimodal foundation model supplies reusable representations or generation across data types. RAG retrieves external sources to ground an answer. Fixed workflow orchestration executes a prescribed sequence of services. A feedback-driven AI agent can change the sequence, request missing evidence, retry or switch tools, abstain, and escalate to a clinician according to the evolving state. Long-term memory is optional and should be limited to an authorized, versioned, patient-specific or institutional state rather than treated as an unrestricted capability.

First, Figure 1 contrasts a conventional specialist synthesis with an agent-coordinated 4-stage workflow. Second, Figure 2 operationalizes the agentic layer as a plan-act-observe-update-check loop. Act coordinates the 4 stages in Figure 1B; update records state and provenance; and check evaluates evidence sufficiency, cross-modal consistency, and whether further action remains within the authorized scope. If these criteria are met, the synthesis proceeds to clinician review. Remediable gaps return the workflow to planning; gaps that cannot be supplemented trigger clinician escalation. The agent-coordinated elements in both figures are prospective design propositions and not validated clinical systems or performance claims.

We applied 3 evidence labels throughout this Viewpoint. Agent-level evidence evaluates a complete loop that plans, uses tools, and produces an integrated output, as in the external oncology evaluation by Ferber et al [11]. Component or infrastructure evidence evaluates a model, retrieval method, data pipeline, or governance mechanism that an agent could invoke [12]. Prospective propositions describe plausible clinical workflows that still require system-level and prospective validation. This hierarchy prevents strong component performance from being misread as proof of a clinically validated agent.

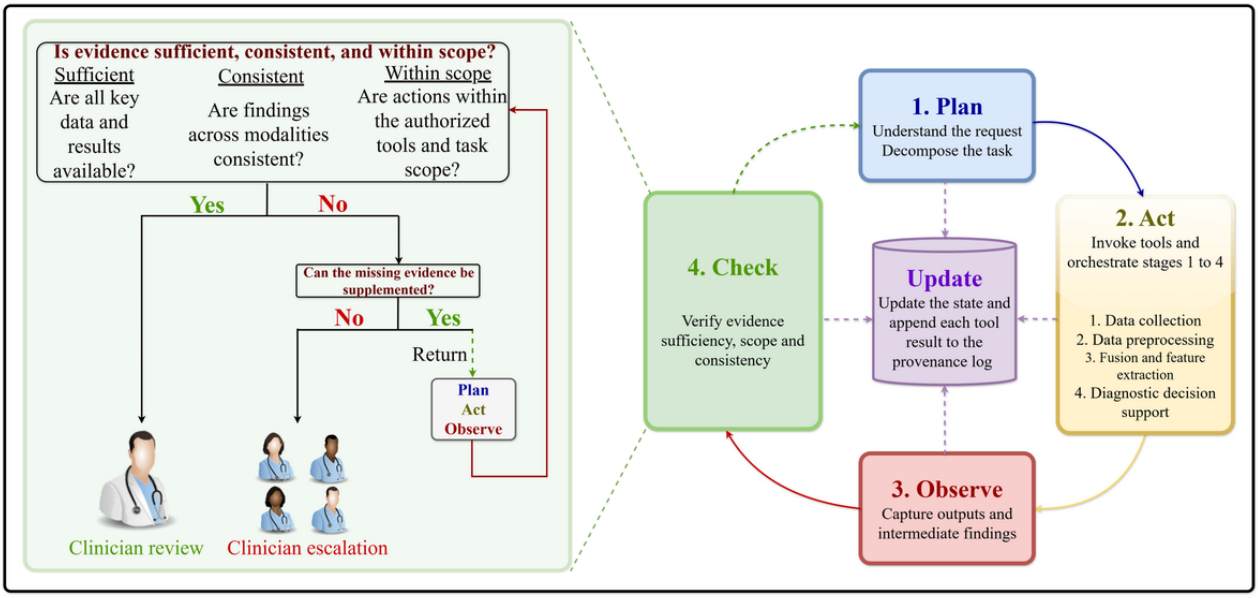
XSL•FO
RenderX

Figure 1. Conventional and AI agent–coordinated multimodal oncology workflows. (A) In the conventional workflow, the attending physician synthesizes modality-specific reports. (B) In the proposed agent-coordinated workflow, the agent coordinates multimodal data collection, preprocessing, fusion and feature extraction, and diagnostic decision support. CT: computed tomography; HIS: hospital information system; LIS: laboratory information system; LLM: large language model; PACS: picture archiving and communication system; PTC: papillary thyroid carcinoma; ROI: region of interest; TI-RADS 4C: Thyroid Imaging Reporting and Data System, category 4C.



XSL•FO
RenderX

Figure 2. Feedback-driven orchestration and evidence-adequacy control for an AI agent in multimodal oncology diagnostics. The control architecture links plan, act, observe, update, and check. Act invokes governed tools across the 4 task stages shown in Figure 1B. Update appends each tool result to the current state and provenance log. Check evaluates whether evidence is sufficient and consistent across modalities and whether further action remains within the authorized scope. If these criteria are met, the synthesis proceeds to clinician review. Remediable gaps return the workflow to planning, whereas gaps that cannot be supplemented trigger clinician escalation.



Multimodal Data Collection

The first stage is the governed assembly of patient-specific evidence. Relevant data may be distributed across a hospital information system (HIS), laboratory information system (LIS), picture archiving and communication system (PACS), electronic medical record (EMR), pathology archives, and genomic platforms. These systems differ in format, ontology, temporal granularity, and access control.

Component-level infrastructure studies show that integration must be engineered. A single-center oncology data supply chain linked clinical, genomic, and imaging data for more than 170,000 patients across 11 cancer types and more than 800 features per case [13]. Its extract-transform-load (ETL), natural language processing (NLP), and quality control procedures support the feasibility of a data layer but not the validation of an agent. Standards such as Health Level Seven Fast Healthcare Interoperability Resources (HL7 FHIR) and Digital Imaging and Communications in Medicine (DICOM) can reduce site-specific translation [14]. Local mapping and governance nevertheless remain necessary.

An agent could use secure interfaces to identify, retrieve, and temporally align data for a specific diagnostic question. It could connect a new scan to prior pathology, biomarker trajectories, sequencing results, and notes [9,15]. Unstructured text and patient-reported or wearable data may add context in selected settings [16]. Each input should carry its source, acquisition and update times, missingness status, and freshness threshold. Contradictory or stale records should trigger clarification rather than silent fusion. Once evidence is assembled, the next risk is whether every modality is technically fit for analysis.

Multimodal Data Preprocessing

Multimodal preprocessing converts heterogeneous inputs into representations that downstream tools can use. Imaging varies by scanner and acquisition protocol. Digital pathology depends on tissue preparation, staining, and slide digitization [6]. Laboratory and genomic data may be incomplete or generated on different platforms, while clinical text often contains duplication and institution-specific shorthand.

A fixed workflow applies the same preprocessing sequence. An agentic workflow could select modality-specific tools according to file type, quality, and clinical purpose. It might invoke imaging quality control and segmentation or pathology stain normalization, tiling, tissue filtering, and feature extraction. Pathology foundation models provide reusable component representations [17,18], while self-supervised medical imaging models illustrate robustness and data-efficiency strategies [19]. Attribution tools such as gradient-weighted class activation mapping can support region-level inspection, although an attribution map is not itself a clinical explanation [20].

Every action should be recorded with the input source, time stamp, parser or model version, quality control result, and review status. If optical character recognition or parsing fails, the system should not fabricate a normalized record. It should retry within a bounded policy, switch to a validated alternate parser, return a partial result with an explicit missingness flag, or request human review. This provenance-aware preprocessing creates the conditions for fusion without concealing upstream uncertainty.

XSL•FO
RenderX

Multimodal Fusion and Representation Learning

Fusion is not the mere presence of multiple data types. Imaging describes macroscopic phenotype, pathology reveals tissue architecture, molecular assays characterize biological drivers, and longitudinal records provide clinical context. Multimodal models can combine these views for characterization and risk estimation [7], but many assume that all modalities are harmonized and simultaneously available.

At the molecular level, machine learning can integrate complementary omics layers and prioritize regulatory interactions. Omidi combined RNA sequencing and microRNA sequencing profiles from 2 public cohorts [21]. The analysis included 78 adrenocortical carcinoma samples from The Cancer Genome Atlas adrenocortical carcinoma (TCGA-ACC) and 250 normal adrenal samples from the Genotype-Tissue Expression (GTEx) project, reconstructing condition-specific competing endogenous RNA networks. Random forest models performed best for predicting microRNA-messenger RNA associations, and predicted interactions were cross-referenced with TargetScan and miRTarBase. This study is an in silico systems biology module that an agent could invoke. Batch effects, cohort imbalance, cross-cohort heterogeneity, and the absence of independent or experimental validation prevent these findings from being interpreted as validation of a multimodal agent.

An agent could select a fusion strategy according to the clinical question, available modalities, and uncertainty. If pathology and imaging are available while genomic testing is pending, it could use a modality-robust model, mark molecular evidence as unavailable, and revise the synthesis later. Explainable multimodal real-world models have identified patient-level marker contributions in large cohorts and undergone independent validation [22]. These results support component-level representation and explanation but not autonomous clinical orchestration. Fusion outputs must therefore pass an evidence-adequacy check before decision support. The evidence-adequacy gate and its return-or-escalate logic are shown in Figure 2.

Diagnostic Decision Support

The final stage translates analyzed evidence into a structured, reviewable diagnostic summary. The purpose is not autonomous diagnosis or treatment selection. It is to state the supported findings, unresolved conflicts, missing evidence, uncertainty, and provenance in a form that clinicians can audit.

A recent feasibility study evaluated Gemini 2.5 Pro (Alphabet Inc) and ChatGPT 4o (OpenAI) on 80 Turkish-language prostate-specific membrane antigen positron emission tomography–computed tomography (PSMA PET-CT) reports [23]. The prompts incorporated the American Joint Committee on Cancer (AJCC) staging, the Chemohormonal Therapy Versus Androgen Ablation Randomized Trial for Extensive Disease (CHAARTED) criteria, and few-shot examples. The models assigned T, N, and M categories and disease volume classes, with overall task accuracies of 93.8% and 91.3%, respectively. Errors clustered around equivocal findings and different interpretive thresholds, including overstaging of ambiguous lesions. The single-center design, 1 expert reference standard, and exploratory sample of 80 reports limit generalizability. The study supports an LLM-based report-to-staging component that requires expert oversight; it does not validate a complete AI agent.

Other component studies illustrate evaluable end points. Multimodal oncology chatbots have been tested on image-containing clinical cases [24]. Explanation types can alter physician trust and diagnostic performance [25]. On the contrary, fabricated details can induce hallucinated clinical elaboration despite simple mitigation attempts [26]. Agent evaluation should therefore include tool selection, evidence citation, contradiction handling, safe refusal, calibration, workflow burden, and final conclusion quality rather than fluency alone.

In practice, an agent might prepare a multidisciplinary team (MDT) case by linking a suspicious lung lesion to pathology, molecular markers, smoking history, and prior imaging. It could rank competing diagnoses, identify a missing confirmatory test, and provide an evidence-linked summary. Proposed MDT assistance and randomized source-attribution studies support case preparation and verifiable retrieval as useful components [27,28]. Clinicians must retain authority over the final judgment, with explicit responsibility assigned across users, institutions, and manufacturers [29].

Challenges and Future Directions

Limits of the Current Evidence

The evidence base remains uneven. Most cited studies validate models, infrastructure, or simulated workflows rather than prospectively deployed agentic systems. The following implementation requirements should therefore be read as a translational agenda and not as established clinical effectiveness.

Data Quality and Interoperability

Interoperability does not guarantee semantic correctness. A deployable system must map local fields to versioned ontologies, record lineage and freshness, and quantify missingness. It must also test performance across scanners, laboratories, institutions, demographic groups, and unavailable-modality patterns [13,14,19]. External validation and subgroup analysis remain necessary because an agent can amplify errors introduced by any upstream component.

Operational Resilience and Graceful Degradation

Real-time MDT use requires an explicit latency budget for retrieval, preprocessing, inference, and rendering. Institutions may precompute stable features, use asynchronous tasks, or choose smaller local models when the expected benefit does not justify the delay. Time-outs should activate circuit breakers rather than indefinite retries. A parser or API failure should trigger a typed error and one of the following governed responses: a bounded retry, use of a validated alternate tool, a partial-result flag, or abstention with human escalation. The audit record should preserve the failed attempt and the fallback

XSL•FO
RenderX

path. This fallback logic corresponds to the supplementation and clinician-escalation branches in Figure 2.

Knowledge Life Cycle and Guideline Synchronization

Oncology guidelines and drug knowledge change frequently. Retrieval stores should preserve source dates, jurisdiction, guideline version, and effective period. Updates require staged ingestion, regression tests against reference cases, comparison with the previous version, canary deployment, and rollback criteria. A response should identify the version used and warn when local policy conflicts with an external guideline. This reduces, but cannot eliminate, legacy hallucinations.

Evaluation and Clinical Integration

Evaluation should follow the clinical workflow and the evidence label. Component metrics include extraction accuracy, calibration, source verifiability, latency, throughput, and failure rate. Agent-level evaluation additionally requires plan appropriateness, tool-use accuracy, recovery from failed steps, contradiction resolution, abstention, clinician workload, and patient safety end points [11,24-28]. Prospective studies should compare the full human-agent workflow with current practice and not only the agent with a static model.

Governance, Cost, and Institutional Adoption

Deployment also depends on computing cost, local infrastructure, cybersecurity, procurement, workflow fit, regulatory status, reimbursement, and institutional adoption. Trustworthy AI principles require fairness, traceability, usability, robustness, and explainability [12]. Precision oncology governance should add access control, deidentification, audit logging, incident response, continuous monitoring, model and knowledge base change control, and explicit accountability [29,30]. These controls are prerequisites for a clinical service and not evidence that the service improves outcomes.

Table 1 summarizes the current evidence supporting AI agents in oncology diagnostics and the limitations that constrain their interpretation and clinical use.

XSL•FO
RenderX

Table 1. Evidence supporting and delimiting AI agents in oncology diagnostics.

Evidence domains	Study design or setting	Main empirical finding	Agentic relevance and boundary
Agent-level oncology decision support	External evaluation on 20 realistic multimodal oncology cases [11]	Reported 87.5% correct tool use, 91% correct clinical conclusions, and 75.5% accurate guideline citation	Agent-level evidence from simulated cases; a useful benchmark but not prospective clinical validation
Real-world multimodal infrastructure	Single-center data supply chain covering >170,000 patients, 11 cancer types, and >800 features per case [13]	Integrated clinical, genomic, and imaging data with extract-transform-load, natural language processing, quality control, and continuously updated records	Infrastructure evidence; supports a callable data layer but does not evaluate agent planning or clinical benefit
Multimodal oncology chatbot benchmarking	Case-based evaluation of multimodal chatbots on 79 image-containing oncology cases [24]	Demonstrated that image-text oncology reasoning can be measured with case-level accuracy	Component evidence; a chatbot benchmark does not establish workflow orchestration or safety
Large language model–assisted prostate cancer staging from unstructured imaging reports	Single-center feasibility study of 80 Turkish-language prostate-specific membrane antigen positron emission tomography–computed tomography reports with embedded American Joint Committee on Cancer and Chemohormonal Therapy Versus Androgen Ablation Randomized Trial for Extensive Disease criteria and few-shot prompting [23]	Overall task accuracy was 93.8% for Gemini 2.5 Pro and 91.3% for ChatGPT 4o; errors clustered around equivocal findings	Component evidence; supports a report-to-staging module requiring expert oversight and not a complete agent
Machine learning integration of molecular regulatory networks	In silico integration of 78 The Cancer Genome Atlas adrenocortical carcinoma tumors and 250 normal adrenal samples from the Genotype-Tissue Expression project using RNA and microRNA sequencing [21]	Random forest best predicted microRNA-messenger RNA associations; predictions were cross-referenced with TargetScan and miRTarBase.	Specialized analytic module; cohort imbalance, batch effects, and the absence of independent or experimental validation limit inference
Tumor MDT[a] workflow	Workflow-focused proposal for AI assistance in tumor MDTs [27]	Identified opportunities for case preparation, synthesis of unstructured data, and rationale capture	Prospective workflow proposition; final decisions remain clinician led and require comparative evaluation
Source-attributed uro-oncology decision support	Randomized reader evaluation of a retrieval-augmented uro-oncology chatbot [28]	Improved recommendation correctness, source attribution, and source verifiability compared with a general chatbot	Component evidence; supports governed retrieval and verifiable sources and not full agent autonomy
Explainable multimodal real-world oncology AI	Multimodal model in 15,726 patients across 38 solid cancer types with external validation in 3288 patients with lung cancer [22]	Estimated patient-level marker contributions and prognostic interactions	Model-level evidence; supports reusable prediction and explanation tools with external validity boundaries
Hallucination and safety	Multimodel simulation of adversarial hallucination during clinical decision support [26]	Fabricated details induced false elaboration despite simple mitigation prompts	Safety evidence supporting contradiction checks, retrieval constraints, abstention, and human escalation
Interoperability	Clinical interoperability analysis across the AI data life cycle [14]	Emphasized standardized information exchange for data entry, processing, sharing, replication, and analytics	Infrastructure evidence; local semantic mapping and provenance remain necessary
Governance, liability, and implementation	Consensus trustworthy AI guidance, liability analysis, and precision oncology implementation review [12,29,30]	Identified requirements for fairness, traceability, usability, robustness, explainability, responsibility, safety, governance, cost, and harmonization	Cross-cutting requirements; governance controls enable evaluation but are not evidence of clinical effectiveness

[a]MDT: multidisciplinary team.

Conclusions

AI agents should be understood as auditable, feedback-driven orchestration systems rather than as another name for multimodal models or RAG. Oncology-specific studies now support selected agent-level, component-level, and infrastructure capabilities, including tool use, report-to-staging, multimodal modeling, and source attribution. However, these findings do not establish the safety or benefit of autonomous diagnosis. The defensible near-term goal is a resilient human-agent workflow that exposes provenance, uncertainty, failures, guideline versions, and escalation decisions. Prospective comparison with current clinical practice is required before broader implementation claims are justified.

XSL•FO
RenderX

Acknowledgments

The authors thank colleagues and collaborators for valuable discussions and conceptual input and acknowledge institutional research platforms that facilitated literature retrieval and manuscript preparation. No other individuals or organizations contributed in a manner that met authorship criteria.

ChatGPT (version 5.6; OpenAI) was used during revision to support language polishing. The authors reviewed and approved all AI-assisted revisions and take full responsibility for the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (grant 72604378), Kunming Medical University Clinical Research Projects (grant MR-KYLY-6), Yunnan Fundamental Research Projects (grant 202501AU070024), Yunnan Provincial Department of Education Science Research Fund Project (grant 2025J0166), Yunnan Provincial Clinical Medical Center for Blood Disease and Thrombosis Prevention and Treatment (grant 2024YNLCYXZX0253), Project of Yunnan Province Clinical Research Center for Hematologic Disease (grant 202505AJ310003), Yunnan Province Major Difficult Diseases Chinese and Western Clinical Cooperation Pilot Project—Leukemia, and the Major Collaborative Project on Clinical Treatment of Critical Illnesses by Traditional Chinese and Western Medicine.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

Conceptualization: LY, LS, XY, ZL
Critical revision of the manuscript: ZL, YW
Funding acquisition: LY, YW
Literature review: XY, RZ
Project administration: YW
Study design: LY
Supervision: YW
Visualization: LS
Writing—original draft: LY
Writing—review and editing: LS, XY, RZ, SF
All authors contributed to the intellectual development and final approval of the manuscript.

Conflicts of Interest

None declared.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263 [FREE Full text] [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Hanna TP, King WD, Thibodeau S, Jalink M, Paulin GA, Harvey-Jones E, et al. Mortality due to cancer treatment delay: systematic review and meta-analysis. BMJ 2020 Nov 04;371:m4087 [FREE Full text] [doi: 10.1136/bmj.m4087] [Medline: 33148535]
3. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. Nat Rev Clin Oncol 2018 Feb;15(2):81-94. [doi: 10.1038/nrclinonc.2017.166] [Medline: 29115304]
4. Lotter W, Hassett MJ, Schultz N, Kehl KL, Van Allen EM, Cerami E. Artificial intelligence in oncology: current landscape, challenges, and future directions. Cancer Discov 2024 May 01;14(5):711-726. [doi: 10.1158/2159-8290.CD-23-1199] [Medline: 38597966]
5. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019 Jan;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
6. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. Nat Rev Clin Oncol 2019 Nov;16(11):703-715 [FREE Full text] [doi: 10.1038/s41571-019-0252-y] [Medline: 31399699]
7. Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, et al. Artificial intelligence for multimodal data integration in oncology. Cancer Cell 2022 Oct 10;40(10):1095-1110 [FREE Full text] [doi: 10.1016/j.ccell.2022.09.012] [Medline: 36220072]

XSL•FO
RenderX

8. Bhinder B, Gilvary C, Madhukar NS, Elemento O. Artificial intelligence in cancer research and precision medicine. Cancer Discov 2021 Apr;11(4):900-915 [FREE Full text] [doi: 10.1158/2159-8290.CD-21-0090] [Medline: 33811123]
9. Shah NH, Entwistle D, Pfeffer MA. Creation and adoption of large language models in medicine. JAMA 2023 Sep 05;330(9):866-869. [doi: 10.1001/jama.2023.14217] [Medline: 37548965]
10. Huang J, Xu Y, Wang Q, Wang QC, Liang X, Wang F, et al. Foundation models and intelligent decision-making: progress, challenges, and perspectives. Innovation (Camb) 2025 May 12;6(6):100948 [FREE Full text] [doi: 10.1016/j.xinn.2025.100948] [Medline: 40528892]
11. Ferber D, El Nahhas OS, Wölflein G, Wiest IC, Clusmann J, Leßmann ME, et al. Development and validation of an autonomous artificial intelligence agent for clinical decision-making in oncology. Nat Cancer 2025 Aug;6(8):1337-1349. [doi: 10.1038/s43018-025-00991-6] [Medline: 40481323]
12. Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ 2025 Feb 05;388:e081554 [FREE Full text] [doi: 10.1136/bmj-2024-081554] [Medline: 39909534]
13. Chang JS, Kim H, Baek ES, Choi JE, Lim JS, Kim JS, et al. Continuous multimodal data supply chain and expandable clinical decision support for oncology. NPJ Digit Med 2025 Feb 27;8(1):128 [FREE Full text] [doi: 10.1038/s41746-025-01508-2] [Medline: 40016534]
14. Rehburg F, Graefe A, Hübner M, Thun S. How interoperability can enable artificial intelligence in clinical applications. Stud Health Technol Inform 2024 Aug 22;316:596-600. [doi: 10.3233/SHTI240485] [Medline: 39176813]
15. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med 2019 Jan;25(1):24-29. [doi: 10.1038/s41591-018-0316-z] [Medline: 30617335]
16. Manz CR, Schriver E, Ferrell WJ, Williamson J, Wakim J, Khan N, et al. Association of remote patient-reported outcomes and step counts with hospitalization or death among patients with advanced cancer undergoing chemotherapy: secondary analysis of the PROStep randomized trial. J Med Internet Res 2024 May 17;26:e51059 [FREE Full text] [doi: 10.2196/51059] [Medline: 38758583]
17. Chen RJ, Ding T, Lu MY, Williamson DF, Jaume G, Song AH, et al. Towards a general-purpose foundation model for computational pathology. Nat Med 2024 Mar;30(3):850-862. [doi: 10.1038/s41591-024-02857-3] [Medline: 38504018]
18. Xu H, Usuyama N, Bagga J, Zhang S, Rao R, Naumann T, et al. A whole-slide foundation model for digital pathology from real-world data. Nature 2024 Jun;630(8015):181-188 [FREE Full text] [doi: 10.1038/s41586-024-07441-w] [Medline: 38778098]
19. Azizi S, Culp L, Freyberg J, Mustafa B, Baur S, Kornblith S, et al. Robust and data-efficient generalization of self-supervised machine learning for diagnostic imaging. Nat Biomed Eng 2023 Jun;7(6):756-779. [doi: 10.1038/s41551-023-01049-7] [Medline: 37291435]
20. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. Int J Comput Vis 2020 Oct 11;128(2):336-359. [doi: 10.1007/s11263-019-01228-7]
21. Omidi J. Integrative machine learning-driven prioritization of ceRNA networks in adrenocortical carcinoma. Intell Oncol 2026 Apr;2(2):100056. [doi: 10.1016/j.intonc.2026.100056]
22. Keyl J, Keyl P, Montavon G, Hosch R, Brehmer A, Mochmann L, et al. Decoding pan-cancer treatment outcomes using multimodal real-world data and explainable artificial intelligence. Nat Cancer 2025 Feb;6(2):307-322. [doi: 10.1038/s43018-024-00891-1] [Medline: 39885364]
23. Ismayilov R, Aktas A, Gencoglu EA, Oguz A, Altundag O, Akcali Z. Staging prostate cancer with AI: a comparative study of large language models and expert interpretation on PSMA PET-CT reports. Mol Imaging Biol 2026 Feb;28(1):93-105. [doi: 10.1007/s11307-025-02072-7] [Medline: 41350971]
24. Chen D, Huang RS, Jomy J, Wong P, Yan M, Croke J, et al. Performance of multimodal artificial intelligence chatbots evaluated on clinical oncology cases. JAMA Netw Open 2024 Oct 01;7(10):e2437711 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.37711] [Medline: 39441598]
25. Prinster D, Mahmood A, Saria S, Jeudy J, Lin CT, Yi PH, et al. Care to explain? AI explanation types differentially impact chest radiograph diagnostic performance and physician trust in AI. Radiology 2024 Nov;313(2):e233261. [doi: 10.1148/radiol.233261] [Medline: 39560483]
26. Omar M, Sorin V, Collins JD, Reich D, Freeman R, Gavin N, et al. Multi-model assurance analysis showing large language models are highly vulnerable to adversarial hallucination attacks during clinical decision support. Commun Med (Lond) 2025 Aug 02;5(1):330 [FREE Full text] [doi: 10.1038/s43856-025-01021-3] [Medline: 40753316]
27. Geukes Foppen RJ, Morkūnas M, Traverso A, Dienstmann R. AI assistance in tumor multidisciplinary teams. ESMO Real World Data Digit Oncol 2026 Feb 12;11:100684 [FREE Full text] [doi: 10.1016/j.esmorw.2026.100684] [Medline: 41930313]
28. Carl N, Hetz MJ, Wies C, Haggenmüller S, Winterstein JT, Mangold MH, et al. Enhancing clinicians' trust in large language models via transparent source attribution: a randomized controlled evaluation in uro-oncology. Eur J Cancer 2026 Jan 17;233:116168 [FREE Full text] [doi: 10.1016/j.ejca.2025.116168] [Medline: 41401634]
29. Gerke S, Simon DA, Roman BR. Liability risks of ambient clinical workflows with artificial intelligence for clinicians, hospitals, and manufacturers. JCO Oncol Pract 2026 Mar;22(3):357-361. [doi: 10.1200/OP-24-01060] [Medline: 40749149]

XSL•FO
RenderX

30. Hamamoto R, Koyama T, Takahashi S, Yasuda T, Kobayashi K, Akagi Y, et al. Implementing generative artificial intelligence in precision oncology: safety, governance, and significance. J Hematol Oncol 2026 Feb 09;19(1):14 [FREE Full text] [doi: 10.1186/s13045-026-01781-y] [Medline: 41664164]

Abbreviations

AJCC: American Joint Committee on Cancer
CHAARTED: Chemohormonal Therapy Versus Androgen Ablation Randomized Trial for Extensive Disease
DICOM: Digital Imaging and Communications in Medicine
EMR: electronic medical record
ETL: extract-transform-load
GLOBOCAN: Global Cancer Observatory
GTEx: Genotype-Tissue Expression
HIS: hospital information system
HL7 FHIR: Health Level Seven Fast Healthcare Interoperability Resources
LIS: laboratory information system
LLM: large language model
MDT: multidisciplinary team
NLP: natural language processing
PACS: picture archiving and communication system
PSMA PET-CT: prostate-specific membrane antigen positron emission tomography–computed tomography
RAG: retrieval-augmented generation
TCGA-ACC: The Cancer Genome Atlas adrenocortical carcinoma

Edited by M Balcarras; submitted 04.Jun.2026; peer-reviewed by R Ismayilov, J Omidi; comments to author 13.Aug.2026; revised version received 19.Aug.2026; accepted 20.Aug.2026; published 15.Sep.2026.

Please cite as:
Yang L, Shan L, Yao X, Zhao R, Li Z, Feng S, Wang Y
AI Agents for Multimodal Oncology Diagnosis: Toward Transparent and Traceable Clinical Decision Support
JMIR Cancer 2026;12:e103545
URL: https://cancer.jmir.org/2026/1/e103545
doi:10.2196/103545
PMID:

©Liuyang Yang, Liyu Shan, Xiangmei Yao, Renbin Zhao, Zengzheng Li, Shuai Feng, Yajie Wang. Originally published in JMIR Cancer (https://cancer.jmir.org), 15.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI-Enhanced Predictive Analytics to Optimize Tele-Oncology Implementation in Rural Settings: Scoping Review

Laiba Husain[1,2], MPH, DPhil; Megan Mullins[1], MPH, PhD; Bella Etingen[1,2], MA, PhD; Raaed Mohammed Zafar[3], BSc; Mediha Siddiqui[4], BSc; George L Jackson[1,5], MHA, PhD

[1]O'Donnell School of Public Health, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX, United States

[2]Research and Development Service, Dallas Veterans Affairs Medical Center, Dallas, TX, United States

[3]School of Behavioral and Brain Sciences, UT Dallas, Dallas, TX, United States

[4]Dedman College of Humanities and Sciences, Southern Methodist University, Dallas, TX, United States

[5]Center of Innovation to Accelerate Discovery and Practice Transformation (ADAPT), Durham VA Health Care System, Durham, NS, United States

Corresponding Author:
Laiba Husain, MPH, DPhil
O'Donnell School of Public Health, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX, United States

Abstract

Background: Tele-oncology addresses geographic barriers to cancer care, but implementation challenges persist in rural settings. AI-enhanced predictive analytics offer opportunities for optimizing deployment through personalized, data-driven strategies; however, evidence in rural tele-oncology contexts remains limited, and critical equity considerations remain underexamined.

Objective: This scoping review aimed to map evidence on AI-enhanced predictive analytics in tele-oncology implementation, with particular attention to rural and underserved populations, to identify research gaps and inform implementation science priorities.

Methods: We searched 5 databases (PubMed, Embase, CINAHL, Web of Science, and IEEE Xplore) using 4 concept domains (tele-oncology, rural implementation barriers, AI or predictive analytics, implementation science) from January 2015 through November 2025. Two independent reviewers screened 330 unique records (title or abstract; Cohen κ=0.78), with the principal investigator resolving conflicts. Of 138 full-text reviews (κ=0.82), 4 studies met inclusion criteria. Data extraction captured study characteristics, AI applications, implementation factors, and outcomes. We used narrative thematic analysis to map findings into three themes: (1) the current tele-oncology implementation landscape in rural and underserved settings, (2) potential AI applications addressing implementation challenges, and (3) implementation considerations for AI systems themselves.

Results: Four included studies (1 pilot feasibility study, 1 proof-of-concept validation study, 1 cross-sectional predictive study, and 1 platform development study; published 2019 - 2025) demonstrated limited evidence at the intersection of AI, tele-oncology, and rural health equity. Patient characteristics predicted telehealth modality preferences with 86.2% accuracy, revealing that male patients exhibited 66% increased odds of video selection versus female patients (*P*=.004), and urban residents showed 101% increased odds compared to rural counterparts (*P*<.001). Liu et al demonstrated that disadvantaged populations engaged with AI-generated health literacy content 2.52-fold more frequently than nondisadvantaged counterparts. However, all 4 studies documented substantial implementation barriers (patient, provider, organizational, and system levels) persisting despite technological sophistication. Organizational threshold effects, where remote monitoring interventions succeeded with adequate provider capacity but failed under resource constraints—suggest that algorithmic innovations cannot overcome structural limitations in rural facilities. No studies explicitly examined algorithmic bias, cross-population validation, or potential harms in rural contexts. Geographic concentration in high-resource countries (United States n=2, Greece n=1, and Singapore n=1) and limited oncology-specific focus underscore structural gaps in knowledge generation for underserved populations.

Conclusions: Current evidence remains insufficient to support definitive practice recommendations. The observed evidence gap may reflect broader structural inequities in knowledge generation: populations with the greatest implementation challenges appear to remain substantially underrepresented in AI and digital health literature. Future research should prioritize comparative effectiveness studies in authentic rural contexts with implementation science outcomes, equity-centered cross-population validation, specification of translation mechanisms linking AI predictions to implementation strategies, health economic analyses, and mechanistic research on sociotechnical integration factors, ensuring technological innovation reduces rather than perpetuates disparities in cancer care.

(JMIR Cancer 2026;12:e78005) doi:10.2196/78005

XSL•FO
RenderX

KEYWORDS

artificial intelligence; predictive analytics; tele-oncology; telemedicine; rural health; implementation science; health equity; scoping review

Introduction

Terminology and Key Definitions

To ensure clarity and consistency throughout this review, we define key terms as follows: we use telehealth as the broader umbrella term encompassing all remote health care delivery using telecommunications technology, including both clinical services (telemedicine) and nonclinical services (tele-education, administrative videoconferencing) [1]. Telemedicine refers specifically to remote clinical services involving diagnosis, treatment, and patient monitoring [1,2]. Tele-oncology denotes the application of telehealth or telemedicine specifically to cancer care delivery, encompassing video consultations for diagnosis and treatment planning, remote symptom monitoring, virtual tumor boards, telephonic follow-up care, and hub-and-spoke models connecting rural facilities with specialized cancer centers [3,4]. AI-enhanced predictive analytics encompasses computational approaches, including supervised machine learning (random forests and gradient boosting), natural language processing, ensemble methods, and deep learning techniques that analyze data to forecast future outcomes, identify patterns, or generate predictions to inform clinical or operational decision-making [5]. Rural is defined broadly to include populations in geographically isolated areas with limited health care access, consistent with heterogeneous definitions across included studies, recognizing that rural classifications vary by country and health care system [6]. Implementation refers to the process of translating evidence-based interventions into routine practice, encompassing activities related to adoption, integration, sustainability, and scale-up of tele-oncology services within health care delivery systems [7].

Tele-oncology has emerged as a critical strategy for expanding access to cancer care, particularly for individuals facing geographic, socioeconomic, or infrastructure-related barriers to traditional in-person services [8,9]. Following rapid implementation during the COVID-19 pandemic [10], tele-oncology has demonstrated feasibility across varied settings and patient populations [4,8,10]. However, substantial implementation challenges persist, including gaps in technological literacy among patients, limitations in broadband infrastructure in rural areas, complexities in workflow integration, barriers to device access, and constraints on health care system resources [11,12].

These implementation barriers disproportionately affect rural and underserved populations, potentially exacerbating existing cancer care disparities [13,14]. Patients with cancer living in rural areas experience lower survival rates and later-stage diagnoses compared to their urban counterparts [15], driven partly by limited access to oncology services, multidisciplinary care teams, and clinical trials [16]. While tele-oncology theoretically addresses geographic access barriers, populations with the greatest potential to benefit from telehealth often demonstrate the lowest usage rates [17]. This implementation gap reflects complex, multilevel barriers spanning individual patient factors (technological literacy, device access, and digital confidence), health care system characteristics (infrastructure capacity, workflow integration, and provider comfort with telehealth), and contextual elements (broadband availability, reimbursement policies, and organizational support) [1,14,18].

Recent advancements in artificial intelligence and predictive analytics present opportunities for addressing these implementation challenges through more personalized, data-driven approaches to tele-oncology deployment [19]. Emerging evidence from adjacent health care domains demonstrates that AI-enhanced systems can accurately predict the preferences of patients for different telehealth modalities based on demographic and clinical characteristics [20], identify individuals at high risk for implementation barriers before they manifest clinically [21], optimize resource allocation by forecasting demand and prioritizing high-need populations [22], and enable remote symptom monitoring with predictive capabilities for early complication detection [23]. These applications suggest potential translational pathways for enhancing tele-oncology implementation, though direct evidence specifically addressing AI applications in rural tele-oncology contexts remains limited.

The integration of AI-enhanced predictive analytics with implementation science frameworks offers a promising approach for optimizing tele-oncology delivery [19,24]. Implementation science provides systematic methods for translating evidence-based interventions into routine practice, with established frameworks for assessing barriers, designing implementation strategies, and evaluating outcomes [7,25]. Recent theoretical work has articulated how AI applications could specifically address implementation science challenges related to speed, sustainability, equity, generalizability, contextual assessment, and causal inference [18,19]. However, the successful translation of AI technologies into clinical practice requires careful attention to implementation considerations, including stakeholder engagement, workflow integration, technical infrastructure requirements, validation across diverse populations, and proactive monitoring for unintended consequences [24].

This scoping review systematically examines current applications of AI-enhanced predictive analytics and digital health technologies across oncology care delivery, with particular attention to implementation considerations relevant for rural and underserved populations. Given the nascent and heterogeneous nature of research at this intersection, we used a scoping review methodology to map existing evidence, identify research gaps, and synthesize findings to inform a suggested research agenda [26,27]. Our specific objectives were to: (1) characterize current applications of AI and predictive analytics in telehealth and digital oncology care; (2) examine empirical evidence on implementation factors, barriers, and facilitators;

XSL•FO
RenderX

and (3) identify critical research gaps and future directions for optimizing tele-oncology implementation in rural contexts.

Methods

This scoping review followed established methodological frameworks [26] and adhered to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines (Checklist 1) [28] and PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension) guidelines (Checklist 2) [29] reporting guidelines. Complete adherence documentation is provided in (PRISMA-ScR checklist) and Multimedia Appendix 1 (full search strategy). No published protocol was developed, consistent with exploratory scoping review methodology. The review was not prospectively registered.

Studies were included if they addressed: (1) tele-oncology implementation across any setting, (2) rural health care barriers or implementation strategies relevant to underserved populations, and (3) AI or machine learning applications in health care delivery with transferable insights for oncology. Consistent with Joanna Briggs Institute methodological standards for scoping reviews, which are designed to synthesize primary empirical evidence, systematic reviews, and scoping reviews were excluded from the eligible study designs [26,27]. Where relevant secondary reviews were identified during screening, specifically Aziz et al [30] and Anderson et al [31], these were retained and referenced in the “Discussion” section to contextualize primary study findings within the broader evidence landscape, consistent with accepted practice in scoping review reporting.

Publications were restricted to English-language, peer-reviewed literature from January 2015 through November 2025, capturing recent AI developments and postpandemic tele-oncology expansion. Complete inclusion or exclusion criteria, along with operational definitions, examples, and rationale, are detailed in Multimedia Appendix 2.

We searched 5 databases selected for disciplinary breadth: PubMed or MEDLINE (biomedical literature), Embase (pharmaceutical and European coverage), CINAHL (nursing and allied health), Web of Science (multidisciplinary sciences and implementation research), and IEEE Xplore (computer science and AI technical literature). The search strategy, developed in consultation with a health sciences librarian, integrated 4 concept domains using controlled vocabulary (MeSH, Emtree, and CINAHL headings) and free-text keywords: (1) tele-oncology or telehealth in cancer care, (2) rural health care and implementation barriers, (3) AI and predictive analytics, and (4) implementation science. These domains were combined using Boolean logic to optimize sensitivity while maintaining specificity.

An initial search was completed on April 9, 2025. Following reviewer feedback indicating limited evidence, the search strategy was comprehensively revised and rerun from inception (January 2015) through November 2025, using broader search terms across all 5 databases (Multimedia Appendix 1). Supplemental strategies included backward citation chasing of included articles and forward citation tracking of seminal papers.

Following duplicate removal via Covidence systematic review software (Veritas Health Innovation), 330 unique records underwent title or abstract screening by 2 independent reviewers (RZ and MS). The principal investigator (PI) (LH) served as the arbiter for conflicts. Calibration exercises on pilot samples demonstrated substantial interrater agreement (Cohen κ=0.78 for title or abstract screening, κ=0.82 for full-text review). Of the 330 records screened, 138 proceeded to full-text review. The same 2 independent reviewers conducted the full-text assessment and data extraction, with the PI resolving any disagreements. Full-text assessment excluded 130 studies primarily for the following reasons: clinical outcomes focus without implementation considerations (n=62, 47.7%), telehealth outside oncology without transferability (n=34, 26.2%), insufficient empirical or theoretical contribution (n=24, 18.5%), absence of rural/barrier focus (n=18, 13.8%), no AI or predictive analytics application (n=12, 9.2%), or inadequate methodological detail (n=8, 6.2%). Four studies met all inclusion criteria. Detailed screening procedures, interrater agreement calculations, and reasons for key exclusions are documented in Multimedia Appendix 2.

Data extraction used a standardized form to capture: study characteristics (design, setting, and population), tele-oncology implementation details (modality, clinical application), AI or predictive analytics applications (techniques, prediction targets, and performance metrics), implementation factors (barriers, facilitators, and strategies), and outcomes (implementation, clinical, and process measures). The extraction form, which was developed iteratively through pilot testing and team discussions, is provided in Multimedia Appendix 3.

Synthesis used narrative thematic analysis, which was appropriate for heterogeneous evidence that precluded meta-analysis. Following a modified framework synthesis, we organized findings into 3 prespecified themes aligned with the review objectives: (1) the current tele-oncology implementation landscape in rural and underserved settings, (2) potential AI-powered predictive analytics applications addressing implementation challenges, and (3) implementation considerations for AI systems themselves. Two independent reviewers (RZ and MS) coded the extracted data, with the principal investigator resolving disagreements and developing subthemes inductively through iterative discussion and consensus. Formal quality appraisal was not conducted, consistent with scoping review methodology [27]. As articulated in foundational frameworks, scoping reviews prioritize evidence mapping and gap identification over intervention effectiveness synthesis, which is the purview of systematic reviews requiring quality assessment.

Results

The systematic search yielded 330 unique records; 138 underwent full-text review, resulting in 4 included studies (Figure 1). This limited evidence base reflects the nascent intersection of AI-powered predictive analytics, tele-oncology implementation, and rural health equity—domains that have

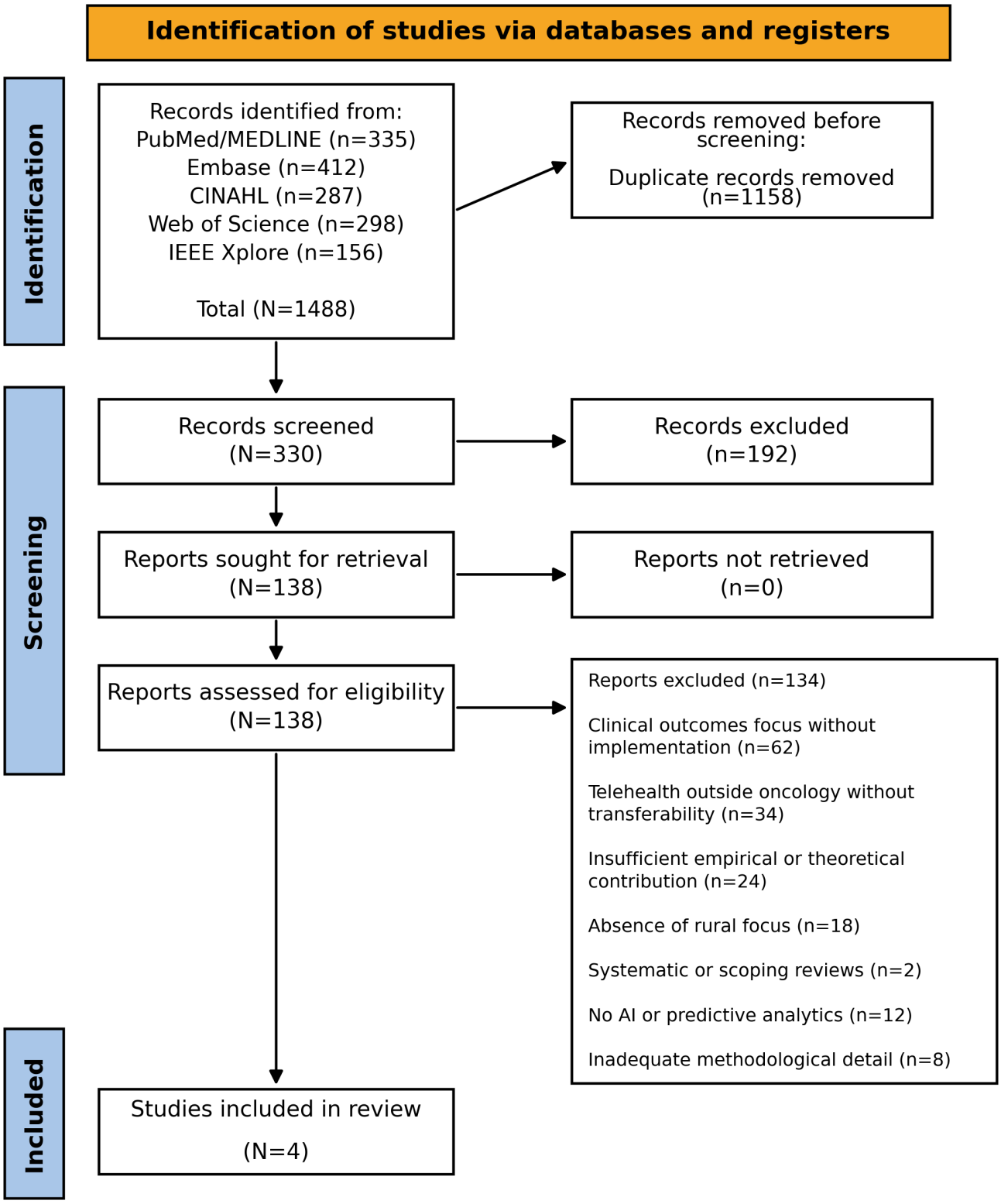
XSL•FO
RenderX

evolved largely in parallel rather than through integrated inquiry. The included studies comprised 2 pilot feasibility studies [32,33], 1 cross-sectional predictive study [34], and 1 platform development study [35], published between 2019 and 2025. Geographic concentration in high-resource settings (United States n=2, Greece n=1, and Singapore n=1) and limited oncology-specific focus underscore structural gaps in evidence generation for underserved populations.

Table 1 presents study characteristics, including methodological approaches, populations, AI applications, and key limitations. Table 2 synthesizes implementation barriers across multilevel ecological domains. Table 3 maps AI-powered predictive analytics applications to specific tele-oncology implementation challenges, articulating translational pathways from current evidence to future research priorities. Figure 2 presents an evidence gap map illustrating the distribution of current evidence across AI application domains, implementation science dimensions, and equity and validation considerations.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. AI: artificial intelligence; IEEE: Institute of Electrical and Electronics Engineers.



XSL•FO
RenderX

Table . Characteristics of included studies (N=4).

Study	Design	Setting or population	AI[a] or analytics application	Key implementation findings	Limitations
Daly et al [32], 2020	Single-arm pilot feasibility study	Academic cancer center regional site; n=100 high-risk patients initiating antineoplastic therapy	ML[b] risk stratification model identifying high-risk patients for preventable acute care; daily PRO[c] symptom surveillance with alert generation	29% enrollment rate; 56% daily response rate sustained 6 months; 93% generated severe alerts; 4 patient archetypes identified; 22% vs 39% acute care use (enrolled vs nonenrolled); provider volume threshold effects	No control group; single-site; resource-intensive (3 FTE[d] for 100 patients); cost-effectiveness unknown
Mammas et al [33], 2022	Proof-of-concept validation study	Peripheral Greece (rural); n=40 patients with breast cancer (32 malignant, 8 benign)	AI-assisted remote multidisciplinary tumor boards integrating teleradiology, telepathology, telecytology with algorithmic interpretation	Mean accuracy: 94.1% surgical planning (95% CI: 85% - 99%), 96.8% chemotherapy, 96.7% hormonal therapy; addresses rural subspecialty access barrier	Small sample; breast cancer only; single health care system; generalizability unclear
Khairat et al [34], 2019	Cross-sectional predictive analytics	Academic medical center; n=1403 virtual urgent care encounters; 44% rural	Multinomial logistic regression predicting patient preferences for mHealth vs telemedicine based on demographics and chief concern	Sex (P=.004) and geographic setting (P<.001) significantly predicted modality choice; 86.2% prediction accuracy; rural patients 50% less likely to choose video; high satisfaction both modalities (91.1%, 89%)	Single-site; limited demographic diversity; selection bias (self-selected virtual care users)
Liu et al [35], 2025	Platform development and pilot study	This block should read single-site US (Emory University); n=40 (20 colorectal cancer patients + 20 caregivers); 55% disadvantaged (income ≤250% Federal Poverty Level, Medicaid, or uninsured)	Generative AI (ChatGPT, Pictory) creating accessible health literacy content at 6th-grade reading level with multimedia	8-week pilot: disadvantaged participants logged in 2.52 × more than nondisadvantaged; 3.98/5 satisfaction; addresses health literacy barrier disproportionately affecting rural or underserved	Short duration; small sample; single cancer type; generative AI accuracy or reliability concerns

[a]AI: artificial intelligence.

[b]ML: machine learning.

[c]PRO: patient-reported outcomes.

[d]FTE: full-time equivalent.

XSL•FO
RenderX

Table . Multilevel implementation barriers identified across the included studies (N=4).

Level	Barrier category	Specific barriers (studies reporting)
Patient	Technological access	Device ownership gaps [33,34]; broadband connectivity limitations, especially in rural areas [33,34]; inadequate infrastructure supporting telehealth access [33]
Patient	Digital literacy	Technological skill deficits [32,34,35]; low digital confidence or self-efficacy [33-35]; platform navigation difficulties [32,34]
Patient	Health literacy	Inability to interpret health information without support [35]; limited understanding of symptom significance [32,35]; educational barriers limiting self-management [35]
Patient	Psychosocial	Cancer-related distress limiting engagement [32]; preference for in-person care [33]; privacy/security concerns about data sharing [34]
Provider	Workflow integration	EMR[b] integration complexities [32,33]; disrupted clinical workflows [32,34]; time requirements for reviewing AI-generated outputs [32,34]
Provider	Algorithmic trust	Algorithmic opacity hindering understanding [34]; uncertainty about appropriate reliance on predictive recommendations [32,34]
Provider	Clinical concerns	Medicolegal uncertainty regarding AI-supported care decisions [32]; inability to perform physical examination remotely [33]; care fragmentation concerns [32]
Provider	Training or support	Insufficient AI literacy among clinicians [32]; inadequate technical support infrastructure [32,34]; limited training on platform use [34]
Organizational	Technical infrastructure	EMR interoperability challenges [32-34]; data standardization deficits [34]; cybersecurity requirements [33]
Organizational	Human resources	Dedicated monitoring team requirements [32]; technical support staffing needs [32,34]; competing clinical priorities limiting engagement [32]
Organizational	Financial	High startup costs [32,34]; uncertain return on investment [32,34]; reimbursement ambiguities [34]; ongoing maintenance costs [32](
Organizational	Organizational culture	Resistance to change in clinical practice [32]; inadequate leadership support for tele-oncology implementation [32,34]; misaligned organizational incentives [34]
System	Infrastructure	Rural broadband inadequacies [33,34]; digital divide persistence in underserved regions [33,35]; geographic infrastructure limitations [33]
System	Regulatory or policy	Licensure barriers across state/national borders [32,34]; reimbursement policy limitations constraining sustainability [34]; data privacy or HIPAA[c] compliance requirements [32]
System	Validation or standards	Lack of standardized validation frameworks for AI[a] systems [32-34]; algorithmic bias in training data [34,35]

[a]AI: artificial intelligence.

[b]EMR: electronic medical record.

[c]HIPAA: Health Insurance Portability and Accountability Act.

XSL•FO
RenderX

Table . Translational pathways: mapping AI-powered predictive analytics to tele-oncology implementation challenges.

Implementation challenge	AI[a] or predictive analytics application	Current evidence base	Research priorities
Low telehealth used among rural or underserved populations despite geographic barriers	Predictive models identifying patient subgroups likely to engage with specific telehealth modalities; personalized platform matching based on demographic and clinical characteristics; algorithmically informed outreach strategies	Khairat et al [34]: 86.2% accuracy predicting mHealth versus telemedicine preference based on sex, geographic setting, and chief concern; urban residents showed 101% increased odds of video selection vs rural (P<.001)	Validate modality prediction models in rural tele-oncology settings; develop actionable implementation protocols translating predictions into targeted outreach; examine sustainability of personalized matching approaches
Preventable acute care use during cancer treatment due to undetected symptom escalation	Risk stratification models identifying high-risk patients; real-time symptom surveillance with predictive escalation alerts; machine learning identifying patients at risk for acute care events	Daly et al [32]: ML[b] risk stratification model targeting high-risk quartile; 93% of enrolled patients generated severe symptom alerts; preliminary data suggested 22% acute care usage vs 39% in nonenrolled high-risk controls	Conduct RCTs[c] comparing AI-enhanced vs standard symptom monitoring; optimize alert threshold sensitivity or specificity; determine cost-effectiveness in rural cancer centers; validate across cancer types; examine long-term sustainability
Lack of subspecialty oncology expertise in rural facilities for multidisciplinary treatment planning	AI-assisted virtual tumor boards integrating imaging or pathology interpretation with algorithmic decision support; remote access to tumor board expertise via AI-augmented consultations	Mammas et al [33]: AI-assisted remote tumor boards achieved 94.1% accuracy for surgical planning (95% CI: 85% - 99%), 96.8% for chemotherapy recommendations, and 96.7% for hormonal therapy decisions in breast cancer cases (n=40)	Expand validation to multiple cancer types beyond breast cancer; validate across diverse health care systems and geographic regions; examine implementation feasibility in US rural cancer centers with limited technical infrastructure
Health literacy barriers limiting self-management capacity, disproportionately affecting rural/underserved populations	Generative AI creating personalized, accessible educational content tailored to individual literacy level, language preferences, and learning modality; AI-generated multimedia for symptom management and treatment adherence	Liu et al [35]: Generative AI platform creating 6th-grade reading level content with multimedia enhancement; disadvantaged participants (income <$50,000 or education ≤ high school) demonstrated 2.52-fold higher engagement compared to nondisadvantaged counterparts	Validate generative AI accuracy and reliability in oncology contexts; test across cancer types and treatment modalities; examine long-term outcomes on self-management and treatment adherence; assess equity in content generation
Inadequate resources for proactive identification and outreach to patients experiencing implementation barriers	Predictive models identifying patients at elevated risk for technology access barriers, digital literacy deficits, and other implementation obstacles using SDOH[d] data; automated risk flagging enabling targeted interventions	Theoretical frameworks emphasized [32,34] but no empirical validation in tele-oncology contexts	Develop and validate barrier prediction models integrating SDOH data; design targeted intervention protocols addressing identified barriers; test implementation in rural cancer centers; examine cost-effectiveness
Workflow integration challenges and provider resistance limiting adoption of tele-oncology platforms	AI-assisted clinical decision support reducing cognitive burden; automated clinical documentation generating summaries from encounters; intelligent prioritization of high-risk patients; systems learning from provider feedback to improve usability	Frameworks emphasize sociotechnical integration [32,34] with emphasis on perceived usefulness and ease of use, but limited empirical testing of specific AI features addressing workflow pain points	Develop AI tools targeting specific workflow pain points identified through provider ethnography; conduct rigorous usability testing; measure time savings and clinician satisfaction; examine adoption and fidelity in rural settings
Algorithmic bias in AI systems potentially exacerbating existing health disparities in rural cancer care	Fairness-aware machine learning approaches; bias detection algorithms monitoring performance across demographic groups; continuous equity monitoring of disparate impact; diverse training data representation; transparent performance reporting disaggregated by population	Theoretical emphasis on equity monitoring [32,34] but minimal empirical application in tele-oncology	Develop equity-focused validation frameworks requiring intersectional analyses; mandate diverse, representative training datasets; implement real-time disparities monitoring; publish disaggregated performance metrics by race, ethnicity, SES[f], geographic location; establish accountability mechanisms

[a]AI: artificial intelligence.

[b]ML: machine learning.

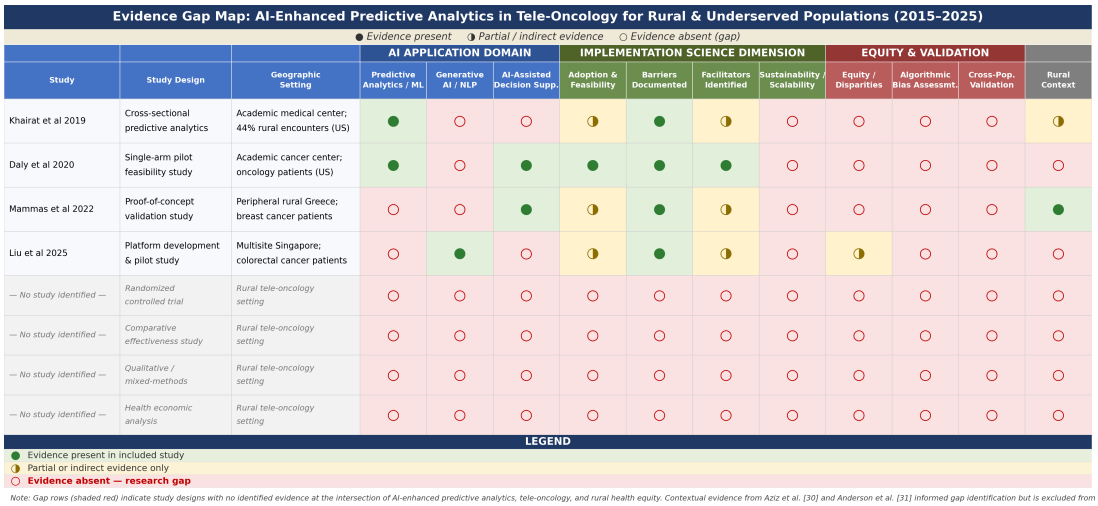
[c]RCT: randomized controlled trial.

XSL•FO
RenderX

[d]SDOH: Social Determinants of Health.

[e]SES: socioeconomic status.

Figure 2. Evidence gap map [32-35]. ML: machine learning; NLP: natural language processing.



Khairat et al [34] demonstrated that patient characteristics predict telehealth modality preferences with sufficient accuracy (86.2%) to inform personalized implementation strategies. Analyzing 1403 virtual care encounters, multinomial logistic regression revealed sex (P=.004) and geographic setting (P<.001) as primary determinants: male patients exhibited 66% increased odds of video telemedicine selection versus female patients, while urban residents showed 101% increased odds compared to their rural counterparts. Chief medical concern independently predicted modality choice (P<.001), with privacy-sensitive conditions (eg, urinary tract infections and gynecological complaints) associating with telephone preference and visual assessments (eg, dermatological concerns and conjunctivitis) predicting video selection. High satisfaction persisted across modalities (91.1% mobile health and 89% telemedicine), suggesting that offering choice and algorithmically matching patients to appropriate platforms enhances acceptability without compromising outcomes.

These findings illuminate how predictive analytics can address the implementation paradox, wherein individuals experiencing the greatest geographic barriers demonstrate the lowest telehealth usage. By identifying patient subgroups likely to engage with specific modalities, health care systems can proactively tailor outreach, technology support, and platform recommendations—moving beyond one-size-fits-all implementation toward precision approaches responsive to intersecting social determinants.

Three studies examined AI applications in oncology remote monitoring, revealing feasibility alongside persistent implementation challenges. Daly et al [32] integrated machine learning risk stratification with digital symptom surveillance for 100 high-risk patients initiating antineoplastic therapy. The predictive model prospectively identified patients at elevated risk for preventable acute care usage, enabling targeted enrollment. Daily patient-reported outcome assessments achieved a 56% sustained response rate over 6 months, with 93% of enrolled patients generating severe symptom alerts requiring clinical intervention. Qualitative analysis identified 4 patient engagement archetypes—Promise Keeper, Data Tracker, Unengaged, and Overwhelmed—each characterized by distinct motivational drivers and implementation needs, underscoring heterogeneity in technology acceptance even among self-selected participants.

Preliminary acute care data suggested a potential benefit (22% usage among enrolled patients vs 39% among nonenrolled high-risk controls), though the absence of a randomized design precludes causal inference. Critical implementation barriers emerged: providers with low patient volumes perceived information overload and care fragmentation, while those with adequate panel penetration valued the intervention, suggesting threshold effects for workflow integration. Organizational requirements—dedicated nursing team (2 registered nurses and 1 nurse practitioner for 100 patients), custom electronic medical record (EMR) integration, and 12-hour daily coverage—raise scalability concerns for rural facilities with constrained resources.

Mammas et al [33] demonstrated the technical feasibility of AI-assisted remote multidisciplinary tumor boards for breast cancer in peripheral Greece, achieving mean accuracy of 94.1% for surgical planning, 96.8% for chemotherapy, and 96.7% for hormonal therapy (95% CI 85%–99% for all 3 estimates), across 40 cases. This model addresses a fundamental rural implementation barrier—the absence of local subspecialty expertise—by leveraging AI to augment telehealth consultations. However, generalizability beyond breast cancer and validation in diverse health care contexts remain unexplored.

Liu et al [35] used generative AI to create accessible health literacy content for colorectal cancer self-management, targeting disadvantaged populations at a single US academic medical center. The platform automatically generated 6th-grade reading level materials with multimedia enhancement. Disadvantaged participants (defined as income ≤250% of the Federal Poverty Level, Medicaid enrollment, or uninsured status) demonstrated 2.52-fold higher platform engagement than nondisadvantaged counterparts, suggesting particular value for populations

XSL•FO
RenderX

experiencing compounding socioeconomic marginalization. This application illustrates how AI can address specific implementation barriers, such as inadequate health literacy, that disproportionately constrain rural and underserved community engagement with tele-oncology.

Cross-study synthesis reveals fundamental evidence gaps. No included studies directly examined AI-powered predictive analytics for identifying or mitigating tele-oncology implementation barriers in rural settings—the core focus of this review. Validation of predictive models across diverse populations, health care systems, and geographic contexts remains limited, with most studies conducted in single academic medical centers serving nonrepresentative populations. Translation of AI predictions into actionable implementation strategies has received minimal empirical examination. Cost-effectiveness analyses comparing AI-enhanced approaches versus standard approaches were notably absent. Longitudinal research examining sustainability, adaptation, and the impact on disparities beyond pilot phases would strengthen the evidence for scaling.

Table 2 synthesizes implementation barriers across patient, provider, organizational, and system levels. Patient-level barriers encompassed technological literacy deficits, device access limitations, broadband connectivity constraints, and digital confidence gaps—challenges disproportionately affecting older adults, individuals from lower socioeconomic populations, and people living in rural communities. Provider barriers included workflow integration complexities, algorithmic opacity concerns, time requirements for reviewing AI-generated outputs, and medicolegal uncertainty regarding appropriate reliance on algorithmic recommendations. Organizational barriers spanned technical infrastructure (EMR integration and data interoperability), staffing resources (dedicated monitoring teams and technical support), and financial sustainability (startup costs, ongoing maintenance, and reimbursement uncertainties). System-level barriers encompassed broadband infrastructure inadequacies in rural areas, regulatory ambiguities, data privacy requirements, and the absence of standardized validation frameworks.

Figure 3 illustrates the relative prominence of each barrier category within each included study, highlighting areas where challenges to implementation challenges were most extensively documented.

Figure 3. Barrier heat map [32-35]. AI: artificial intelligence.

Multilevel implementation barrier coverage across included studies

Shading reflects prominence of each barrier category within each study. Darker green = more extensively documented.

Barrier level	Barrier category	Khairat et al 2019	Daly et al 2020	Mammas et al 2022	Liu et al 2025
Patient	Technological access	Prominent	Mentioned	Prominent	Mentioned
	Digital literacy	Prominent	Prominent	Prominent	Prominent
	Health literacy	—	Prominent	—	**Central**
	Psychosocial / engagement	Mentioned	Prominent	Prominent	—
Provider	Workflow integration	Prominent	**Central**	Prominent	Mentioned
	Algorithmic trust / opacity	Prominent	Prominent	—	—
	Clinical / medicolegal concerns	—	Prominent	Prominent	—
	AI training and literacy	Mentioned	Prominent	—	Mentioned
Organizational	Technical infrastructure	Prominent	**Central**	Prominent	—
	Human resources	—	**Central**	—	—
	Financial / cost	Prominent	Prominent	—	—
	Organizational culture	Mentioned	Prominent	—	—
System	Rural broadband / infrastructure	Prominent	—	Prominent	—
	Regulatory / policy / licensure	Prominent	Prominent	—	—
	Algorithmic bias / validation standards	Mentioned	Mentioned	Mentioned	Mentioned

Legend — Barrier prominence

— — — Not documented in this study

Mentioned — Referenced as a barrier

Prominent — Discussed in depth as a substantive challenge

Central — Primary implementation challenge in the study

Facilitators enabling implementation success emerged consistently (Table 3): strong clinician champions actively reinforcing program value with patients, a responsive technical support infrastructure, demonstrable clinical benefits driving

XSL•FO
RenderX

stakeholder buy-in, and sustained organizational commitment beyond pilot funding. The evidence suggests that AI implementation success depends less on technological sophistication per se than on systematic attention to sociotechnical integration—aligning AI capabilities with clinical workflows, organizational contexts, and patient needs through iterative, stakeholder-engaged development and deployment processes.

Discussion

Principal Findings

This scoping review sought to (1) characterize current applications of AI and predictive analytics in telehealth and digital oncology care; (2) examine empirical evidence on implementation factors, barriers, and facilitators; and (3) identify research gaps and future directions for optimizing tele-oncology implementation in rural contexts. Across all 3 objectives, the overarching finding is one of substantial evidence scarcity. Of 330 unique records screened, only 4 primary studies [32-35] met inclusion criteria, none of which directly tested the integration of AI-enhanced predictive analytics within a rural tele-oncology implementation context. The included studies offered selective proof-of-concept across discrete AI applications (modality preference prediction, remote symptom monitoring, AI-assisted multidisciplinary tumor boards, and generative health literacy tools): yet each was conducted in a well-resourced or academic setting, with limited applicability to the rural contexts that are the focus of this review. Multilevel implementation barriers were consistently documented across all 4 studies, and no study examined algorithmic bias, cross-population validation, or equity-related harms. These findings suggest that the field remains at an early, exploratory stage, and that current evidence is insufficient to support practice recommendations or policy guidance for rural tele-oncology implementation.

Technical Feasibility and the Gap to Implementation Readiness

The included studies collectively suggest that AI-based tools may, under specific conditions, address discrete implementation challenges in oncology care delivery. Khairat et al [34] found that patient demographic and clinical characteristics were sufficiently predictive of telehealth modality preference to potentially inform personalized outreach strategies, a finding with plausible relevance for rural populations, who exhibit lower rates of video-based telehealth adoption in the broader literature [14,17]. Daly et al [32] demonstrated the operational feasibility of machine learning risk stratification integrated with digital symptom surveillance, with preliminary signals suggesting reduced acute care use among enrolled patients, though the absence of a control group precludes causal inference. Mammas et al [33] reported the technical accuracy of AI-assisted remote tumor boards for breast cancer management in rural Greece, offering a possible model for extending subspecialty access to geographically isolated facilities [15,16]. Liu et al [35] showed that generative AI can produce health literacy materials that may engage socioeconomically disadvantaged populations more effectively than standard content, though this finding was observed in a single cancer type in Singapore, and its generalizability to rural Western contexts warrants caution.

However, this provisional feasibility evidence coexists with documented implementation fragility. Daly et al [32] observed organizational threshold effects, whereby the remote monitoring intervention functioned adequately when provider panel penetration was sufficient, but generated information overload and care fragmentation concerns under lower-volume conditions—conditions more representative of rural cancer facilities [14]. The intervention's staffing requirements raise additional scalability questions for resource-constrained settings [32]. These tensions are echoed in the broader contextual literature. Aziz et al [30], synthesizing 11 studies of AI-driven remote patient monitoring in oncology, concluded that available evidence remains insufficient to establish superiority over traditional care delivery, identifying consistent barriers including technical infrastructure requirements, gaps in literacy among patients and providers, and workflow integration complexities. The fact that this conclusion emerged from studies largely conducted in well-resourced academic environments further underscores the depth of the evidence gap for rural contexts. Anderson et al [31] noted that only 26% of the telehealth predictive analytics studies they reviewed focused on predictive systems, with a notable absence of resource optimization models applicable to rural or underresourced implementation contexts. This suggests that the orientation of existing research may be poorly matched to the implementation challenges that rural settings face [18,19,31].

Persistence of Multilevel Implementation Barriers

A consistent pattern across the primary studies and contextual review literature is that implementation barriers operate simultaneously across patient, provider, organizational, and system levels, and that technological sophistication does not, in itself, resolve barriers operating at other levels [7,14,18]. At the patient level, device access limitations, broadband connectivity gaps, digital literacy deficits, and health literacy constraints appear to intersect in compounding ways, with particular salience for older adults, individuals with lower incomes, and people living in geographically isolated areas [13,14,17]. Khairat et al [34] found that people living in rural areas were significantly less likely to select video-based telehealth even when it was available, suggesting that geographic context independently shapes technology engagement beyond questions of mere access. At the provider level, algorithmic opacity (clinicians' difficulty in interpreting and appropriately acting on AI-generated outputs) emerges as a recurrent concern across studies [32,34], and is identified by Aziz et al [30] as a barrier to adoption of AI-enhanced monitoring across multiple health care systems. Reddy [24] similarly argues that translating AI tools into clinical practice requires governance structures and workflow integration processes that most current deployment efforts have not adequately addressed.

Organizational and system-level barriers may be especially consequential for rural settings. The dedicated staffing, EMR integration, and sustained technical support required by the Daly et al [32] intervention represent structural prerequisites that are unlikely to be available in many rural facilities [14,15].

XSL•FO
RenderX

Anderson et al [31] noted the absence of resource optimization models in the telehealth analytics literature, suggesting that the field has yet to systematically engage with how AI tools might be deployed sustainably under resource constraints. At the system level, rural broadband inadequacies, cross-border licensure ambiguities, and reimbursement policy limitations add further layers of complexity compounding facility-level constraints and potentially limiting the transferability of findings from academic settings [13,14,19].

Equity Considerations and Gaps in the Evidence Base

A notable gap across all 4 primary studies is the absence of equity-centered analysis. None of the included studies explicitly examined algorithmic bias, conducted cross-population validation of AI system performance across racial, socioeconomic, or geographic subgroups, or assessed potential harms associated with AI deployment in underserved populations. This gap is particularly concerning given documented evidence that AI systems trained predominantly on data from majority or well-resourced populations may perform differentially—and potentially harmfully—when applied to underrepresented communities [5,18,19]. Individuals with cancer living in rural and lower-income areas, who are already subject to compounding health care disadvantages, including later-stage diagnoses and poorer survival outcomes [15,16], may be disproportionately exposed to such risks.

Aziz et al [30] explicitly identified insufficient attention to equity impacts as a limitation of the AI-enhanced remote monitoring literature, observing that reviewed studies rarely examined whether benefits were distributed equitably across patient subgroups or whether the implementation burden fell disproportionately on already-disadvantaged populations. Anderson et al [31] similarly found that the vast majority of telehealth predictive analytics studies were conducted in single academic medical centers serving nonrepresentative populations, with minimal rural implementation focus and no systematic reporting of outcomes disaggregated by socioeconomic or demographic characteristics. Taken together, these observations suggest that populations with the greatest potential need for equitable tele-oncology implementation may be systematically underrepresented in the research base intended to inform it [5,17,18].

The implications of this gap extend beyond research design. Maw et al [18] argue that achieving equitable AI implementation in health care likely requires pragmatic implementation science approaches that explicitly center population diversity, contextual adaptation, and community engagement—an orientation that appears largely absent from the current evidence base. Trinkley et al [19] further caution that AI applications in implementation science carry the risk of perpetuating or amplifying existing disparities if equity considerations are not prospectively embedded in system design, validation protocols, and ongoing monitoring frameworks. Without deliberate attention to whose data informs AI models and whose outcomes are measured, there is a meaningful risk that AI-enhanced tele-oncology tools could widen rather than narrow existing cancer care disparities for rural and underserved populations [5,18,19].

Implications for Future Research

Table 4 summarizes the identified research priorities, mapped to the gap type, supporting evidence, and ratings of urgency, feasibility, and equity relevance.

XSL•FO
RenderX

Table . Research priorities matrix.

Research priority	Gao type	Primary evidence (this review)	Contextual support	Urgency	Near-term feasibility	Equity relevance
Comparative effectiveness studies measuring implementation outcomes (adoption, fidelity, penetration, and sustainability) in authentic rural health care settings, with longitudinal follow-up beyond initial pilot phases	Methodological gap	Daly et al [32], Mammas et al [33]	[7,25,26]	High	Low	High
Cross-population validation of AI[a] systems, with performance disaggregated by rural versus urban status, race or ethnicity, socioeconomic position, and digital literacy level	Equity or validation gap	All 4 studies (absent in each)	Aziz et al [30], Anderson et al [31]	High	Medium	High
Research specifying translation mechanisms linking AI predictions to actionable implementation strategies, examining how predictions are acted upon by clinicians and organizations	Implementation science gap	Khairat et al [34], Daly et al [32]	Trinkley et al [19]; [7,19,24]	High	Medium	Medium
Health economic analyzes comparing AI-enhanced versus standard implementation approaches in rural settings, informing infrastructure investment, workforce, and reimbursement decisions	Economic evidence gap	Daly et al [32], Khairat et al [34]	Absent from current literature	Medium	Medium	Medium
Qualitative and mixed methods research on sociotechnical integration in resource-constrained settings, and development or testing of implementation science–informed governance frameworks	Mechanistic or governance gap	Daly et al [32]	Reddy [24]; [12,24]	Medium	High	Medium

[a]AI: artificial intelligence.

The gaps identified through this review point to several interrelated priorities for future investigation. Perhaps most foundationally, the field would benefit from comparative effectiveness studies conducted in authentic rural health care settings (rather than academic medical centers) that measure implementation science outcomes, including adoption, fidelity,

XSL•FO
RenderX

penetration, and sustainability, alongside clinical end points [7,25,26]. Scoping reviews such as this one can identify where evidence is absent, but only prospectively designed trials and implementation studies can determine whether AI-enhanced approaches offer meaningful advantages over standard care under realistic rural conditions. Longitudinal designs capable of capturing sustainability beyond initial pilot phases are likely necessary, given that existing feasibility studies were of limited duration and scope and cannot speak to whether observed benefits persist once dedicated research support is withdrawn [32,33].

Cross-population validation of AI systems represents a further priority of considerable urgency, with performance reporting disaggregated by rural versus urban status, race and ethnicity, socioeconomic position, and digital literacy level. Both Aziz et al [30] and Anderson et al [31] identify this as largely unaddressed in current literature, and the absence of such validation means that the equity implications of deploying existing AI tools in rural oncology contexts remain genuinely unknown. Relatedly, research that specifies the translation mechanisms linking AI predictions to actionable implementation strategies would address a theoretical gap articulated by Trinkley et al [19] but not yet empirically examined in rural oncology contexts. Understanding not just whether AI can predict an outcome, but how those predictions are acted upon by clinicians and organizations, and under what contextual conditions, is essential for implementation science to move beyond feasibility demonstration [7,19,24].

Health economic analyses comparing AI-enhanced and standard implementation approaches in rural settings are notably absent from the current evidence base [32,34], yet they are likely to be important for policy decisions regarding infrastructure investment, workforce development, and reimbursement, ultimately determining whether AI-enhanced tele-oncology becomes accessible to rural populations at scale. Finally, qualitative and mixed-methods research examining sociotechnical integration in resource-constrained settings would provide mechanistic insight into the organizational and cultural factors enabling or constraining implementation success—insight that quantitative approaches alone cannot offer [12,24]. Reddy [24] argues that implementation science-informed governance frameworks are a prerequisite for responsible AI deployment in health care; developing and testing such frameworks in rural oncology contexts represents a meaningful and tractable research contribution.

Limitations

Several limitations of this review should be considered when interpreting its findings. The small number of included primary studies (n=4), their heterogeneous designs, and their concentration in single academic or high-resource sites preclude comparative synthesis and limit generalizability. Geographic concentration in the United States, Greece, and Singapore means that findings may not reflect implementation contexts in low-income and middle-income countries or in rural settings with more severe infrastructure limitations. The English-language restriction may have introduced systematic bias by excluding relevant non-Anglophone literature. Formal quality appraisal was not conducted, consistent with scoping review methodology [26,27], which means the internal validity of individual study findings remains unassessed. The rapid pace of AI development also means that applications currently under development or reported in gray literature may not yet be captured in the peer-reviewed evidence base, and implementation-focused research characteristically lags behind technology development [19,24]. Finally, while Aziz et al [30] and Anderson et al [31] are used as contextual evidence throughout this discussion, their conclusions are themselves subject to the limitations of the primary studies they synthesized, including heterogeneous quality and limited rural-specific analysis.

Conclusions

This scoping review suggests that the potential of AI-enhanced predictive analytics to improve tele-oncology implementation in rural and underserved settings, while theoretically plausible, remains empirically underdeveloped. The 4 included primary studies offer isolated proof-of-concept evidence but collectively illustrate a pattern in which technological capability may outpace implementation readiness, and in which the structural constraints most characteristic of rural settings receive comparatively limited research attention. The broader contextual literature reinforces this assessment: Aziz et al [30] found that AI-enhanced remote monitoring in oncology has not yet demonstrated clear superiority over standard care, and Anderson et al [31] observed that predictive analytics research in telehealth has largely not addressed the resource optimization and implementation science questions most relevant to populations in underserved areas.

The systematic underrepresentation of rural, low-income, and racially diverse populations in AI and digital health research is not merely an empirical gap—it may reflect and perpetuate structural inequities in knowledge production that have downstream consequences for health equity [5,18]. Deploying AI tools with unexamined performance characteristics in contexts where patients already experience compounding disadvantages carries meaningful risks that the field has not yet adequately engaged with [15,16,19]. Reddy [24] and Trinkley et al [19] both suggest that responsible AI implementation in healthcare requires equity monitoring, community engagement, and governance frameworks that remain underdeveloped in current practice. Until such frameworks are established and empirically validated in rural oncology contexts, the risk that AI-enhanced tele-oncology reproduces rather than reduces existing disparities warrants serious consideration.

Advancing this field equitably will likely require a reorientation of the research agenda toward community-engaged, implementation science-informed study designs conducted in authentic rural contexts, with explicit attention to equity, sustainability, and population diversity [7,18,25,26]. Rural communities should ideally participate as collaborators in the design and governance of AI tools intended to serve them, ensuring that research questions are grounded in authentic implementation challenges [18,19]. Funding priorities, editorial standards, and implementation science frameworks that incentivize disaggregated outcome reporting, long-term

XSL•FO
RenderX

sustainability research, and equity-centered validation may support this reorientation [5,7,26]. Only through such an approach can the theoretical promise of AI-enhanced tele-oncology be meaningfully evaluated and responsibly translated for the populations who may stand to benefit most.

Acknowledgments

The authors declare the use of generative artificial intelligence (AI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to generative AI tools under full human supervision: structuring and formatting data extraction tables, including organizing extracted study characteristics, implementation barriers, and translational pathway content into standardized tabular formats. All AI-assisted outputs were critically reviewed, verified against source materials, and revised by the authors. Generative AI was not used for literature searching, screening, data interpretation, or the drawing of conclusions. The views expressed in this publication are those of the authors and do not reflect the position or policy of the Department of Veterans Affairs, United States government, State of Texas, or other organizations with which authors may be affiliated.

Funding

This scoping review received no specific grant from any public, commercial, or not-for-profit funding agency. Institutional support was provided through startup fund resources affiliated with the authors' institution. The authors received no direct financial compensation for the conduct of this research or preparation of this manuscript.

Data Availability

All data supporting this scoping review are available in the manuscript and supplementary files (PRISMA-ScR checklist, complete search strategies, inclusion/exclusion criteria, and data extraction form). Screening decisions and extracted data are available from the corresponding author upon reasonable request. All included articles are publicly available through their respective publishers as cited in the References. No primary data were collected; all data were extracted from published literature.

Authors' Contributions

LH conceived and designed the scoping review, developed the protocol, led the systematic search strategy across all 7 databases, performed data extraction and synthesis, conducted the thematic analysis, drafted the manuscript, and served as the principal investigator responsible for resolving screening conflicts. MM contributed to the conceptualization of the review framework, provided methodological guidance on implementation science frameworks, and critically revised the manuscript for important intellectual content. BE contributed to the conceptualization and design of the review, provided expert input on implementation science and health equity considerations, and critically revised the manuscript. RZ conducted independent title/abstract and full-text screening, performed data extraction, assisted with the expanded database searches following peer review, and contributed to drafting and revising manuscript sections and supplementary materials. MS conducted independent title/abstract and full-text screening, performed data extraction, contributed to the expanded search and supplementary materials development, and assisted with drafting and revising manuscript sections. GJ provided expert oversight on tele-oncology and rural health content, contributed to the interpretation of findings, and critically revised the manuscript for important intellectual content. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 20 KB - cancer_v12i1e78005_app1.docx]

Multimedia Appendix 2
Inclusion and exclusion criteria.
[DOCX File, 22 KB - cancer_v12i1e78005_app2.docx]

Multimedia Appendix 3
Data extraction form.
[DOCX File, 17 KB - cancer_v12i1e78005_app3.docx]

Checklist 1

XSL•FO
RenderX

PRISMA-ScR checklist.
[DOCX File, 12 KB - cancer_v12i1e78005_app4.docx]

Checklist 2
PRISMA-S checklist.
[DOCX File, 10 KB - cancer_v12i1e78005_app5.docx]

References

1. Roy J, Levy DR, Senathirajah Y. Defining telehealth for research, implementation, and equity. J Med Internet Res 2022 Apr 13;24(4):e35037. [doi: 10.2196/35037] [Medline: 35416778]
2. Sodhi N, Weinstein RS, Stewart K, Doarn CR. Analysis of telehealth versus telemedicine terminology in the Telemedicine and e-Health journal between 2010 and 2020. Telemed J E Health 2022 Dec;28(12):1861-1865. [doi: 10.1089/tmj.2022.0073] [Medline: 35417253]
3. Doolittle GC, Allen A. Practising oncology via telemedicine. J Telemed Telecare 1997;3(2):63-70. [doi: 10.1258/1357633971930869] [Medline: 9206275]
4. Nopour R. Barriers and facilitators of using tele-oncology in cancer care: a scoping review. BMC Health Serv Res 2025 Apr 14;25(1):541. [doi: 10.1186/s12913-025-12731-8] [Medline: 40229656]
5. Ogundeko-Olugbami O, Ogundeko O. AI-enhanced predictive analytics systems combatting health disparities while driving equity in U.S. healthcare delivery. World J Adv Res Rev 2025;25(1):2067-2084. [doi: 10.30574/wjarr.2025.25.1.0298]
6. Williams AM, Cutchin MP. The rural context of health care provision. J Interprof Care 2002 May;16(2):107-115. [doi: 10.1080/13561820220124120] [Medline: 12028892]
7. Leeman J, Rohweder C, Lafata JE, et al. A streamlined approach to classifying and tailoring implementation strategies: recommendations to speed the translation of research to practice. Implement Sci Commun 2024 Jun 17;5(1):65. [doi: 10.1186/s43058-024-00606-8] [Medline: 38886763]
8. Zullig LL, Raska W, McWhirter G, et al. Veterans Health Administration national teleoncology service. JCO Oncol Pract 2023 Apr;19(4):e504-e510. [doi: 10.1200/OP.22.00455] [Medline: 36649579]
9. Zullig LL, Jeffreys AS, Raska W, et al. Quality of care in Veterans Affairs health care system in-person and National TeleOncology Service-delivered care. JCO Oncol Pract 2025 Dec;21(12):1838-1846. [doi: 10.1200/OP-24-01040] [Medline: 40233294]
10. Shirke MM, Shaikh SA, Harky A. Implications of telemedicine in oncology during the COVID-19 pandemic. Acta Biomed 2020 Sep 7;91(3):e2020022. [doi: 10.23750/abm.v91i3.9849] [Medline: 32921719]
11. Alpert JM, Taylor G, Hampton CN, Paige S, Markham MJ, Bylund CL. Clinicians' perceptions of the benefits and challenges of teleoncology as experienced through the COVID-19 pandemic: qualitative study. JMIR Cancer 2022 Feb 24;8(1):e34895. [doi: 10.2196/34895] [Medline: 35142622]
12. Turner K, Bobonis Babilonia M, Naso C, et al. Health care providers' and professionals' experiences with telehealth oncology implementation during the COVID-19 pandemic: a qualitative study. J Med Internet Res 2022 Jan 19;24(1):e29635. [doi: 10.2196/29635] [Medline: 34907900]
13. Nesbitt TS, Cole SL, Pellegrino L, Keast P. Rural outreach in home telehealth: assessing challenges and reviewing successes. Telemed J E Health 2006 Apr;12(2):107-113. [doi: 10.1089/tmj.2006.12.107] [Medline: 16620164]
14. Gurupur VP, Miao Z. A brief analysis of challenges in implementing telehealth in a rural setting. mHealth 2022;8:17. [doi: 10.21037/mhealth-21-38] [Medline: 35449506]
15. Bhatia S, Landier W, Paskett ED, et al. Rural-urban disparities in cancer outcomes: opportunities for future research. J Natl Cancer Inst 2022 Jul 11;114(7):940-952. [doi: 10.1093/jnci/djac030] [Medline: 35148389]
16. Monroe AC, Ricketts TC, Savitz LA. Cancer in rural versus urban populations: a review. J Rural Health 1992;8(3):212-220. [doi: 10.1111/j.1748-0361.1992.tb00354.x] [Medline: 10121550]
17. Morris BB, Rossi B, Fuemmeler B. The role of digital health technology in rural cancer care delivery: a systematic review. J Rural Health 2022 Jun;38(3):493-511. [doi: 10.1111/jrh.12619] [Medline: 34480506]
18. Maw AM, Trinkley KE, Glasgow RE. The role of pragmatic implementation science methods in achieving equitable and effective use of artificial intelligence in healthcare. J Gen Intern Med 2024 May;39(7):1242-1244. [doi: 10.1007/s11606-023-08580-y] [Medline: 38172408]
19. Trinkley KE, An R, Maw AM, Glasgow RE, Brownson RC. Leveraging artificial intelligence to advance implementation science: potential opportunities and cautions. Implement Sci 2024 Feb 21;19(1):17. [doi: 10.1186/s13012-024-01346-y] [Medline: 38383393]
20. Rossi M, Rehman S. Integrating artificial intelligence into telemedicine: evidence, challenges, and future directions. Cureus 2025 Aug;17(8):e90829. [doi: 10.7759/cureus.90829] [Medline: 40988810]
21. Sharma S, Rawal R, Shah D. Addressing the challenges of AI-based telemedicine: best practices and lessons learned. J Educ Health Promot 2023;12:338. [doi: 10.4103/jehp.jehp_402_23] [Medline: 38023098]

XSL•FO
RenderX

22. Wu H, Lu X, Wang H. The application of artificial intelligence in health care resource allocation before and during the COVID-19 pandemic: scoping review. JMIR AI 2023;2. [doi: 10.2196/38397]
23. Pozzar RA. Integrating remote symptom monitoring, person-centered analytics, and artificial intelligence to advance precision health symptom science in oncology. Semin Oncol Nurs 2025 Aug;41(4):151901. [doi: 10.1016/j.soncn.2025.151901] [Medline: 40368651]
24. Reddy S. Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. Implement Sci 2024 Mar 15;19(1):27. [doi: 10.1186/s13012-024-01357-9] [Medline: 38491544]
25. Handley MA, Gorukanti A, Cattamanchi A. Strategies for implementing implementation science: a methodological overview. Emerg Med J 2016 Sep;33(9):660-664. [doi: 10.1136/emermed-2015-205461] [Medline: 26893401]
26. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
27. Pollock D, Tricco AC, Peters MDJ, et al. Methodological quality, guidance, and tools in scoping reviews: a scoping review protocol. JBI Evid Synth 2022 Apr 1;20(4):1098-1105. [doi: 10.11124/JBIES-20-00570] [Medline: 34446668]
28. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
29. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev 2021 Jan 26;10(1):39. [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
30. Aziz F, Bianchini D, Olawade DB, Boussios S. The impact of AI-driven remote patient monitoring on cancer care: a systematic review. Anticancer Res 2025 Feb;45(2):407-418. [doi: 10.21873/anticanres.17430] [Medline: 39890180]
31. Anderson E, Lennon M, Kavanagh K, et al. Predictive data analytics in telecare and telehealth: systematic scoping review. Online J Public Health Inform 2024 Aug 7;16:e57618. [doi: 10.2196/57618] [Medline: 39110501]
32. Daly B, Kuperman G, Zervoudakis A, et al. InSight Care pilot program: redefining seeing a patient. JCO Oncol Pract 2020 Oct;16(10):e1050-e1059. [doi: 10.1200/OP.20.00214] [Medline: 32468925]
33. Mammas CS, Mamma AS, Papaxoinis G, Georgiou I. Remote AI supported E-multidisciplinary oncology conference in breast cancer as a technology and method to optimize outcomes in the peripheries. Stud Health Technol Inform 2022 Jan 14;289:309-312. [doi: 10.3233/SHTI210921] [Medline: 35062154]
34. Khairat S, Liu S, Zaman T, Edson B, Gianforcaro R. Factors determining patients' choice between mobile health and telemedicine: predictive analytics assessment. JMIR mHealth uHealth 2019 Jun 8;7(6):e13772. [doi: 10.2196/13772] [Medline: 31199332]
35. Liu D, Lin Y, Yan R, Wang Z, Bold D, Hu X. Leveraging artificial intelligence for digital symptom management in oncology: the development of CRCWeb. JMIR Cancer 2025 Jun 16;11:e68516. [doi: 10.2196/68516] [Medline: 40523210]

Abbreviations

AI: artificial intelligence
EMR: electronic health record
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension
PRISMA-SCR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by M Balcarras; submitted 18.Jun.2025; peer-reviewed by A Billis, P Codyre; revised version received 24.Feb.2026; accepted 02.Apr.2026; published 16.Jul.2026.

Please cite as:
Husain L, Mullins M, Etingen B, Zafar RM, Siddiqui M, Jackson GL
AI-Enhanced Predictive Analytics to Optimize Tele-Oncology Implementation in Rural Settings: Scoping Review
JMIR Cancer 2026;12:e78005
URL: https://cancer.jmir.org/2026/1/e78005
doi:10.2196/78005

© Laiba Husain, Megan Mullins, Bella Etingen, Raaed Mohammed Zafar, Mediha Siddiqui, George L Jackson. Originally published in JMIR Cancer (https://cancer.jmir.org), 16.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Perspectives of Adolescents and Young Adults, Caregivers, and Health Care Providers on Regional Cancer Care: Qualitative Study

Joanne Tay[1*], MPH, RN, PhD; Mohammad Jarrar[2*], MD; Telford Yeung[3*], MD, PhD; Jessica C Kichler[4*], CPsych, PhD

[1]Faculty of Nursing, University of Windsor, Windsor, ON, Canada

[2]Department of Pediatrics, Windsor Regional Hospital, Windsor, ON, Canada

[3]Department of Pediatrics, University of Calgary, Calgary, AB, Canada

[4]Department of Psychology, University of Windsor, Windsor, ON, Canada

[*]all authors contributed equally

Corresponding Author:
Joanne Tay, MPH, RN, PhD
Faculty of Nursing
University of Windsor
401 Sunset Ave
Windsor, ON, N9B 3P4
Canada
Phone: 1 519 253 3000 ext 6618
Email: jtay@uwindsor.ca

Abstract

Background: Adolescents and young adults with cancer have distinct developmental and psychosocial needs that require care models bridging pediatric and adult oncology systems. While survival outcomes have improved, there is growing recognition of the need to strengthen coordinated care, psychosocial support, and survivorship services. In Ontario, regional and community-based cancer programs play a central role in delivering accessible care and are well-positioned to support adolescents and young adults closer to home. However, variation in infrastructure, workforce capacity, and system-level coordination influences how developmentally appropriate oncology care is implemented. Existing literature has focused on tertiary and specialty centers, highlighting the need to leverage and strengthen regional systems, providers, and community resources to support high-quality adolescent and young adult cancer care.

Objective: This study aimed to explore how structural barriers and supports shape the experiences of adolescents and young adults with cancer, their caregivers, and health care providers (HCPs) in accessing and navigating regional cancer care in Ontario.

Methods: An exploratory qualitative study was conducted between July 2024 and March 2025, guided by Bronfenbrenner's Ecological Systems Theory and a social constructionist orientation. Semistructured interviews were conducted with 14 participants (adolescents and young adults, caregivers, and HCPs) in Southwestern Ontario. Participants were recruited through purposive sampling. Data were analyzed using reflexive thematic analysis. Trustworthiness and rigor were strengthened through reflexive team discussions and audit trail documentation.

Results: A total of 2 adolescents and young adults, 6 caregivers, and 6 HCPs were interviewed. Youth were in their early 20s and represented different cancer diagnoses. Most caregivers (n=5) were mothers, and HCPs had 2-20 years of experience. Six key themes were generated: (1) structural gaps in coordinated care, (2) reliance on individual champions, (3) developmental and environmental misalignment, (4) barriers to youth engagement and autonomy, (5) gaps in psychosocial, palliative, and survivorship support, and (6) local system enablers. Participants described fragmented care poorly suited to adolescents' and young adults' needs. Care environments often defaulted to child- or adult-oriented models, limiting autonomy and engagement. Psychosocial, palliative, and survivorship supports were unevenly available, particularly during transitions off treatment. Despite these challenges, local satellite clinics and community organizations reduced travel burden, fostered continuity, and provided long-term emotional and financial support.

Conclusions: Adolescent and young adult cancer care in regional settings is delivered within evolving systems adapting to youth's developmental and psychosocial needs. Our findings suggest that targeted investments in workforce development, coordinated communication structures, and age-appropriate care models can further enhance regional capacity to support adolescents

XSL•FO
RenderX

and young adults across the cancer trajectory. By shifting from reliance on informal solutions toward more structured, team-based approaches, regional programs can build resilient, youth-responsive systems that promote high-quality, equitable adolescent and young adult cancer care.

International Registered Report Identifier (IRRID): RR2-10.2196/76877

(JMIR Cancer 2026;12:e85096) doi:10.2196/85096

KEYWORDS

adolescent; health services accessibility; health services; neoplasm; qualitative research

Introduction

Adolescents and Young Adults With Cancer: Challenges Within Care

Adolescents and young adults with cancer represent a distinct population whose needs have historically been overlooked within pediatric and adult oncology systems [1]. While overall cancer survival has improved, adolescents and young adults continue to experience slower gains and face persistent inequities in access, supportive care, and survivorship services [1]. Cohort and qualitative studies published over the past several years indicate that these inequities persist despite increased recognition of adolescent and young adult–specific needs, with ongoing gaps in psychosocial support, care coordination, and survivorship planning across cancer systems [2-4]. In Ontario, policy efforts have sought to promote developmentally appropriate care, yet regional and community settings often lack the infrastructure and resources to deliver it [5]. Within this context, adolescents and young adults continue to encounter challenges that extend beyond medical treatment.

Cancer diagnosis during adolescence and young adulthood presents unique challenges extending beyond medical treatment to encompass educational, social, and developmental disruptions [6,7]. Globally, adolescents and young adults with cancer experience structural misalignment between health care systems designed for children or adults and the developmental needs of youth in transition [6,8]. Unlike younger children who rely on parental advocacy or adults who navigate systems independently, adolescents and young adults are in a distinct developmental stage characterized by emerging autonomy, identity formation, and increasing self-determination [9]. However, health care structures are often inadequately designed to recognize this developmental stage, creating what literature has described as a dual disadvantage where youth experience gaps in both pediatric and adult care systems [10]. Despite growing evidence documenting high psychosocial distress, disrupted life trajectories, and unmet developmental needs among adolescents and young adults with cancer, much of this literature remains concentrated in tertiary or specialty care contexts, offering limited insight into how these challenges are navigated within regional and community-based systems [11-14].

Systemic Barriers Beyond Clinical Care

The structural inadequacy extends beyond clinical care into broader systems surrounding adolescent and young adult cancer treatment. Educational institutions face challenges accommodating youth with fluctuating health needs, employers demonstrate limited understanding of cancer's long-term impacts, and social support systems frequently default to child-focused or adult-oriented models that inadequately address adolescent developmental needs [3]. The convergence of these structural barriers creates cumulative disadvantages that may persist beyond treatment completion, impacting educational achievement, career development, and psychosocial outcomes [15-17]. Within the Canadian health care context, accessible care delivery is a core mandate for Ontario's community hospitals [18,19]; yet, achieving equitable access to cancer services between community and tertiary oncology centers remains an ongoing challenge, particularly for adolescents and young adults [12].

An Ecological Systems Lens

Bronfenbrenner's Ecological Systems Theory provides a framework for examining how systemic barriers operate across multiple levels to constrain adolescent and young adult cancer care delivery [9]. At the exosystem level, institutional policies, funding structures, and professional training deficits establish structural conditions that either facilitate or constrain quality care delivery. Research applying this framework has documented how institutional constraints, including unclear mandates, underresourced settings, and workforce limitations, compromise the system's capacity to address youth-specific needs [2,20]. Despite efforts to decentralize tertiary care pediatric oncology services through initiatives such as the Pediatric Oncology Group of Ontario (POGO) and its satellite clinics, substantial community-based inequities in adolescent and young adult care persist. While some satellite sites provide essential follow-up and supportive care for younger patients, adolescents and young adults older than 18 years or those requiring specialized interventions must access tertiary centers for diagnosis, treatment initiation, and subspecialist consultation [16,21]. For youth navigating developmental transition characterized by increasing autonomy, identity formation, and life planning, these structural care disruptions may be particularly destabilizing [22]. This theoretical lens guided our examination of system-level conditions, policies, organizational processes, and service structures, rather than focusing solely on individual experience.

Theoretical Underpinning

A social constructionist epistemological orientation and relativist ontological stance shaped the study, recognizing that participants' accounts reflect socially situated interpretations of their interactions with health systems. By incorporating perspectives from youth, caregivers, and health care providers (HCPs), the design enabled an integrative view of the

XSL•FO
RenderX

organizational and structural dynamics influencing adolescent and young adult cancer care.

Resource Constraints in Community Hospitals

In community health care settings, these systemic challenges are amplified by infrastructure limitations and resource constraints. Community hospitals are mandated to deliver outpatient and supportive care; however, many lack the necessary infrastructure, staffing, and specialized services to provide comprehensive, developmentally appropriate adolescent and young adult oncology care [16,23,24]. Consequently, youth and their families in smaller communities require travel, often repeatedly, to tertiary centers for diagnosis, treatment, and acute care management [21,25]. These care disruptions create substantial logistical, financial, and emotional burdens, particularly for families with limited socioeconomic resources, single-parent households, newcomers to Canada, and those with reduced informal support networks [3].

Community hospitals typically operate with limited oncology-specific resources, creating challenges for clinical teams who must adapt adult or pediatric protocols for adolescent and young adult patients without specialized training or institutional support [8]. Care providers often lack training in adolescent and young adult oncology, with limited guidance on managing complex psychosocial issues or coordinating with tertiary care teams [2,20]. Systemic limitations often necessitate reliance on informal champions, typically nurses or allied health professionals, who also function as care coordinators in the absence of more structured support. While their contributions are essential, this reliance on individual practitioners highlights deeper systemic vulnerabilities, including insufficient interdisciplinary teams, the absence of contingency planning, and inadequate infrastructure to support workforce sustainability. High provider turnover, burnout, and service discontinuity represent recurring challenges, particularly in settings where care delivery depends on limited personnel [8,12]. Community hospitals serving adolescents and young adults frequently lack resources, training, and capacity to deliver developmentally appropriate care [7,26]. Youth perspectives are often marginalized in these settings, where reliance on parental involvement in care inadequately addresses developmental needs for autonomy and self-advocacy within an adolescent and young adult population [17].

Knowledge Gaps in Community-Based Adolescent and Young Adult Oncology

Much of the existing research has examined travel burdens and navigation barriers in tertiary and adult oncology settings, leaving the experiences of adolescents and young adults and their families in community settings comparatively underexplored [26,27]. Care providers in these contexts face considerable challenges in addressing service gaps and supporting youth through complex medical and psychosocial transitions [8,20]. Understanding how systemic barriers operate across multiple ecological levels in community contexts is essential for developing interventions that address the underlying causes of care fragmentation.

Study Objectives

To explore how structural barriers and supports shape the experiences of adolescents and young adults with cancer, their families, and HCPs in accessing and navigating regional cancer care.

Methods

Study Design

We used an exploratory qualitative design informed by Bronfenbrenner's Ecological Systems Theory [9]. This design was chosen to help investigate and understand adolescents and young adults', caregivers', and HCPs' experiences with oncology care within the complex health care systems. The study is reported in accordance with the Consolidated Criteria for Reporting Qualitative Research.

Sampling and Recruitment

Recruitment took place between July 2024 and March 2025. Purposive sampling was used to recruit participants who represented diverse social contexts, including variation in gender identity, family composition, and neighborhood socioeconomic background. Eligibility criteria for youth included (1) diagnosis of cancer between January 2018 and December 2024, (2) age 15-18 years at time of diagnosis, (3) residence in Southwestern Ontario and receipt of care through the pediatric oncology satellite clinic, (4) completion of treatment and at least 3 months in remission prior to interview, (5) adequate communicative and cognitive ability to participate, and (6) fluency in English. Youth currently on active treatment, recently diagnosed (<3 months), or identified by clinicians as being at risk of psychological harm from participation were excluded. Parents or caregivers of eligible youth were invited to participate, with English fluency as the only inclusion criterion. HCPs were recruited using the same approach, with study information distributed through community hospital networks and professional contacts, and interested individuals contacting the research team directly. Eligible HCPs included physicians, nurses, social workers, and child life specialists with direct clinical responsibilities in adolescent and young adult cancer care within regional hospitals or POGO satellite clinics. Eligibility was confirmed during initial contact through brief screening questions aligned with the study's inclusion criteria prior to scheduling interviews. Data analysis was conducted with all participants as a whole rather than as 3 separate groups (adolescents and young adults, caregivers, and HCPs), as each group had fewer than the required 8 participants. Data saturation was met when no new data yielded new insights.

Data Collection

Recruitment materials (information sheet and poster) were distributed through the POGO satellite clinic in Southwestern Ontario and via closed social media networks, including ChildCan and POGO. Interested individuals responded by contacting the research team directly using the contact information provided on the study materials. This recruitment approach supported purposive sampling by enabling the identification of participants with direct experience of adolescent and young adult cancer care within community hospital settings,

XSL•FO
RenderX

including youth, caregivers, and HCPs involved in care delivery. Semistructured interviews were conducted individually by JT or MM based on participants' expressed availability. Two youth who were interviewed were young adults and preferred to be interviewed independently without their caregiver. Microsoft Teams meeting invites were emailed to participants for access to the meeting; contact information was provided if any technical difficulties arose. Verbal consent for study participation and audio recording was obtained during the meeting prior to proceeding with the interview. Consent was treated as an ongoing process, and participants were reminded that they could decline to answer questions or withdraw from the interview at any time without consequence. Interviews were audio recorded on Microsoft Teams and lasted 45-90 minutes, after which they were transcribed and verified for accuracy prior to analysis. Additional procedural details are available in the published protocol [28].

Youth interviews focused on their daily lives and routine, along with their transition throughout treatment. Caregiver interviews revolved around the emotional impact of navigating the health care system and experiences with caregiving. HCP interviews focused on care delivery, navigating systemic barriers, and recommendations for care. See Multimedia Appendices 1 and 2 for interview guides.

Data Analysis

Analysis followed Braun and Clarke's [29] 6-phase reflexive thematic analysis, which supported the identification of recurring patterns while attending to the contextual meanings expressed by participants. Transcripts were deidentified and managed in NVivo (version 15; Lumivero) to help organize and manage codes and find patterns for theme development. The Bronfenbrenner's Ecological System framework was used to help with data analysis and theme development to structure themes around adolescent and young adult oncology care within the complex and multifaceted layers of the health care system [9].

In Phase 1, team members individually reviewed transcripts (JT, MM, and BG) to identify features related to systemic processes, access pathways, and coordination challenges. In Phase 2, data were coded with attention to organizational and structural dynamics, rather than individual coping or emotional responses. Initial codes were inductively developed through iterative engagement with the data, involving the application and refinement of descriptive labels to meaningful text segments and constant comparison across transcripts. For example, segments describing repeated travel to tertiary centers due to unavailable local services were initially coded as "care fragmentation" and later integrated into a broader theme reflecting systemic discontinuity in regional adolescent and young adult cancer care. In Phases 3-5 (theme generation, review, and refinement), codes were iteratively developed into potential themes and reviewed across youth, caregiver, and HCP accounts to ensure coherence and depth at the system level. As part of this process, data from youth, caregivers, and HCPs were examined both within and across participant groups to identify areas of convergence, complementarity, and divergence in experiences of regional adolescent and young adult cancer care. The final phase synthesized themes into a thematic structure that reflects how system organization, developmental alignment, and community-level conditions influence adolescent and young adult cancer care experiences. Both inductive and deductive reasoning informed analysis, drawing on existing adolescent and young adult oncology models to refine interpretation.

Early coding differentiated between individual experiences and system-level processes. Through team discussion, codes related to travel burden, delayed communication, and unavailable services were grouped as structural issues rather than individual concerns, reflecting their shared system-level origins. Themes were retained based on their presence across youth, caregiver, and HCPs' accounts, rather than their frequency within a single group, to support relevance across roles and settings. Divergent cases were examined to refine themes and ensure the findings reflected variation in regional adolescent and young adult cancer care experiences rather than a single dominant perspective. Any disagreements were discussed until a consensus was reached. Regular team debriefings supported critical reflection on emerging patterns and facilitated discussions across the diverse participant groups, including youth, caregivers, and HCPs.

Ethical Considerations

This study received ethical clearance from the Windsor Regional Hospital Research Ethics Board (REB 24-477) and the University of Windsor Research Ethics Board (REB 24-027). All participants were English-speaking and resided or worked in the Windsor-Essex region. Informed consent was received prior to proceeding with the interview and audio recording. Both interviewers, a registered pediatric nurse and a research assistant trained in qualitative interviews, are trained in detecting signs of distress and pausing interviews if needed. Mental health resources were also provided to participants during interviews. Interviews were then transcribed verbatim by trained research assistants, deidentified, and stored on a secure computer. Participants were then given an electronic gift card via email.

Rigor and Reflexivity

To strengthen the rigor and trustworthiness of our findings, we used several strategies, including methodological alignment, role-based triangulation, maintaining an audit trail, and ongoing reflexive practice. Three research team members (JT, MM, and BG), each with qualitative research experience, collaboratively led the coding and thematic development, bringing complementary perspectives to the analysis. We systematically recorded analytic decisions and individual reflections in an audit trail, revisiting these notes and consulting with the research team to ensure interpretations remained closely tied to the data. To enhance credibility and transferability, the research team documented and discussed key analytic decisions throughout the analysis. Throughout the study, we remained reflexive by journaling and openly examining how our disciplinary training and varying clinical involvement in adolescent and young adult oncology shaped our analytic lens and helped mitigate potential biases. This reflexive process helped us foreground participant voices and mitigate the influence of preexisting assumptions.

XSL•FO
RenderX

Results

Participant Characteristics

A total of 16 participants signed up for the study, but 2 did not meet the eligibility criteria. As a result, 14 participants were interviewed, including 2 youth, 6 caregivers, and 6 HCPs. Among the caregivers, the majority (5/6) were mothers, with 1 father. Youth participants were in their early 20s and represented different cancer diagnoses. HCPs included professionals with varied roles, such as child life specialist, family support liaison, pediatrician, pharmacist, and registered nurses. There was a range of 2 to >20 years of experience in adolescent and young adult oncology among HCPs. Full demographic details are provided in Table 1.

Table 1. Participant characteristics.

Variable	Value
Parents (n=6)	
Age in years, mean (range)	47 (39-53)
Ethnicity, n (%)	
White	5 (83.3)
Prefer not to answer	1 (16.7)
Education, n (%)	
High school diploma	1 (16.7)
College/trade school	2 (33.3)
University	1 (16.7)
Postgraduate degree	1 (16.7)
Other	1 (16.7)
Employment status, n (%)	
Not employed	1 (16.7)
Part-time	2 (33.3)
Full-time	3 (50)
Relationship to youth who has cancer, n (%)	
Mother	5 (83.3)
Father	1 (16.7)
Youth (n=2)	
Age in years, mean (range)	23 (22-24)
Ethnicity, n (%)	
White	2 (100)
Education, n (%)	
High school diploma	1 (50)
College/trade school	1 (50)
Employment status, n (%)	
Not employed	1 (50)
Full-time	1 (50)
Health care providers (n=6)	
Length of time providing oncology care to young adults and youth, n (%)	
<5 years	2 (33.3)
5-10 years	1 (16.7)
>10 years	3 (50)
Number of young adults and youth in current practice, mean (range)	14 (5-25)

XSL•FO
RenderX

Overview of Themes

Thematic Summary

Analysis generated 6 interconnected themes that illuminate how structural barriers operate across ecological system levels to constrain adolescent and young adult cancer care. These themes demonstrate the complex relationship between individual experiences and systemic deficiencies that characterize regional adolescent and young adult oncology. These themes are (1) structural gaps in coordinated and equitable oncology care; (2) overreliance on individual champions in an underresourced system; (3) developmental and environmental misalignment in youth cancer care; (4) barriers to youth engagement, autonomy, and voice; (5) gaps in psychosocial, palliative, and survivorship supports; and (6) system enablers that support youth and families. Findings reflect converging and contrasting perspectives across youth, caregivers, and HCPs, highlighting shared challenges as well as role-specific experiences within regional adolescent and young adult cancer care.

Structural Gaps in Coordinated and Equitable Oncology Care

HCPs consistently mentioned fragmented communication and uneven care quality between tertiary and regional centers as key barriers to coordinated adolescent and young adult oncology care. Despite the need for youth-centered care, participants described a disconnect across sites, leaving both HCPs and caregivers underinformed and youth vulnerable to suboptimal care.

Structural gaps in coordinated and equitable oncology care were described as multifaceted and interconnected. Across participant groups, these gaps encompassed fragmented communication between tertiary and regional centers, uneven access to clinical expertise and oncology-trained personnel, and limited availability of medications and treatment protocols in community settings. Communication breakdowns included inconsistent written documentation, delayed information sharing, and limited direct access to tertiary specialists, leaving regional teams and families to bridge gaps informally. Together, these structural deficiencies undermined continuity, timely decision-making, and eroded confidence in the coordination of care across settings, particularly within regional follow-up and shared care pathways, even as confidence in tertiary oncology expertise remained high. One regional HCP articulated this disconnection vividly:

> *[The youths] don't tell me anything. They will go to their tertiary care centre and then divulge all this information to their oncologist, but that sometimes doesn't get back to me.* [HCP01]

These communication gaps hindered the delivery of timely and responsive care, placing added pressure on families who were left to navigate disconnected systems on their own. Providers described encountering patients during follow-up appointments without access to critical discharge information, leading to delays and uncertainty:

> *Patients come back for follow-up, and it is not in the discharge doctor's note. I don't know what I am supposed to do. Patients have to wait until we figure it out ... there are a lot of barriers there and parents get upset.* [HCP04]

Rather than functioning as part of an integrated care system, regional teams were often left reactive, informed only when issues arose:

> *There's room for improvement in terms of more consistent check-ins and streamlined communication between the care teams in [tertiary center] and [regional center]. Communication usually occurs only when there's an issue.* [HCP06]

These accounts illustrate how fragmented communication extended beyond interpersonal exchanges to include gaps in written discharge documentation, delayed information transfer, and limited direct communication channels between regional providers and tertiary oncology teams. While adolescents and young adults and families continued to seek and trust tertiary oncologist expertise, confidence in the system's ability to function as a coordinated whole was diminished, especially during transitions and regional follow-up.

Beyond communication, participants described structural gaps related to limited clinical expertise and staffing capacity in regional settings, which constrained the ability to deliver consistent, oncology-informed care. Parents, particularly those with clinical backgrounds, described regional teams as underresourced and underprepared for the demands of pediatric oncology:

> *We need trained, more educated nurses who know what they're doing with respect to oncology pediatric patients ... if there was none on the floor, they had to call the adult oncology nurse to come.* [P06]

From the provider perspective, limited caseloads and infrastructure posed ongoing challenges to maintaining expertise:

> *You want experts in the field who are seeing this daily. So, if you can't support that with volume ... the care is watered down.* [HCP06]

Structural inequities were also evident in access to medications and treatment protocols, with families encountering limitations to locally available therapies and inconsistent access to chemotherapy regimens. These constraints often required families to travel to tertiary centers or accept suboptimal symptom management in regional settings. One parent described the frustration of discovering that their regional center would not carry a nausea medication that had been effective for their child:

> *There was a medication which they figured out for my son's nausea [but] our cancer center does not have it ... they will not carry it or even order it for you.* [P06]

Others described having to travel to a tertiary treatment center to access specific chemotherapy not available at the local satellite clinic:

> *He could get some [chemo] here at the satellite clinic, but there were certain kinds that ... he'd have to travel*

XSL•FO
RenderX

> *to [tertiary treatment center] for the day and then come back.* [P05]

Continuity of care was frequently sustained through the consistent engagement and relationship-building efforts of individual providers. Their ongoing presence created stability for families within care pathways that were not always seamlessly connected, highlighting that relational continuity often compensated for, rather than resolved, underlying structural fragmentation, and underscoring opportunities to strengthen system-level structures that can better support this work.

Overreliance on Individual Champions in an Underresourced System

Youth cancer care in smaller cities was often sustained by a few dedicated champions, who served as the relational and procedural anchors in the absence of coordinated infrastructure. These longstanding relationships became vital for families navigating a complex and emotionally charged care journey. Nurses who had followed youth through multiple treatment cycles were often seen as irreplaceable. One parent reflected on the central role of a trusted nurse:

> *[The nurse] is someone who is crucial for [the cancer care in the city] ... [she] was a very big part in like, I guess the good part of the cancer diagnosis ... Built a relationship with [youth] as well, so that was a familiar face.* [P04]

Similarly, a provider described the unique role of a long-serving staff member who became the "pillar" of the clinic:

> *I will say that our [satellite] clinic is fantastic in terms of our recent lead staff, who had just retired. I think she was ... the pillar of the clinic. And so, I can definitely appreciate her longitudinal relationships with these children, and you know that consistency and comfort that she brought to the clinic.* [HCP06]

Another provider described the relational disconnect that arose when covering for a colleague on leave:

> *They trust the [healthcare provider] very much, and I have a harder time bonding with that family. I always try really hard to get the parents to open up about how they're feeling.* [HCP04]

While the relational model provided meaningful continuity of care, it placed considerable strain on those stepping into these "individual champion" roles temporarily. Relief staff were expected to handle not only complex clinical tasks such as port access and chemotherapy preparation, but also to meet the relational expectations usually fulfilled by familiar providers. The same provider further reflected on this challenge, noting, "when it's not that primary person, it is difficult and very nerve-wracking for me" (HCP04), highlighting the implicit comparison to a trusted key team member. Participants described how high parental expectations and the lack of team-based backup intensified provider stress:

> *Parents were constantly hovering and commenting, 'will the staff be able to do it like [nurse] does?' ... and I was stressed. There are not enough people who are confident with ports ... I did not have a backup.* [HCP05]

This fragility in having trusting relationships when receiving regional cancer care was not limited to nursing. Other HCPs also reported similar vulnerabilities due to understaffing and role overload. In contrast to tertiary centers, which had designated personnel for oncology services, regional providers often worked in isolation across multiple service areas:

> *Tertiary center has two [healthcare providers of the same role] just dedicated to oncology. I am doing all of these alone ... How do we make it so that it's not just one person?* [HCP05]

While these individual efforts were essential to maintaining continuity, they also underscored the extent to which the care experience relied on personal relationships rather than systemic design. Familiarity with a trusted provider was a powerful buffer against an environment that otherwise felt disjointed from the needs of youth. Without systemic structures in place to ensure developmental responsiveness, care often defaulted to rigid categories (ie, child or adult) that failed to reflect the developmental realities of adolescents and young adults. This disconnect surfaced not only in staffing and coordination but also in the broader environments where care and recovery unfolded.

Developmental and Environmental Misalignment in Youth Cancer Care

Youth, parents, and providers consistently described both regional and tertiary cancer care environments as misaligned with the developmental and psychosocial needs of adolescents and young adults. Designed mainly for children or adults, current systems often leave adolescents and young adults without appropriate spaces or supports, reinforcing feelings of isolation, exclusion, and disempowerment. Many participants described health care environments that were either child-centric or adult-oriented, with no age-appropriate options in between. A provider noted, "We have a pediatric or an adult cancer ward. But what about the young adults?" (HCP05). This forced placement into inappropriate programs often made adolescents and young adults feel out of place and visibly different. One parent recalled the discomfort of being placed in an adult support group:

> *They had registered us for the adult [cosmetic] program ... and so we were in a room full of women ... and my daughter felt very uncomfortable.* [P02]

Beyond issues of age-appropriateness, participants also pointed to systemic limitations in accommodating youth under immunosuppressive protocols. One youth described the social isolation of being physically present but barred from meaningful interaction:

> *Even when I did go back, I wasn't allowed to go around people. I was just able to see them from a distance.* [Y01]

Even when youth were physically present, the surrounding structures seldom provided space for independent voice, reflection, or choice, deepening their sense of invisibility within the system.

XSL•FO
RenderX

This disconnect within exosystems extended beyond health care into education systems as well, where youth and families encountered similar structural gaps in communication and developmental support. For some, physical inaccessibility in school environments compounded the recovery process and emotional toll:

> *Her mobility was hugely restricted ... and I found that the school was not very understanding ... didn't provide extra time ... loaded her up with heavy textbooks.* [P02]

Educational disruption and poor reintegration into school activities were among the most frequently raised concerns. Youth described being met with apathy or misunderstanding from teachers, contributing to a sense of abandonment at a time when support was most needed:

> *It'd be nice just for [teachers] to be a little more supportive, not just kind of shun you or shove you off.* [Y01]

Parents echoed these frustrations, noting the absence of structured communication between hospital-based care and educational institutions. In their view, the lack of coordinated roles and clear processes left families to navigate these transitions on their own:

> *It would be nice to have more collaboration between the contact person at the hospital and the schools ... There wasn't really a process or a system of support in place that I found was really adequate.* [P06]

Even when educational support was available, it was often not tailored to the developmental stage of older youth. As one provider explained, volunteer tutoring programs, while well-intentioned, frequently fell short of meeting the academic needs of senior high school students:

> *We have a volunteer tutoring program ... but for kids that are in grades 11 and 12... It's very difficult to find someone to help who is qualified.* [HCP02]

Youth also described the emotional burden of having to repeatedly explain their cancer history to peers and teachers, an experience that underscored the lack of systemic communication mechanisms:

> *It would be nice if the school knew [so] I didn't have to restate everything and relive all the things that happened.* [Y01]

These mismatches between care environments and developmental needs not only left youth feeling out of place in schools and hospitals but also reinforced patterns of exclusion that limited their ability to participate meaningfully in their own care.

Barriers to Youth Engagement, Autonomy, and Voice

As mentioned earlier, adolescence is a crucial period for developing autonomy and identity. Yet, youth, parents, and providers described barriers to engaging adolescents and young adults as active participants in their care. Young people often felt sidelined by the need for constant parental presence in a hospital setting and health care systems that lacked the structure to center the youth voices in decision-making. The most immediate barrier was the structural absence of private, youth-directed spaces in clinical settings. Providers noted that youth rarely had opportunities to speak independently, as parents were nearly always present during appointments:

> *When they're in the clinic or admitted to the hospital, there is always a parent present. They don't have any time or opportunity to express themselves without a parent listening.* [HCP01]

Even when youth were physically present, the surrounding environment often lacked opportunities for independent voice, reflection, or choice, which may have contributed to a sense of invisibility within the system. In these situations, clinicians observed that parents often took the lead in communication, sometimes speaking on behalf of the youth.

> *You know, it was ... trying to ask them if they had any questions ... kind of engage them in their care versus their parents ... [instead] I get the parents' answer.* [HCP05]

These missed opportunities for confidential conversations with clinicians were seen as limiting not only for youth but also for providers trying to build rapport:

> *[Private spaces or time alone with youths] would be better at helping them open up ... but parents are always there with the kids* [HCP02]

The consequences of these missed youth-provider interpersonal interactions were twofold. First, youth often disengaged entirely, either due to discomfort or lack of perceived agency:

> *With the youth and adolescent population, I struggle ... I feel like I don't necessarily always have the time to explore all the underlying issues ... It's more superficial.* [HCP06]

Second, parents often became the sole conduit for emotional disclosures, an arrangement that left providers relying on second-hand reports to assess psychosocial needs. One provider noted:

> *The biggest issue is that [youth] don't talk about their feelings; parents will usually reach out to me.* [HCP02]

This not only limited the ability to engage youth authentically but also risked reinforcing dependency at a time when adolescents are developmentally primed to begin asserting independence and autonomy. Providers emphasized the importance of structured, developmentally appropriate approaches to foster youth voice and participation:

> *As they're trying to navigate adolescence and personal identity ... young adult transition clinics would be quite important ... they need to build that trust before being able to actually open up.* [HCP06]

Some pointed to peer-facilitated spaces as promising platforms for agency and shared learning:

> *There needs to be a peer-on-peer family group ... having a facilitator could ensure accuracy and support.* [HCP04]

XSL•FO
RenderX

Yet even when youth sought to engage, the absence of integrated psychosocial, palliative, and survivorship supports left many of their broader needs unmet.

Gaps in Psychosocial, Palliative, and Survivorship Supports

The psychosocial aspects of youth cancer care, especially mental health, end-of-life support, and survivorship systems, were described as fragmented, underresourced, and overlooked. Although essential for long-term well-being, these supports are often less developed and less available than medical services. Mental health emerged as a particular area of concern. Parents described the cumulative emotional trauma usually experienced by their children and families and lamented the lack of proactive mental health resources or referrals.

> *There's a lot of trauma. And for no sort of mental health support to be recommended for you to get referrals, no one coming in to talk about it, that is very concerning ... Not just for the child, but for parents, caregivers, and siblings too.* [P06]

Access to services was also marked by inequities. Providers and parents pointed to long waitlists for youth counseling and the financial barriers to accessing private care. One provider and one parent noted:

> *I don't really think that [mental health] support is widely accessible. There are long waitlists for accessing mental health services, and going private depends on whether they have the money or benefits.* [HCP02]

> *I think if there was counselling later that was offered, that wasn't a cost or whatever ... She definitely does have some PTSD [post-traumatic stress disorder] ... and OHIP [Ontario Health Insurance Plan] should cover counselling for critical conditions.* [P05]

For many, the systemic gaps in support became particularly stark during the survivorship and transition periods for the youth. Parents and HCPs both noted that psychosocial support tended to decrease once treatment ended:

> *You're 17, and everything is organized. At 18, you're expected to navigate it all on your own ... Youth transition clinics would be quite important. Smaller cities would definitely benefit from that.* [HCP06]

> *It almost seems like there needs to be a period of transition into adult care ... not just jump from a tertiary children's centre to the adult oncology.* [P03]

This abrupt transition was especially difficult for youth experiencing developmental vulnerability:

> *The transition wasn't gradual; it was sudden. It was becoming stressful for me because he didn't belong there. He belongs around more fun things. Not the adult cancer centre.* [P02]

Equally concerning were the significant gaps in palliative care and bereavement support. Families reported being left to navigate end-of-life planning and grief without structured guidance or psychological care:

> *Try to imagine planning a funeral and a birthday party three days apart ... There is absolutely nothing there [to support a family member after death]... You are so alone.* [P02]

Even when community programs existed, they were often underfunded or unavailable to many:

> *The grief program for kids through CMHA [Canadian Mental Health Association] has no additional funding ... It's purely [based] on fundraising.* [P06]

The challenge of finding trauma-informed therapy was ongoing and deeply taxing for some parents.

> *All the trauma he went through and all the nightmares. Thank God, we took him for therapy ... but it was not easy to find.* [P02]

While many families faced challenging moments in the cancer journey without adequate psychosocial or palliative support systems, their experiences also revealed what mattered when care felt responsive. Participants highlighted local, community-based, and satellite-driven supports that eased the burdens of treatment and recovery.

> *The [satellite] clinic was very helpful ... for blood work or the chemo he could get here instead of us having to travel to [tertiary treatment center].* [P04]

> *It was really good to be close to home, because I could go home and rest after ... And [nurse name] was just always there ... It wasn't some new person every time. It made it feel like people cared, like I wasn't just a number.* [Y01]

These were not incidental benefits but crucial system enablers that, when present, enhanced access, continuity, and emotional well-being for youth and their families.

System Enablers That Support Youth and Families

While participants acknowledged systemic gaps in cancer care, they also highlighted key supports that eased the burden for youth and families. Local satellite clinics, financial assistance, and community-based resources provided critical support. Access to nearby care reduced travel strain, though participants emphasized that proximity must be matched by adequate staffing and resources to ensure safe, high-quality care.

> *It would be nice to have more care closer to home, provided we could, if we had the resources to do it safely and the necessary support staff. We are a satellite center, and we must acknowledge that we are not the primary one, but it would be nice.* [HCP05]

Youth also voiced a desire for greater local capacity, envisioning a future where advanced therapies, like stem cell transplants, might eventually be delivered in regional settings.

> *More local treatment on a grander scale, more options locally to do more treatment here, even so far as like the stem cell transplants.* [Y02]

Alongside local access, financial and transportation support were key. For many families, unexpected levels of financial assistance, offered through local charities and hospital

XSL•FO
RenderX

partnerships, helped offset the high costs of treatment, travel, and time away from work. As one parent shared: "I was very surprised by how much financial support there really is" (P03). Providers also noted that these systemic supports extended beyond the treatment period, with some organizations offering educational bursaries to youth survivors pursuing postsecondary education: "[Community organization] gives bursaries for postsecondary students." (HCP01).

The long-term engagement model offered by community organizations like ChildCan and Fight Like Mason was especially valued [30]. These groups were not seen as peripheral supports, but as integral components in the care ecosystem, offering emotional, logistical, and financial assistance that extended well beyond hospital walls. Youth and families described these relationships as enduring, often continuing for years after active treatment had ended. One provider noted: "Even ten years later, [survivors] can apply for a bursary" (HCP02). This long-term community support was especially valuable in smaller cities, where formal survivorship and mental health services were limited. Community organizations provided a sense of continuity and connection when formal systems withdrew, ensuring that families did not feel abandoned at the end of treatment.

These accounts highlight the importance of flexible, locally embedded supports that bridge gaps in formal care. While satellite clinics, financial aid, and community organizations may not replace the comprehensive health care infrastructure of tertiary centers, they serve as crucial scaffolding, helping youth and families stay connected, supported, and hopeful throughout and beyond the cancer care journey.

Discussion

Principal Findings

This study explored how adolescents and young adults with cancer, their caregivers, and HCPs experienced regional cancer care in Ontario, and identified key barriers to developmental alignment, care coordination, and equitable access within these systems. Our findings reveal structural gaps that create challenges for delivering equitable care to adolescents and young adults with cancer. Communication between tertiary and regional centers is inconsistent, leaving families to rely on individual champions rather than reliable systems. Fragmented communication had implications beyond information exchange, shaping relationships, trust, and continuity of care. Inadequate or delayed discharge documentation limited direct communication channels, and inconsistent access to oncology specialists left regional providers uncertain about care plans and families unsure whom to rely on for guidance. These gaps shifted the burden of coordination onto caregivers and strained therapeutic relationships, particularly for nurses who were expected to provide continuity without access to complete or timely information. From a nursing perspective, such breakdowns undermine relational care by constraining providers' ability to anticipate needs, follow through on care plans, and build confidence with youth and families. These findings highlight the need for more formalized, standardized communication processes between tertiary and regional teams to support continuity, accountability, and relationship-centered care in adolescent and young adult oncology. Care environments often default to binary pediatric or adult models, with adolescent and young adult autonomy and psychosocial needs frequently unaddressed. Gaps in mental health, palliative care, and survivorship services were also evident throughout the care journey, reflecting broader misalignments within existing systems. In this context, families described drawing on local supports and community organizations outside of the formal care system. Our findings point to an opportunity to better align existing communication processes with the needs of adolescent and young adult oncology care across institutional boundaries, drawing on structured handover approaches already used in other areas of health care.

In regional settings, individual providers often took on responsibilities beyond their formal roles to deliver adolescent and young adult cancer care. While this is consistent with rural oncology research indicating that committed professionals sustain care continuity [8,31], our findings suggest that more systematic infrastructure could help strengthen these efforts. This pattern may reflect broader challenges, including infrastructure gaps, insufficient system-level guidance, and limited provider education in developmentally appropriate adolescent and young adult care [8,32]. Reliance on key individuals rather than a sustainable infrastructure means that a single provider's departure can disrupt service continuity and erode trust among patients and families [33,34]. While countries like the United Kingdom and Australia have made strides toward centralized adolescent and young adult care frameworks, implementation remains inconsistent across many jurisdictions, including Canada, leaving individual champions to fill systematic gaps [8,20]. This has implications for workforce sustainability and system resilience, as documented challenges arise when providers operate beyond their scope without adequate support [35-38]. Our findings suggest that while individual champions remain important and should be celebrated, there is significant potential to evolve toward well-resourced, team-based models that ensure more consistent, equitable adolescent and young adult cancer care [6,20]. These dedicated efforts could be enhanced through more systematic support to improve adolescent and young adult care quality and access across regional settings.

Another key insight from our study is the mismatch between existing care models and the developmental realities of adolescents and young adults. Pediatric care often centers on parental involvement, which can delay the development of independence, while adult care assumes emotional and cognitive maturity that many adolescents and young adults are still developing [2,4,39]. Such gaps contribute to disengagement and disruptions in care continuity. The misalignment is further complicated when adolescents and young adults are treated in spaces that do not foster trust or reflect their evolving identities [12,15]. Without youth-appropriate space and support, many are reluctant to raise sensitive issues or participate fully in care decisions [10]. These gaps extend beyond clinical care into important life domains, with our participants describing educational interruptions, limited attention to fertility, and a lack of structured support for work reintegration. These patterns

XSL•FO
RenderX

echo the BRIGHTLIGHT cohort findings of high psychosocial distress among adolescents and young adults due to ineffective communication and insufficient age-appropriate guidance [17], highlighting the need for developmentally relevant supports in sexuality, education, and career planning often overlooked in adult oncology settings [32].

Psychosocial, palliative, and survivorship care remains inconsistently delivered in cancer care settings despite growing awareness of their importance [17,26]. Our findings demonstrate that regional variation in care quality and access creates equity barriers even in publicly funded systems like Canada [27,40,41]. Extending this work, our study also highlights how these gaps are experienced within regional and community-based settings, where limited mental health infrastructure and workforce shortages further constrain access to counselling, bereavement, and trauma-informed care, and our study participants described feeling abandoned during care transitions and struggling with unresolved trauma and grief without access to counselling or bereavement support [20,31,38]. These gaps were particularly significant in rural areas, where mental health infrastructure and workforce shortages were limited [12,25,31]. Psychosocial supports are often overlooked as biomedical treatment takes priority over emotional recovery and life-stage reintegration, partly because of gaps in integrating trauma-informed psychosocial care due to inadequate funding and lack of structural support [10,32]. As a result, families are left to navigate fragmented supports independently, adding emotional strain and worsening inequities.

Strengths and Limitations

This study is limited by its focus on 1 region (Southwest Ontario), which may affect transferability to other health care contexts. A key strength of this study lies in the integration of perspectives across youth, caregivers, and HCPs, enabling examination of converging and contrasting accounts of regional adolescent and young adult cancer care. This sample is predominantly White, which may affect generalization to other populations. While purposive sampling was used to ensure variation across key social characteristics, such as age, socioeconomic background, and geographical location, it may not fully capture diversity across other rural and urban regions in Canada. Although the study included 2 adolescents and young adults with different diagnoses, the sample size is limited, which may be due to the study's specificity and inclusion criteria. This may not reflect the perspectives of youth with other oncology diagnoses or those entirely disengaged from cancer care, such as those lost to follow-up or experiencing ongoing structural exclusion. Due to resource constraints, only parents or caregivers with English fluency were included within the study, which can limit the diversity of this sample.

Comparison With Prior Work

Much of the existing literature has focused on broad national or international perspectives, including the Institute of Medicine workshop summary [10] and the *Canadian Partnership Against Cancer's Adolescents and Young Adults with Cancer: A Reference Report* [12]. Our study offers a regional perspective on how these systemic issues manifest in day-to-day cancer care delivery. While large-scale cohort investigations such as BRIGHTLIGHT emphasize psychosocial outcomes [17] and qualitative studies of advanced disease foreground existential concerns [11], our analysis focuses on how structural barriers and developmental misalignments are experienced within a smaller, underserved Canadian region. This regional perspective complements existing models-of-care reviews [20,32] by illustrating how resource limitations and workforce reliance shape the care landscape in ways not fully captured in national or tertiary-level studies.

Recommendations for Further Research

Future research should explore findings across broader geographic contexts, including interprovincial comparisons of adolescent and young adult oncology models, and examine how structural determinants, such as income, Indigeneity, and gender, may intersect with access to cancer care and subsequent outcomes. Mixed method studies integrating patient-reported outcomes, service use data, and system-level indicators could also deepen understanding of how developmental misalignment and fragmentation affect quality, equity, and survivorship trajectories. Longitudinal research is needed to assess the effectiveness of targeted system interventions, such as youth navigation roles or regional adolescent and young adult coordinators.

Implications for Policy and Practice

To advance equity and quality in adolescent and young adult oncology, policy and practice may benefit from prioritizing care models that reflect young people's developmental contexts, including age-appropriate education, counseling, and clinical environments that support autonomy. Although models such as age-specific care settings and coordinated transition services have shown promise, their implementation has been uneven and often limited to well-resourced settings [6,20]. Sustained improvements will likely require investment in cross-institutional coordination, provider training, and supportive policies, with examples showing that dedicated funding and leadership can drive progress [8,20].

Within this study, nurses were consistently identified as central to supporting adolescents and young adults, often taking on coordinating and advocacy roles despite limited formal recognition or specialized training. These findings suggest opportunities for strengthening nursing contributions in adolescent and young adult care, such as (1) integrating developmental psychology into nursing education, (2) establishing advanced practice adolescent and young adult care coordinator roles, (3) fostering youth-friendly care environments, (4) incorporating trauma-informed approaches to address mental health needs, and (5) developing adolescent and young adult–specific interventions through nursing-led research. Advancing these areas will require institutional support, including dedicated funding, protected coordination time, and recognition of nursing expertise, to move from informal efforts toward more structured, developmentally responsive systems of care.

Conclusion

Adolescents and young adults with cancer receiving care in regional settings often face services that are fragmented, not

XSL•FO
RenderX

fully aligned with developmental needs, and unevenly supported across systems. Families described depending on the efforts of individual clinicians, particularly nurses, to bridge service gaps, while psychosocial, palliative, and survivorship supports were less consistently available. Local clinics and community organizations played an important role in providing accessible and ongoing support, particularly close to home. These findings suggest a need to transition from individually driven solutions to more reliable, coordinated, and developmentally responsive models of adolescent and young adult cancer care that integrate equity, sustainability, and youth voice.

Acknowledgments

We are deeply grateful to the youth, caregivers, and health care providers who participated in this study for their time, trust, and willingness to share their experiences. We also thank Melissa Miljanovski, Brianna Grandi, Bayan Nasra, and Adelia Scofield for their contributions to data transcription and cleaning, and Joanne Ta for preparing the manuscript for submission and for proofreading.

Data Availability

The datasets generated during or analyzed during this study are available from the corresponding author on reasonable request.

Funding

This study was funded by Windsor Cancer Centre Foundation through the WE-SPARK Health Institute Igniting Discovery Grant.

Authors' Contributions

JT was responsible for conceptualization, methodology, investigation, formal analysis, data curation, project administration, funding acquisition, and drafting the original manuscript. MJ contributed to conceptualization and methodology and assisted with manuscript review and editing. TY contributed to conceptualization. JCK contributed to conceptualization, methodology, and investigation, conducted formal analysis, and assisted with manuscript review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide for health care providers.
[DOCX File , 22 KB - cancer_v12i1e85096_app1.docx]

Multimedia Appendix 2
Interview guide for youth and parents/caregivers.
[DOCX File , 27 KB - cancer_v12i1e85096_app2.docx]

References

1. Ting C, Chen X, Poh YN, Yeo ZZ, Ang QR, Chong PH. 'I really feel like I am ageing early': Aa interpretative phenomenological analysis of the lived experience of young adults with life-limiting conditions (YALLC). BMJ Public Health 2024;2(1):e000343. [doi: 10.1136/bmjph-2023-000343] [Medline: 40018141]
2. Hovsepyan S, Hoveyan J, Sargsyan L, Hakobyan L, Krmoyan L, Kamalyan A, et al. The unique challenges of AYA cancer care in resource-limited settings. Front Adolesc Med 2023;1:1279778. [doi: 10.3389/fradm.2023.1279778]
3. Braun I, Friedrich M, Morgenstern L, Sender A, Geue K, Mehnert-Theuerkauf A, et al. Changes, challenges and support in work, education and finances of adolescent and young adult (AYA) cancer survivors: a qualitative study. Eur J Oncol Nurs 2023;64:102329. [doi: 10.1016/j.ejon.2023.102329] [Medline: 37172338]
4. Avutu V, Lynch KA, Barnett ME, Vera JA, Glade Bender JL, Tap WD, et al. Psychosocial needs and preferences for care among adolescent and young adult cancer patients (ages 15-39): a qualitative study. Cancers (Basel) 2022;14(3):710 [FREE Full text] [doi: 10.3390/cancers14030710] [Medline: 35158975]
5. Chiu M, Ali A, Leung F, Dong C, Pechlivanoglou P, Hodgson D, et al. Sociodemographic factors and utilization of pediatric oncology satellite clinics in Ontario, Canada. JAMA Netw Open 2024;7(12):e2452063 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.52063] [Medline: 39724373]
6. Meeneghan MR, Wood WA. Challenges for cancer care delivery to adolescents and young adults: present and future. Acta Haematol 2014;132(3-4):414-422. [doi: 10.1159/000360241] [Medline: 25228567]
7. Stoneham SJ. AYA survivorship: the next challenge. Cancer 2020;126(10):2116-2119 [FREE Full text] [doi: 10.1002/cncr.32774] [Medline: 32129904]

XSL•FO
RenderX

8. Fardell JE, Patterson P, Wakefield CE, Signorelli C, Cohn RJ, Anazodo A, et al. A narrative review of models of care for adolescents and young adults with cancer: barriers and recommendations. J Adolesc Young Adult Oncol 2018;7(2):148-152. [doi: 10.1089/jayao.2017.0100] [Medline: 29298105]
9. Harper CR, Steiner RJ, Brookmeyer KA. Using the social-ecological model to improve access to care for adolescents and young adults. J Adolesc Health 2018;62(6):641-642 [FREE Full text] [doi: 10.1016/j.jadohealth.2018.03.010] [Medline: 29784107]
10. Nass SJ, Beaupin LK, Demark-Wahnefried W, Fasciano K, Ganz PA, Hayes-Lattin B, et al. Identifying and addressing the needs of adolescents and young adults with cancer: summary of an institute of medicine workshop. Oncologist 2015;20(2):186-195 [FREE Full text] [doi: 10.1634/theoncologist.2014-0265] [Medline: 25568146]
11. Barton KS, Steineck A, Walsh CA, Lau N, O'Donnell MB, Rosenberg AR. "I won't get to live my life the way I planned it": a qualitative analysis of the experiences of adolescents and young adults with advanced cancer. Pediatr Blood Cancer 2023;70(10):e30554 [FREE Full text] [doi: 10.1002/pbc.30554] [Medline: 37438862]
12. Adolescents and young adults with cancer. Canadian Partnership Against Cancer. 2017. URL: https://www.partnershipagainstcancer.ca/topics/adolescents-young-adults-with-cancer/ [accessed 2025-06-02]
13. Burgers VWG, van der Graaf WTA, van der Meer DJ, McCabe MG, Rijneveld AW, van den Bent MJ, et al. Adolescents and young adults living with an uncertain or poor cancer prognosis: the "new" lost tribe. J Natl Compr Canc Netw 2021;19(3):240-246. [doi: 10.6004/jnccn.2020.7696] [Medline: 33668023]
14. Garaszczuk R, Yong JHE, Sun Z, de Oliveira C. The economic burden of cancer in canada from a societal perspective. Curr Oncol 2022;29(4):2735-2748 [FREE Full text] [doi: 10.3390/curroncol29040223] [Medline: 35448197]
15. Larsen MH, Hansson KE, Larsen EH, Fridh MK, Petersen NN, Mellblom AV, et al. The gap between expectations and reality: a qualitative study of psychosocial challenges of young childhood cancer survivors from the PACCS study. Eur J Cancer Care (Engl) 2022;31(6):e13696 [FREE Full text] [doi: 10.1111/ecc.13696] [Medline: 36029045]
16. Smith M, Kurup S, Devaraja K, Shanawaz S, Reynolds L, Ross J, et al. Adapting an adolescent and young adult program housed in a quaternary cancer centre to a regional cancer centre: creating equitable access to developmentally tailored support. Curr Oncol 2024;31(3):1266-1277 [FREE Full text] [doi: 10.3390/curroncol31030095] [Medline: 38534928]
17. Korenblum C, Taylor RM, Fern LA, Hough R, Wickramasinghe B. Factors affecting psychosocial distress in adolescents and young adults with cancer: BRIGHTLIGHT cross-sectional and longitudinal cohort study results. Cancers (Basel) 2025;17(7):1196 [FREE Full text] [doi: 10.3390/cancers17071196] [Medline: 40227784]
18. Ontario health regions. Ontario Health. 2025. URL: https://www.ontariohealth.ca/about-us/our-programs/ontario-health-regions [accessed 2025-06-02]
19. Public Hospitals Act, R.S.O. Government of Ontario. 1990. URL: https://www.ontario.ca/laws/statute/90p40 [accessed 2025-06-02]
20. Osborn M, Johnson R, Thompson K, Anazodo A, Albritton K, Ferrari A, et al. Models of care for adolescent and young adult cancer programs. Pediatr Blood Cancer 2019;66(12):e27991. [doi: 10.1002/pbc.27991] [Medline: 31524328]
21. 2020 POGO Surveillance Report. Pediatric Oncology Group of Ontario. 2020. URL: https://www.pogo.ca/research-data/data-reports/2020-pogo-surveillance-report/ [accessed 2025-06-02]
22. Alberts DS, Hess LM. Fundamentals of Cancer Prevention. Berlin Heidelberg: Springer; 2008.
23. Adolescents and young adults with cancer. Canadian Cancer Society. URL: https://cancer.ca/en/about-us/our-health-equity-work/adolescents-and-young-adults-with-cancer [accessed 2025-06-02]
24. Framework for supporting adolescents and young adults with cancer. Canadian Partnership Against Cancer. 2019. URL: https://www.partnershipagainstcancer.ca/topics/framework-adolescents-young-adults/ [accessed 2025-06-02]
25. Payne S, Jarrett N, Jeffs D. The impact of travel on cancer patients' experiences of treatment: a literature review. Eur J Cancer Care (Engl) 2000;9(4):197-203. [doi: 10.1046/j.1365-2354.2000.00225.x] [Medline: 11829366]
26. Avery J, Campbell KL, Mosher P, Al-Awamer A, Goddard K, Edwards A, et al. Advanced cancer in young adults (YAs): living in a liminal space. Qual Health Res 2024;34(1-2):72-85 [FREE Full text] [doi: 10.1177/10497323231204182] [Medline: 37844970]
27. Booker R. Building bridges between clinic and community: supporting patients and caregivers living in rural and remote Canada. Can Oncol Nurs J 2023;33(4):509-516 [FREE Full text] [Medline: 38919591]
28. Tay J, Jarrar M, Yeung T, Kichler J. Cancer care experiences among adolescents, caregivers, and health care providers in a regional Canadian context: protocol for a qualitative study. JMIR Res Protoc 2025;14(1):e76877 [FREE Full text] [doi: 10.2196/76877] [Medline: 40953833]
29. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. Qual Psychol 2022;9(1):3-26. [doi: 10.1037/qup0000196]
30. Childcan. Childcan. URL: https://childcan.com [accessed 2025-06-02]
31. Levit LA, Byatt L, Lyss AP, Paskett ED, Levit K, Kirkwood K, et al. Closing the rural cancer care gap: three institutional approaches. JCO Oncol Pract 2020;16(7):422-430. [doi: 10.1200/op.20.00174]
32. Gupta AA, Papadakos JK, Jones JM, Amin L, Chang EK, Korenblum C, et al. Reimagining care for adolescent and young adult cancer programs: moving with the times. Cancer 2016;122(7):1038-1046 [FREE Full text] [doi: 10.1002/cncr.29834] [Medline: 26848554]

XSL•FO
RenderX

33. Priority 4: eliminate barriers to people getting the care they need. Canadian Partnership Against Cancer. URL: https://www.partnershipagainstcancer.ca/cancer-strategy/strategic-priorities/priority-4-eliminate-barriers/ [accessed 2025-06-02]
34. Rush KL, Seaton CL, Burton L, Smith MA, Li EP. The healthcare experiences of rural-living canadians with and without a primary care provider: a qualitative analysis of open-ended cross-sectional survey responses. Prim Health Care Res Dev 2025;26:e1. [doi: 10.1017/s1463423624000677]
35. Grunfeld E, Whelan TJ, Zitzelsberger L, Willan AR, Montesanto B, Evans WK. Cancer care workers in Ontario: revalence of burnout, job stress and job satisfaction. CMAJ 2000;163(2):166-169 [FREE Full text] [Medline: 10934978]
36. Harry ML, Sudak NL, Engels MJ, Horn KK, Dean K, Poplau S, et al. Physician and advanced practice clinician burnout in rural and urban settings. J Am Board Fam Med 2024;37(1):43-58. [doi: 10.3122/jabfm.2023.230233r1]
37. Singh S, Farrelly A, Chan C, Nicholls B, Nazeri-Rad N, Bellicoso D, et al. Prevalence and workplace drivers of burnout in cancer care physicians in Ontario, Canada. JCO Oncology Practice 2022;18(1):e60-e71. [doi: 10.1200/op.21.00170]
38. Zon R, Figlin R. ASCO president speaks out on "huge problem" of oncologist burnout. Oncology News Central. 2025. URL: https://www.oncologynewscentral.com/podcast/oncology/asco-president-speaks-out-on-huge-problem-of-oncologist-burnout [accessed 2025-06-02]
39. Haase JE, Stegenga K, Robb SL, Hooke MC, Burns DS, Monahan PO, et al. Randomized clinical trial of a self-care and communication intervention for parents of adolescent/young adults undergoing high-risk cancer treatment: a report from the children's oncology group. Cancer Nurs 2022;45(4):316-331. [doi: 10.1097/ncc.0000000000001038]
40. Janardan SK, Wechsler DS. Caught in the in-between: challenges in treating adolescents and young adults with cancer. JCO Oncol Pract 2021;17(6):299-301 [FREE Full text] [doi: 10.1200/OP.21.00178] [Medline: 33914610]
41. Redondo-Sánchez D, Petrova D, Rodríguez-Barranco M, Fernández-Navarro P, Jiménez-Moleón JJ, Sánchez MJ. Socio-economic inequalities in lung cancer outcomes: n overview of systematic reviews. Cancers (Basel) 2022;14(2):398 [FREE Full text] [doi: 10.3390/cancers14020398] [Medline: 35053559]

Abbreviations

HCP: health care provider
POGO: Pediatric Oncology Group of Ontario

Edited by M Balcarras; submitted 02.Oct.2025; peer-reviewed by D Abudu-Birresborn, PN Boakye; comments to author 04.Dec.2025; revised version received 03.Feb.2026; accepted 04.Feb.2026; published 26.Feb.2026.

Please cite as:
Tay J, Jarrar M, Yeung T, Kichler JC
Perspectives of Adolescents and Young Adults, Caregivers, and Health Care Providers on Regional Cancer Care: Qualitative Study
JMIR Cancer 2026;12:e85096
URL: https://cancer.jmir.org/2026/1/e85096
doi:10.2196/85096
PMID:

©Joanne Tay, Mohammad Jarrar, Telford Yeung, Jessica C Kichler. Originally published in JMIR Cancer (https://cancer.jmir.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Health Interventions and Psychosocial Outcomes in Women With Breast Cancer: Systematic Review

Rocío Zúñiga-Tapia[1,2], MSN; Mercè Boixadós[2], PhD; Eulàlia Hernández Encuentra[2], PhD

[1]Facultad de Enfermería, Universidad Andres Bello, Talcahuano 7100, Concepción, Chile

[2]Psychology and Educational Sciences Department, Universitat Oberta de Catalunya, Barcelona, Spain

Corresponding Author:
Rocío Zúñiga-Tapia, MSN
Facultad de Enfermería, Universidad Andres Bello, Talcahuano 7100, Concepción, Chile

Abstract

Background: Breast cancer and its treatment can negatively affect women's quality of life (QoL), including physical, psychological, and social well-being. Digital health interventions have been increasingly used to provide supportive care, but their effectiveness on QoL and related psychosocial outcomes remains unclear.

Objective: This review aimed to evaluate the effectiveness of digital health interventions on QoL in women with breast cancer and to assess their effects on anxiety, depression, and self-efficacy.

Methods: This systematic review was registered in PROSPERO (CRD420250578307) and conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement. PsycINFO, Web of Science, Scopus, LILACS, and SciELO were searched in September 2025. Eligible studies were randomized controlled trials (RCTs) or controlled quasi-experimental studies with preintervention and postintervention assessments that evaluated digital health interventions in women with breast cancer during or up to 1 year after active treatment and reported QoL outcomes. Two reviewers independently screened all records (interrater agreement >90%) and extracted data using a standardized template, with independent verification by a second reviewer. Risk of bias was assessed using the Cochrane Risk of Bias tool (RoB 2) for RCTs and the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I) tool for the quasi-experimental study; findings were synthesized narratively following the Synthesis Without Meta-analysis (SWiM) guideline.

Results: A total of 408 records were identified, and 15 studies were included: 14 RCTs and 1 quasi-experimental study. Sample sizes ranged from 35 to 394 participants. Ten studies evaluated internet-based approaches such as websites or apps; 3 used text messaging (eg, WhatsApp [WhatsApp LLC] or WeChat [Tencent Technology, Shenzhen Company Limited]), 1 used virtual reality, and 1 used an email listserve. The direction of effect favored the digital intervention for QoL in 8 of the 15 studies. Self-efficacy was assessed in 5 studies and favored the intervention in only 1, with the effect not maintained at follow-up. Anxiety, depression, or emotional distress were assessed in 13 studies, with a favorable direction of effect in 5. Retention rates ranged from 68.6% to 100%. Among the 14 RCTs, 3 were judged at high overall risk of bias, and 11 raised some concerns; none was at low risk. The quasi-experimental study had a serious overall risk of bias. Certainty of evidence was low for QoL and emotional outcomes and very low for self-efficacy.

Conclusions: Digital health interventions may provide modest, short-term improvements in some dimensions of QoL in women with breast cancer. Favorable effects appeared more often in interventions with active therapeutic components or professional support; however, this pattern was based on indirect comparisons across heterogeneous studies. Effects on self-efficacy and mental health were inconsistent, and acceptability was insufficiently evaluated. Although promising, current evidence does not support strong or generalizable conclusions.

(JMIR Cancer 2026;12:e96858) doi:10.2196/96858

KEYWORDS

breast neoplasms; quality of life; digital health; telemedicine; mobile apps; self-efficacy; mental health; anxiety; depression; systematic review

Introduction

The World Health Organization (WHO) identifies breast cancer as one of the leading public health challenges globally and the most prevalent malignancy in both high-income and low- and middle-income countries [1]. According to estimates from the Global Cancer Observatory, more than 2 million cases of breast cancer were recorded worldwide in 2022, with an age-standardized rate of 46.8 per 100,000 women [2]. In addition, projections indicate that global breast cancer cases

XSL•FO
RenderX

will increase by 38% by 2050, and annual deaths from this disease will rise by 68% [3].

Women with breast cancer undergoing oncological treatment experience a range of physical, psychological, and social consequences [4,5]. Fatigue, pain, lymphedema, sexual dysfunction, and sleep disturbances are frequently reported as physical sequelae of breast cancer [6-8]. Psychological consequences such as anxiety, depression, and fear of cancer recurrence have also been demonstrated, with negative impacts on health and quality of life (QoL) [9,10]. In this context, psycho-oncological interventions have shown efficacy in improving coping capacity, enhancing self-efficacy skills, and reducing stress and emotional distress [11-13].

In light of the foregoing, digital health—defined by WHO [14] as the "field of knowledge and practice associated with the development and use of digital technologies to improve health, encompassing other uses of digital technologies for health such as the Internet of Things, advanced computing, big-data analytics, artificial intelligence, including machine learning and robotics"—emerges as a platform through which diverse interventions can be deployed. Before the COVID-19 pandemic, digital health had not been widely incorporated into standard oncological care worldwide [15]; however, its use has increased substantially in recent years. Systematic reviews have examined the impact of digital health among patients with cancer, revealing that digital interventions—such as informational websites including patient narratives, mindfulness intervention apps, physical activity programs, virtual reality (VR) for pain management, and videoconferencing—produce significantly positive outcomes in patient-reported symptoms, fatigue and pain levels, health-related QoL, functional capacity, and depression levels compared with care delivered without such tools [16-18].

AI and other advanced digital technologies are increasingly being applied across the breast cancer care continuum, including screening, diagnostic support, risk stratification, treatment planning, symptom monitoring, and survivorship support [19]. However, this review focuses specifically on digital health interventions directed at supportive care and psychosocial outcomes rather than diagnostic or treatment-decision technologies.

Given the growing use of digital health interventions in oncology, and the heterogeneity in the available findings, this systematic review aimed to evaluate the effectiveness of digital health interventions on QoL in women with breast cancer, and to assess their effects on self-efficacy and mental health outcomes.

Methods

Overview

This systematic review was registered in PROSPERO (CRD420250578307) [20] and its protocol was previously published [21]. It was performed in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [22] and following the Population, Intervention, Comparison, Outcomes, and Context (PICOC) [23]. This review addressed the following question (PICOC): in women with breast cancer during active treatment or up to 1 year after completion of active treatment (Population), do digital health interventions (Intervention), compared with usual care, standard treatment, or a waitlist (Comparator), improve QoL and, secondarily, self-efficacy, anxiety, and depression (Outcomes), and eligible interventions were primarily directed at women with breast cancer, irrespective of geographic location or care setting (Context)?

The following amendments to the registered protocol, made during the conduct of the review, are reported in accordance with PRISMA 2020 (item 24c) as detailed in Checklist 1. First, it is specified that the study designs correspond to both randomized and nonrandomized studies (randomized controlled trials [RCTs] and controlled quasi-experimental studies with preintervention and postintervention assessments). Second, Scopus was searched instead of MEDLINE, representing a deviation from the registered protocol; this was complemented by hand-searching the reference lists of the included studies. Third, risk of bias was assessed according to study design, using the revised Cochrane Risk of Bias tool (RoB 2) [24] for RCTs and the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I) tool [25] for the quasi-experimental study, in accordance with the amended protocol. Fourth, the certainty of the randomized evidence was rated using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach [26,27], an assessment not specified in the protocol, to inform confidence in the findings. Fifth, eligibility was restricted to women from diagnosis through active treatment and up to 1 year after completion of active treatment; studies of long-term survivors (more than 1 year after completing active treatment) were excluded, a temporal restriction of the population that was not specified in the registered protocol. In addition, references were managed in Rayyan [28] (Rayyan Systems Inc; rather than the registered Mendeley) and PsycINFO was searched through the ProQuest platform (ProQuest LLC; rather than EBSCOhost); these operational changes did not affect the eligibility criteria or the synthesis.

Search Strategy

The search was conducted in September 2025 using PsycINFO (ProQuest), Web of Science, Scopus, LILACS, and SciELO. Database-specific free-text terms were adapted to the syntax and indexing fields of each database. The complete search strategies are provided in Multimedia Appendix 1. All articles retrieved from the databases were exported to Rayyan software for screening [28]. The auto-resolve function was used, and any records with less than 95% similarity were manually reviewed by the lead researcher (RZ-T).

Title and abstract screening were conducted independently and in duplicate by 2 reviewers (RZ-T and MB); articles that did not meet the criteria were removed, and those with insufficient information to make a decision were carried forward to the full-text screening stage.

The full text of all potentially eligible records was likewise assessed independently and in duplicate by the same 2 reviewers. Interrater agreement was high (>90%) [29], and any

XSL•FO
RenderX

disagreements were resolved through discussion or, when needed, with a third reviewer (EHE).

Eligibility Criteria

Overview

In this review, digital health interventions were defined as the use of digital technologies—such as computers, mobile devices, wearables, and the internet—to deliver or support health-related information, services, or care. Because this is not a uniform category, interventions were interpreted according to their function: informational patient portals or websites, moderated peer-support forums, symptom monitoring and remote tracking, telerehabilitation, digitally delivered psychological therapy (eg, cognitive behavioral therapy [CBT] or mindfulness), and VR.

Inclusion Criteria

The inclusion criteria were as follows:

1. Articles in English, Spanish, or Portuguese.
2. Studies (papers or theses in databases) with research designs: RCTs and controlled quasi-experimental studies with preintervention and postintervention assessments.
3. Women with breast cancer, from diagnosis through active treatment and up to 1 year after completion of active treatment.
4. Digital interventions had to deliver or support a health-related component through a digital device or platform and report at least one dimension of health-related QoL as an outcome (eg, mindfulness, exercise, and self-efficacy activities).
5. Outcomes must include any dimension of health-related QoL (eg, physical, social, psychological, and symptom dimensions).
6. Studies must provide details about the study design, sample, and intervention components that allow the analysis to be carried out.

Exclusion Criteria

The exclusion criteria were as follows:

1. Nonempirical studies (eg, theoretical or conceptual articles)
2. Reviews, editorials, letters, expert opinions, protocols, or preprints
3. Studies focusing on long-term survivors (more than 1 year after completing active treatment), patients in palliative or end-of-life care, or male patients.

Data Extraction

Data were extracted by the lead author (RZ-T) using a Microsoft Excel template designed to collect details about each digital health intervention and were independently verified by a second reviewer (MB) for all included studies, with any discrepancies resolved by consensus. The data included (1) study characteristics (country, design, sample size, target population, recruitment setting, objective, and methods), (2) primary outcomes of the digital health intervention (mean change and effect size, if applicable), and (3) secondary outcomes of the digital health intervention (mean change and effect size, if applicable).

Quality Assessment

The risk of bias of each included study was assessed according to its design. RCTs were assessed using the revised Cochrane RoB 2 tool for randomized trials [24], which evaluates 5 domains—the randomization process, deviations from the intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result—through signaling questions, yielding a domain-level judgment and an overall judgment of “low risk,” “some concerns,” or “high risk.” The nonrandomized quasi-experimental study was evaluated using the ROBINS-I tool [25], as specified in the amended protocol. ROBINS-I assesses bias due to confounding, selection of participants, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selection of the reported result. Each study was assessed independently by 2 reviewers (RZ-T and MB), and disagreements were resolved by consensus or with a third reviewer (EHE).

Data Synthesis

Meta-analysis was not undertaken because the included studies showed substantial clinical and methodological heterogeneity in intervention type and content, comparator conditions, outcome constructs and measurement instruments, and assessment time points. In addition, comparable numerical effect estimates and measures of precision were not consistently reported across studies. Therefore, the findings were synthesized narratively in accordance with the Synthesis Without Meta-analysis (SWiM) reporting guideline [30] and Cochrane guidance for synthesizing findings using methods other than meta-analysis [31]. Studies were grouped according to the outcome domain—QoL (primary) and, secondarily, self-efficacy, anxiety, and depression—and, within each domain, by intervention function (informational portals or websites, moderated peer-support forums, symptom monitoring, telerehabilitation, digitally delivered psychological therapy, and VR), level of human support (guided vs unguided), and length of follow-up. The direction of effect was determined primarily from between-group comparisons at the principal postintervention assessment reported in each study. Findings were classified as favoring the digital intervention, favoring the comparator, or showing no clear direction of effect. Statistical significance was reported when available but was not used as the sole criterion for determining the direction of effect. The certainty of the evidence for each outcome was rated using the GRADE approach [26,27]. The quasi-experimental study was included in the narrative synthesis but was not incorporated into the GRADE certainty assessments.

Results

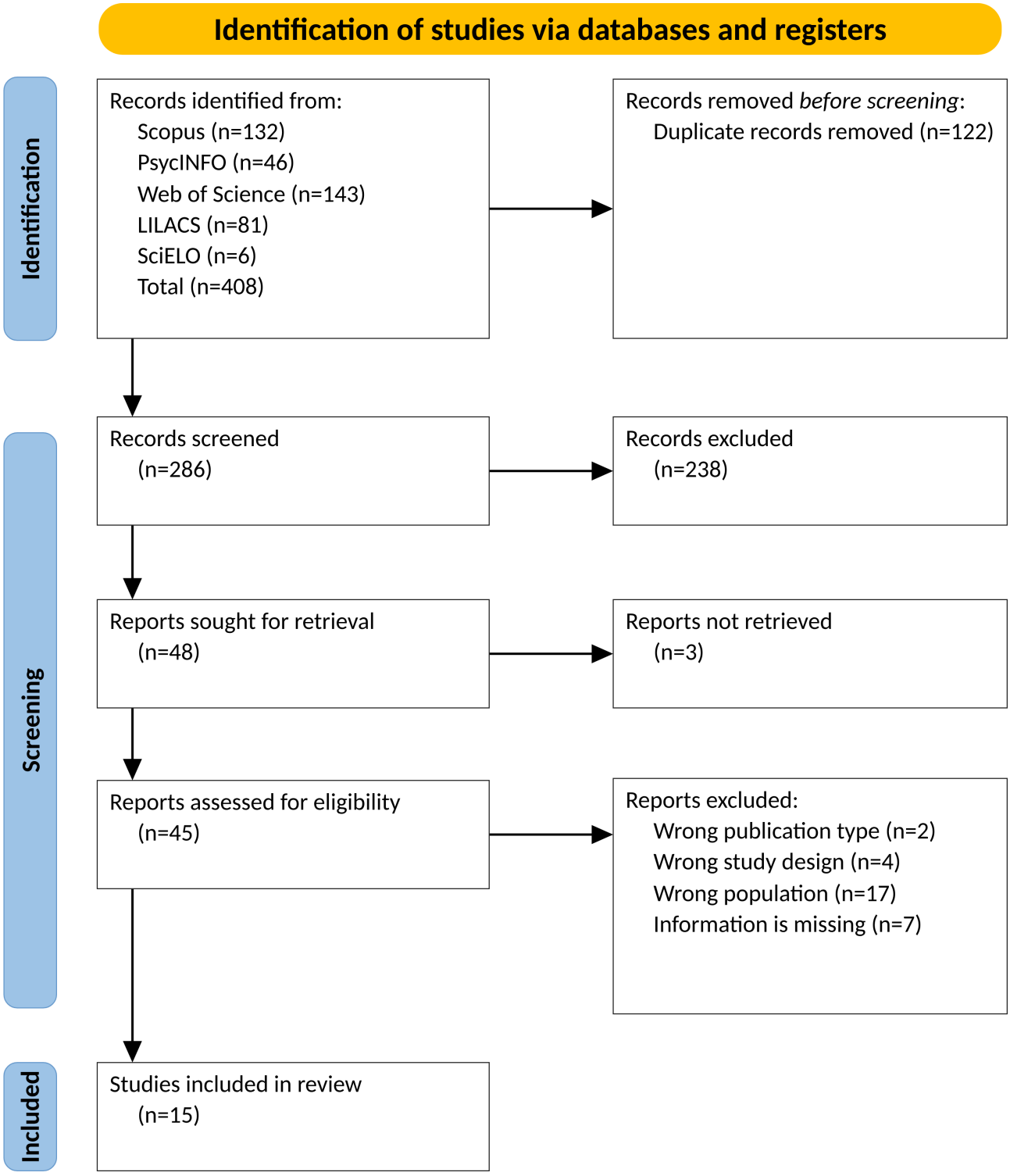
Study Selection

A total of 408 records were identified from 5 databases and uploaded to Rayyan. After removal of 122 duplicate records, 286 records underwent title and abstract screening, of which 238 were excluded. Forty-eight reports were sought for retrieval; 3 could not be retrieved, and 45 full-text reports were assessed for eligibility. Thirty reports were excluded: 2 because of publication type, 4 because of study design, 17 because of population, and 7 because insufficient information was available

XSL•FO
RenderX

to confirm eligibility. Fifteen studies met the eligibility criteria and were included in the review [32-46]. The study selection process is presented in Figure 1, and study-specific reasons for full-text exclusion are provided in Multimedia Appendix 2.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram for new systematic reviews, which included searches of databases and registers only.



Study Characteristics

Table 1 includes a summary of the participant characteristics of the included studies. The 15 included studies [32-46] had a sample size ranging from 35 to 394, with studies being conducted in China (4/15, 26.7%) [35-37,46], Australia (3/15, 20%) [34,38,43], the United States (2/15, 13.3%) [32,41], and Turkey (2/15, 13.3%) [39,44], and one study each in Finland [33], Japan [42], Brazil [40], and Singapore (1/15, 6.7% each) [45], and were published between 2010 and 2025. Participants were adult women, with mean ages 43.6 (SD 5.0) to 63.9 (SD 7.8) years. In terms of research methodology, 14 studies were RCTs [32-45], and 1 was a quasi-experimental study with pre- and postintervention assessments [46].

XSL•FO
RenderX

Table . Participant characteristics of the included studies.

Author and year	Country (design)	Sample size, randomized, enrolled, and analyzed, n (IG[a]:CG[b])	Age (years), IG:CG, mean (SD)	Breast cancer stage	Breast cancer therapy
Salzer et al (2010) [32]	The United States (RCT[c])	78 (51:27)	Not reported	I, II	Surgery, chemotherapy, radiotherapy, endocrine therapy
Ryhänen et al (2013) [33]	Finland (RCT)	98 (50:48)	54.4 (40-67):55.7 (41-69)[d]	I, II, III, IV	Surgery, chemotherapy, radiotherapy, endocrine therapy, targeted therapy
White et al (2018) [34]	Australia (RCT)	394 (212:182)	43.6 (5.0):43.9 (5.3)	I, II, III	Surgery, chemotherapy, radiotherapy, endocrine therapy, targeted therapy
Zhu et al (2018) [35]	China (RCT)	114 (57:57)	46.2 (8.5):48.2 (8.1)	I, II, III, IV	Surgery, chemotherapy
Wang et al (2022) [36]	China (RCT)	103 (51:52)	45.37 (7.59):48.17 (8.05)	0, I, II, III, IV	Surgery, chemotherapy, radiotherapy, immunotherapy
Zhang et al (2022) [37]	China (RCT)	98[e]; 90[f] (45:45)	52.29 (7.69):51.03 (7.98)	I, II, III, IV	Chemotherapy
Singleton et al (2023) [38]	Australia (RCT)	Randomized: 160 (80:80) Included in the principal analysis: 156 (78:78)	53.8 (9.6):55.7 (12.1)	I, II, III	Surgery, chemotherapy, radiotherapy, endocrine therapy, targeted therapy
Bahçaci et al (2024) [39]	Turkey (RCT)	64 (33:31)	49.1 (8.8):44.7 (10.0)	I, II, III, IV	Surgery, chemotherapy
de Aviz et al (2024) [40]	Brazil (RCT)	48 (24:24)	50.7 (11.6):49.3 (10.1)	I, II, III	Surgery, chemotherapy, radiotherapy, endocrine therapy
Graetz et al (2024) [41]	The United States (RCT)	304 (200:104)[g]	58.7 (10.5):58.3 (11.4)[g]	0, I, II, III	Chemotherapy, radiotherapy
Okuyama et al (2024) [42]	Japan (RCT)	Randomized: 130 (65:65) Analyzed: 125 (61:64)	63.9 (7.8):62.7 (7.4)	I, II, III	Surgery, chemotherapy, radiotherapy, endocrine therapy
Rigg et al (2024) [43]	Australia (RCT)	35 (17:18)	58.42 (10.88):56.73 (8.44)	IV	Surgery, chemotherapy, radiotherapy
Salman Saraç et al (2024) [44]	Turkey (RCT)	88 (44:44)	48.8 (9.0):49.3 (9.4)	Not reported	Surgery
Pang et al (2025) [45]	Singapore (RCT)	123 (62:61)	56.6 (11.7):55.4 (10.9)	0, I, II, III	Surgery
Geng et al (2024) [46]	China (quasi-experimental study)	60 (30:30)	52.50 (45.75-59.00):57.50 (44.50-62.25)[h]	I, II, III, IV	Surgery, chemotherapy

[a]IG: intervention group.

[b]CG: control group.

[c]RCT: randomized controlled trial.

[d]range, not SD.

[e]Recruited.

[f]Analyzed.

[g]Used a 3-arm design (IG1, n=98: mean age 59.4, SD 10.3; IG2, n=102: mean age 58.1, SD 10.7).

[h]median (IQR), not mean (SD).

XSL•FO
RenderX

Treatment Conditions

A total of 10 (66.7%) studies primarily evaluated internet-based interventions, including websites or mobile apps for patients with breast cancer [33-35,40-46]. One of these studies combined an application with tailored SMS support [46]. The remaining 3 (20%) studies used SMS text messages through different apps such as SMS, WeChat, WhatsApp, or others [36,38,39]. One study used head-mounted VR glasses [37], and one study used an email listserve [32].

The intervention methods included 6 (40%) studies that used psychological intervention strategies as the core component, including online peer support programs [32], mindfulness [36], stress reduction techniques through VR [37], and an online CBT-based psychosocial program [43], along with 2 studies focused on self-efficacy [35,45]. Regarding educational strategies, they were primarily included in 2 (13.3%) studies, providing cancer-related information, symptom management education, and emotional support [34,44]. Two studies worked with empowerment strategies [33,38] (13.3%), 2 studies involved remote symptom monitoring [41,42], and 3 studies focused on physical exercises [39,40,46]. The duration of these interventions ranged from 29 days to 12 months, depending on the study. The professionals involved in these interventions were nurses, psychologists, doctors (physicians or oncologists), physiotherapists, among others. Table 2 summarizes the intervention characteristics of the included studies.

XSL•FO
RenderX

Table . Summary of study characteristics.

Author and year	Intervention group	Contents	Control group	Intervention duration (data collection points)	Instrument of measurement	Results
Salzer et al (2010) [32]	Unmoderated, closed internet listserve (peer-to-peer group)	• Support and emotional experiences shared by participants without professional structure	Review of information on a cancer-related website	12 months (baseline, 4 months, and 12 months)	• FACT-B[a] • SESES-C[b] • HSCL-25[c] • IES[d] • POMS[e] • HHI[f] • MOS-SSS[g]	• No significant differences in QoL[h], self-efficacy, hope, cancer-specific stress, and social support between groups • Psychological distress and mood improved in CG[i]
Ryhänen et al (2013) [33]	Internet-based breast cancer patient pathway portal	• Education program which includes flow chart diagram of the patient pathway during breast cancer treatment, from diagnosis through surgery, chemotherapy, and follow-up	Usual care	9 months (baseline, surgery, after surgery, meeting with oncologist, before and after chemotherapy, before and after radiotherapy, and 12 months postdiagnosis)	• QoL Scale Breast Cancer[j] • STAI[k] • Own questionnaire of side effects	• No significant differences in QoL, anxiety, or side effects between groups. • Anxiety decreased in both groups
White et al (2018) [34]	"Information for Me" website	• Information resources on diagnosis, surgery, chemotherapy, radiotherapy, and survivorship, emotional responses, and support services	Usual care	6 months (baseline, 3 months, and 6 months)	• FACT-B • HADS[l] • SCNS-BC[m]	• General QoL showed an improvement in the IG[n] than the CG at 3 months, but not at 6 months • No significant differences between groups in anxiety, depression, QoL specific, and unmet needs
Zhu et al (2018) [35]	Breast cancer e-Support app	• Learning, discussion, ask-the-expert, and personal stories forums moderated by health care professionals	Usual care	3 months (baseline, 3 months, and 6 months follow-up)	• FACT-B • SICPA[o] • HADS • MDASI[p] • MSPSS[q]	• At 3 months, QoL, self-efficacy, and social support improved in IG • At 6 months, no significant differences were observed between the groups

XSL•FO
RenderX

Author and year	Intervention group	Contents	Control group	Intervention duration (data collection points)	Instrument of measurement	Results
Wang et al (2022) [36]	Internet-delivered mindfulness-based cancer recovery program delivered through WeChat	• Mindfulness techniques including body scan, mindful breathing, self-compassion, integrating mindfulness into daily life, materials, and audios provided	Usual care with health education on managing cancer symptoms, therapy side effects, diet, and exercise	1 month (baseline, postintervention, and 1-month follow-up)	• FACT-B • MDASI	• Significant improvements in QoL and symptom burden in the IG compared to the CG. Emotional distress decreased only at postintervention, but not at 1 month follow-up
Zhang et al (2022) [37]	VR-CALM[r] virtual reality–based psychotherapy intervention for symptom management	• Psychological support (symptom management, self, relationships, meaning, spirituality, and future hope)	Usual care	3 months (baseline and 3 months)	• FACT-B • DT[s] • SAS[t] • SDS[u] • CARS[v] • PFS[w] • PSQI[x]	• QoL, distress, anxiety, depression, fatigue, and sleep outcomes favored the VR-CALM group after the intervention
Singleton et al (2023) [38]	EMPOWER-SMS text messages delivered	• Physical activity, healthy diet, social and emotional well-being, endocrine therapy adherence, and side-effects management	Usual care plus 1 welcome SMS and 1 follow-up interview reminder SMS	6 months (baseline and 6 months)	• EORTC QLQ-C30[y] or EORTC QLQ-BR23[z] • SEMCD[aa] • DASS-21[ab] • GPAQ[ac] • Own self-reported endocrine therapy adherence • BIPQ[ad] • BMI	• No clear between-group differences were observed for QoL, self-efficacy, depression, anxiety, stress, physical activity, BMI, or body composition. Therapy adherence improved in IG
Bahçaci et al (2024) [39]	Tele-rehabilitation-based relaxation exercises delivered via WhatsApp	• Jacobson's progressive relaxation focusing on muscle relaxation with breathing exercises	Relaxation exercise brochure	1.5 months (baseline and 1.5 months)	• EORTC QLQ-C30 • BPI[ae] • HADS • FIS[af] • FACT-Cog[ag] • PSQI	• IG showed a significant improvement in QoL (symptom scale), depression, anxiety, sleep quality, and cognitive function. Fatigue and pain decreased in IG
de Aviz et al (2024) [40]	Asynchronous telerehabilitation via messaging app	• Exercise videos focusing on shoulder mobility, strength, and aerobic activity	Usual care	1.5 months (baseline, day 15, and 1.5 months)	• FACT-G[ah] • FACIT-F[ai] • VAS[aj]	• QoL showed significant enhancement in IG • Fatigue and pain were significantly reduced in the IG

XSL•FO
RenderX

Author and year	Intervention group	Contents	Control group	Intervention duration (data collection points)	Instrument of measurement	Results
Graetz et al (2024) [41]	Mobile app	• App: symptom and medication adherence monitoring, with weekly symptom reporting and alerts to the oncology team • App and feedback weekly tailored text messages	Usual care	6 months (baseline and 12 months)	• SF-12[ak] • PROMIS SEMS[al] • FACT-ES[am] • CPPPA[an] • Own self-reported adherence therapy	• QoL, self-efficacy, and symptom burden with no significant differences between groups • App and feedback group showed decreases in total and high-cost health care encounters and office visits
Okuyama et al (2024) [42]	ePRO monitoring app	• PRO-CTCAE symptoms (insomnia, joint pain, headache, anxiety, and hot flashes) plus daily status/medication; staff followed up as needed	Usual care	3 months (baseline and 3 months)	• FACT-B/B-TOI[ao] • EORTC QLQ-CO-MU26[ap] • PRO-CTCAE	• No significant differences in QoL and communication between health providers • High ePRO response maintained to week 10
Rigg et al (2024) [43]	Finding My Way-Advanced: self-directed online CBT-based program (website)	• Managing fear of progression, emotional distress, and physical symptoms with role functioning and social support	Usual care	1.5 months (baseline and 1.5 months)	• QLQ-C30 • DASS-21 • FOP[aq] • SCNS - Short Form • PTSD Symptom Scale-Self Report[ar] • SCNS	• QoL and role functioning, cancer-specific distress, and fear of progression showed improvements in the IG • Mental QoL, general distress, and social functioning deteriorated in the IG vs CG
Salman Saraç et al (2024) [44]	"Breast Cancer Surgery Information Guide" informative mobile app	• Surgery, treatment, and recovery information. It contained videos, text, and reminders	Routine clinic care/training only; no app	1 month (baseline and 3 weeks after surgery)	• FACT-G • HAD-A only anxiety • NCCN-DT-PL[as]	• QoL showed no significant differences between groups. Anxiety and distress showed significant reduction in IG
Pang et al (2025) [45]	iCareBreast mobile app	• Perioperative care guidance, breast cancer education, psychological support, and social support	Usual care with educational pamphlets	29 days around surgery; measures at T0, T1, T2 (baseline, 2 weeks post surgery, and 2.5 months follow-up)		

XSL•FO
RenderX

Author and year	Intervention group	Contents	Control group	Intervention duration (data collection points)	Instrument of measurement	Results
					• EORTC QLQ-BR23 • GSES[at] • HADS • MFSI[au] • Own Perioperative Care Satisfaction Scale.	• No clear between-group differences were observed for QoL, self-efficacy, anxiety, depression, and fatigue at follow-up. Satisfaction was higher in the IG at the final assessment, but the between-group difference was not statistically significant
Geng et al (2024) [46]	"Breast Care" app combined with tailored SMS-based physical activity support	• Physical activity intervention based on social cognitive theory, self-efficacy theory, and the theory of planned behavior, with personalized goals, self-monitoring, feedback, tailored messages, and support	Usual care plus printed educational material, face-to-face health education during chemotherapy visits	3 months (baseline and postintervention assessment at 3 months)	• FACT-B • HADS; IPAQ-SF[av]	• At 3 months, QoL and anxiety favored the IG, mainly because outcomes deteriorated in the CG; no significant within-group improvement was observed in the IG

[a]FACT-B: Functional Assessment of Cancer Therapy – Breast.
[b]SESES-C: Stanford Emotional Self-Efficacy Scale – Cancer.
[c]HSCL-25: Hopkins Symptom Checklist-25.
[d]IES: Impact of Event Scale.
[e]POMS: Profile of Mood States.
[f]HHI: Herth Hope Index.
[g]MOS-SSS: Medical Outcomes Study-Social Support Survey.
[h]QoL: quality of life.
[i]CG: control group.
[j]QoL Breast Cancer: Quality of Life Instrument - Breast Cancer Patient Version.
[k]STAI: The State-Trait Anxiety Inventory.
[l]HADS: Hospital Anxiety and Depression Scale.
[m]SCNS-BC: Supportive Care Needs Survey-Breast Cancer.
[n]IG: intervention group.
[o]SICPA: Stanford Inventory of Cancer Patient Adjustment.
[p]MDASI: MD Anderson Symptom Inventory.
[q]MSPSS: Multidimensional Scale of Perceived Social Support.
[r]VR-CALM: Virtual Reality Comprehensive Assisted Learning and Management.
[s]DT: distress thermometer.
[t]SAS: Self-Rating Anxiety Scale.
[u]SDS: Self-Rating Depression Scale.

XSL•FO
RenderX

[v]CARS: Concerns about Recurrence.
[w]PFS: Piper Fatigue Scale.
[x]PSQI: Pittsburgh Sleep Quality Index.
[y]EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire –Core 30.
[z]EORTC QLQ-BR23: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Breast Cancer Module 23.
[aa]SEMCD: Self-Efficacy for Managing Chronic Disease Scale.
[ab]DASS-21: Depression Anxiety and Stress Scale.
[ac]GPAQ: Global Physical Activity Questionnaire.
[ad]BIPQ: Brief Illness Perception Questionnaire.
[ae]BPI: Brief Pain Inventory.
[af]FIS: Fatigue Impact Scale.
[ag]FACT-Cog: Functional Assessment of Cancer Therapy – Cognitive Function.
[ah]FACT-G: Functional Assessment of Cancer Therapy – General.
[ai]FACIT-F: Functional Assessment of Chronic Illness Therapy – Fatigue.
[aj]VAS: Visual Analogue Scale.
[ak]SF-12: Short Form 12-Item Health Survey (Physical and Mental Component Summary).
[al]PROMIS SEMS: Patient-Reported Outcomes Measurement Information System Self-Efficacy for Managing Symptoms.
[am]FACT-ES: Functional Assessment of Cancer Therapy - Endocrine Symptoms.
[an]CPPPA: 10-item Communication: Patient and Physician Peer Assessment.
[ao]FACT-B /B-TOI: FACT-B Trial Outcome Index.
[ap]EORTC QLQ-COMU26: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Communication 26.
[aq]FOP: Fear of Progression Questionnaire.
[ar]PTSD Scale-Self Report: Post-Traumatic Stress Scale-Self Report.
[as]NCCN- DTPL: NCCN Distress Thermometer and Problem List.
[at]GSES: General Self-Efficacy Scale.
[au]MFSI: Multidimensional Fatigue Symptom Inventory.
[av]IPAQ-SF: International Physical Activity Questionnaire–Short Form.

Primary Outcome

The reviewed articles all delivered interventions that had the goal, either primary or secondary, of improving QoL, consistent with the inclusion criteria. All included studies assessed QoL as either a primary or secondary outcome, in accordance with the eligibility criteria. The studies used both breast cancer–specific and generic validated QoL instruments. FACT instruments were used in 9 (60%) studies [32,34-37,40,42,44,46], European Organisation for Research and Treatment of Cancer Quality of Questionnaire–Core 30 (EORTC QLQ-C30) or European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Breast Cancer Module 23 (EORTC QLQ-BR23) instruments in 4 (26.7%) studies [38,39,43,45], the Quality-of-Life Instrument–Breast Cancer Patient Version in 1 study [33], and the Short Form 12-Item Health Survey (SF-12) in 1 study [41]. Where studies used multiple QoL-related instruments, classification was based on the measure covering the broadest range of QoL dimensions; therefore, the single-domain FACT-Cog [39] and FACT-ES [41] subscales were excluded from the FACT tally to avoid double-counting. Following the SWiM approach, findings were summarized by the direction of effect rather than by statistical significance alone. Seven [34-37,39,40,43] of the 14 RCTs [32-45] showed a favorable direction of effect, whereas the remaining RCTs showed no clear between-group benefit or mixed findings [32,33,38,41,42,44,45]. The controlled quasi-experimental study also favored the intervention group, although this primarily reflected less deterioration relative to the control group rather than a clear within-group improvement [46]. Favorable effects appeared more frequently among interventions containing active therapeutic, behavioral, or rehabilitation components. However, because intervention formats, populations, and comparators differed substantially, this pattern should be regarded as exploratory rather than evidence of a moderating effect. Overall, the randomized evidence suggested a possible short-term benefit for QoL, with low certainty.

Other Outcomes: Self-Efficacy

Self-efficacy was assessed in 5 [32,35,38,41,45] out of the 15 (33.3%) studies [32-46] using different validated instruments. In this regard, general self-efficacy was measured using the General Self-Efficacy Scale (GSES) in 1 study [45]; cancer-specific self-efficacy was measured in 2 studies through the Stanford Emotional Self-Efficacy Scale – Cancer (SESES-C) [32] and the Stanford Inventory of Cancer Patient Adjustment (SICPA) [35]; self-efficacy for symptom management was assessed with the Patient-Reported Outcomes Measurement Information System (PROMIS) Self-Efficacy for Managing Symptoms (SEMS) [41] and the Self-Efficacy for Managing Chronic Disease Scale (SEMCD) in another study [38]. These tools were designed to assess participants' confidence in managing various aspects of their health, such as fatigue, emotional distress, and treatment-related challenges. Regarding effectiveness, the direction of effect favored the intervention in only 1 of the 5 trials. In that study, the professionally moderated Breast Cancer e-Support intervention improved cancer-related

XSL•FO
RenderX

self-efficacy at 3 months, but the effect was not maintained at 6 months [35]. The remaining studies showed no clear between-group benefit. The certainty of evidence for self-efficacy was very low (Multimedia Appendix 3).

Anxiety, Depression, and Emotional Distress

Anxiety was assessed in 11 (73.3%) out of the total selected studies [33-35,37-39,42-46] and was most commonly measured with the HADS-A, used in 6 studies [34,35,39,44-46]. Additionally, the Self-Rating Anxiety Scale (SAS) was used in 1 study [37], the Depression Anxiety and Stress Scale (DASS-21) in 2 studies [38,43], the State-Trait Anxiety Inventory in 1 study [33], and the anxiety item of the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events in 1 study [42]. Depression, on the other hand, was assessed in 8 (53.3%) studies [34,35,37-39,43,45,46], using the HADS-D in 5 studies [34,35,39,45,46], the DASS-21 [38,43], and the SDS in 1 study [37]. Emotional distress was assessed in 6 (40%) studies using the Profile of Mood States (POMS) and Hopkins Symptom Checklist-25 (HSCL-25) [32], the MD Anderson Symptom Inventory [36], the Distress Thermometer [37,44], and the stress subscale of the DASS-21 [38,43]. The direction of effect favored the intervention in 5 [36,37,39,44,46] of 13 studies [32-39,42-46]. Four of the 12 RCTs favored the intervention, whereas the remaining RCTs showed no clear benefit, effects favoring the comparator, or mixed findings. In the quasi-experimental study, anxiety favored the intervention group at 3 months, although no clear within-group reduction was observed and depression did not differ between groups [46]. Overall, the findings were heterogeneous and generally limited to short-term follow-up. The certainty of the randomized evidence was low.

Acceptability

Regarding acceptability, only 4 studies reported information on this [38,43-45]. Two studies reported good acceptability among participants in the intervention groups [38,43]. Satisfaction was assessed in 2 studies [44,45]; however, one found no statistically significant between-group difference in perioperative care satisfaction [45]. Because acceptability and satisfaction were inconsistently defined and measured, no overall conclusion regarding intervention acceptability could be established.

Retention Rate

Regarding attrition, the lowest retention rate was 68.6% [43], whereas the highest rate was 100% [46].

Quality Appraisal

The RoB 2 assessment included only the 14 RCTs and is presented in Figure 2. No trial was considered to be at low overall risk of bias. Three trials were judged to be at high overall risk of bias, mainly because of missing outcome data [34,39,43], including a pilot trial with a 34% loss to follow-up and differences between those who completed the study and those who did not [43]. The remaining 11 trials raised some concerns. Concerns were identified regarding outcome measurement because the main outcomes were subjective and self-reported by participants who were aware of their allocation to the intervention, and this knowledge could have influenced their reports. Across domains, most studies showed an adequate randomization process (D1), with allocation concealment unclear in a few; deviations from the intended interventions (D2) were generally limited but unblinded; missing outcome data (D3) was the main driver of high risk, with insufficient retention or incomplete handling of attrition in several trials; and selection of the reported result (D5) raised few concerns, while outcome measurement (D4) was limited throughout by self-reported, participant-aware assessment.

The quasi-experimental study was assessed separately using ROBINS-I and was judged to have a serious overall risk of bias [46]. The main concerns arose from the nonconcurrent sequential allocation of the control and intervention groups, the absence of statistical adjustment for potential confounding, and the measurement of subjective self-reported outcomes in participants who were aware of their intervention assignment. Risk was judged to be low for intervention classification and missing data, moderate for participant selection, deviations from intended interventions, and selection of the reported result, and serious for confounding and outcome measurement.

XSL•FO
RenderX

Figure 2. Risk of bias of the randomized controlled trials (RoB 2) [32-45].



Certainty of the Evidence

The certainty of the evidence was rated for each outcome using the GRADE approach [26,27]. Starting from high certainty, the evidence was downgraded for risk of bias and inconsistency across studies. No randomized trial was judged to be at low overall risk of bias, and 3 were at high risk. Certainty was low for QoL and for anxiety, depression, and emotional distress, and very low for self-efficacy, for which the evidence base was small and the instruments were heterogeneous. These ratings indicate that the short-term benefits observed should be interpreted with caution. Further details on the certainty assessments are provided in Multimedia Appendix 3.

XSL•FO
RenderX

Discussion

General Overview

Considering the advancement of digital health interventions in their various forms, it is essential to conduct critical reviews of the existing evidence, focusing on studies with high-quality scientific designs. This systematic review aimed to analyze the published research on digital health interventions in women with breast cancer and critically evaluate the effectiveness of these interventions. Specifically, it assessed the effects of digital health interventions on QoL and other psychological and behavioral variables, such as self-efficacy, anxiety, and depression.

Principal Findings

QoL

With respect to QoL—the primary outcome of this review—the direction of effect favored the intervention in 8 [34-37,39,40,43,46] of the 15 studies [32-46], although the certainty of this evidence was low. Positive results were observed primarily in interventions such as digitally delivered mindfulness [36], VR-assisted psychotherapy [37] or via a website [43], telerehabilitation or guided relaxation [39,40], and a tailored mobile health (mHealth) physical activity intervention grounded in social cognitive theory (SCT), self-efficacy theory, and the theory of planned behavior (TPB) [46]. These interventions incorporated active therapeutic or behavioral components and, in most cases, some degree of professional guidance or personalized support. However, in the quasi-experimental study, total QoL did not significantly improve within the intervention group, and the between-group difference mainly reflected deterioration in the control group during chemotherapy, suggesting a relative preservation of QoL rather than a clear improvement attributable to the intervention; therefore, these findings should be interpreted cautiously given the nonrandomized design and serious overall risk of bias [46]. By contrast, passive interventions generally showed no clear between-group differences in QoL, with the exception of those incorporating health care team support through alert-triggered calls or professional moderation [35]. This trend has been corroborated by recent meta-analyses of digital interventions, which demonstrate that clinician-led support yields superior outcomes compared to stand-alone passive interventions [47-49].

Furthermore, improvements in QoL were primarily short-term. Follow-up beyond 6 months was uncommon, although 3 studies included assessments at or around 12 months [32,33,41], a duration generally recommended as a baseline for assessing the durability of effects and avoiding the overestimation of results [50,51]. This lack of longitudinal data, compounded by the complexity of the oncological process, where symptoms arising from both the cancer and its treatments converge, alongside factors such as digital fatigue due to the loss of the "novelty effect," the absence of continuous human interaction, and the limited integration of digital interventions into long-term daily routines, calls into question the long-term clinical relevance of these interventions [52-54]. These observations suggest that sustaining benefits beyond the initial intensive phase may require ongoing human support or periodic reinforcement rather than one-off, self-guided exposure. However, given the low certainty of the current evidence and the absence of trials directly comparing such strategies, the optimal format, intensity, and timing of reinforcement remain undetermined and should be examined in future research before specific recommendations can be made.

Self-Efficacy

Across the selected studies, few effects were reported. Only one article showed, among its results, improvements in self-efficacy through the use of a supportive mobile app, which were observed particularly at 3 months of using said application; however, such effects were not maintained over time, with no significant improvements at the 6-month assessment [35]. In all the other interventions conducted, those that incorporated SMS delivery [38], an unmoderated peer listserve via email [32], mobile apps for symptom monitoring [41], and perioperative support [45], no differences between groups were evidenced. The foregoing, according to recent literature [55,56], could originate in the conception of the intervention design, since self-efficacy, in accordance with SCT, is sustained by mastery experiences, gradual achievements, and continuous feedback [57]; therefore, the limited active support of the intervention may affect the confidence acquired by patients, which could ultimately explain both the null effects and, at the same time, their limited persistence over time.

Anxiety, Depression, and Emotional Distress

Of the total number of articles, 13 evaluated anxiety, depression, or emotional distress, yielding mixed results [32-39,42-46]; the direction of effect favored the intervention in a minority of them (5/13) [36,37,39,44,46], and the certainty of this evidence was low. In this regard, interventions that incorporated telerehabilitation based on relaxation exercises delivered through messaging apps [39] and psychotherapy via VR [37] showed significant reductions in anxiety, depression, or distress; while mindfulness delivered by messaging generated initial changes in emotional distress; however, the results did not persist at the 1-month follow-up [36]. Favorable effects on anxiety or distress were also observed with an informative mobile app [44] and a theory-based tailored mHealth physical activity intervention [46]; however, in the latter study, anxiety did not significantly decrease within the intervention group, and no between-group difference was observed for depression. By contrast, informational or self-guided interventions delivered via the internet or a website [33-35], empowerment message delivery [38], symptom monitoring [42], and perioperative support [45] did not show differences compared with the control group, while psychological distress and mood favored the control group in one study [32] and another study reported poorer outcomes in the intervention group [43], although its small sample size limits conclusions regarding efficacy.

The foregoing could be explained by the degree of immersion and interactivity of the digital tool used, insofar as psychotherapy via VR provides a multisensory stimulus that could be capable of interrupting the anxiety cycle [58]. In this sense, these findings are consistent with previous evidence [59], as well as with relaxation videos grounded in Mayer's modality

XSL•FO
RenderX

principle [60], whereby the induction of calm through videos replaces the cognitive load associated with reading with an immersive experience that facilitates emotional regulation. In contrast, informational and monitoring platforms may be insufficient to address the affective component, while, as observed in the mindfulness intervention, one-directional support via messaging yields benefits that are not maintained in the long term, because digital mental health depends on proactive human support [61,62]. Across outcomes, the level of human support emerged as a key interpretive axis: guided interventions—those embedding professional facilitation or interaction—were more consistently associated with favorable effects than unguided, self-directed formats, in line with evidence that human guidance enhances the impact of digital health interventions [63,64].

Strengths and Limitations

The main strength of this study lies in the predominance of RCTs, which represented 14 [32-45] of the 15 included studies [32-46], together with the inclusion of one eligible controlled quasi-experimental study in accordance with the predefined eligibility criteria. The inclusion of articles in English, Spanish, and Portuguese also expanded the knowledge base. The findings allow for a systematization of the current state of the art regarding the implementation of different types of digital health tools used to support women with breast cancer and their effects compared with standard care. Furthermore, this review underscores the need for the development of more rigorous randomized trials with greater statistical power in breast cancer populations throughout the entire oncological process.

Despite the contributions of this review, several limitations must be acknowledged. First, there was substantial heterogeneity in the measurement tools used to assess outcomes, which generated discrepancies in how QoL and other psychological variables were operationalized. Additionally, there was wide variability in the types of digital health interventions used, including mobile apps, text messaging, web-based platforms, and even VR devices. This variability extended to the technological formats, content, duration, level of human support, and timing of the interventions, limiting direct comparison across studies and reducing the ability to draw generalizable conclusions. Risk of bias was also a concern; no RCT was judged to be at low overall risk, and 3 trials were judged to be at high risk, primarily because of missing outcome data. The remaining trials raised some concerns, including concerns related to subjective, self-reported outcomes assessed by participants who were aware of their intervention assignment. The quasi-experimental study was judged to have a serious overall risk of bias and was therefore interpreted cautiously.

Finally, the search did not include MEDLINE or a systematic search of clinical trial registries or gray literature, which may have limited the retrieval of relevant or unpublished studies. Furthermore, the comparison between guided and unguided interventions was exploratory and based on indirect comparisons rather than a formal subgroup analysis.

Conclusion

Digital health interventions for women with breast cancer may provide partial and mostly short-term benefits in some domains of health-related QoL and emotional well-being. Favorable effects were more frequently observed in interventions containing active therapeutic components or professional support; however, this pattern was based on indirect comparisons across heterogeneous studies. Nevertheless, several caveats temper these findings. Acceptability was formally assessed in only a few studies, so it cannot yet be considered established; effects on self-efficacy and on anxiety, depression, and emotional distress were limited and inconsistent, and were generally not maintained over time. Moreover, the certainty of the evidence was low to very low because of concerns regarding risk of bias and inconsistency across studies. Taken together, these findings should be interpreted with caution: digital health interventions are a promising adjunct in this population, but current evidence does not yet support strong or generalizable conclusions. Future research should prioritize methodologically rigorous, adequately powered trials with longer follow-up, standardized outcomes, and attention to adherence and the role of human support.

Acknowledgments

We used generative AI tools (Microsoft Copilot 365) in a limited capacity to assist with grammar, stylistic refinement, and clarity of manuscript text. We did not use AI tools to design this study, collect data, analyze data, interpret findings, or create original research content. All substantive content was drafted, critically reviewed, and verified by the authors, who take full responsibility for the integrity and accuracy of this paper.

Funding

The authors declared no financial support was received for this work.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

Conceptualization: RZ-T
Data curation: RZ-T, MB
Formal analysis: RZ-T

XSL•FO
RenderX

Methodology: RZ-T, MB, EHE
Supervision: MB, EHE
Validation: RZ-T, MB, EHE
Visualization: RZ-T
Writing – original draft: RZ-T
Writing – review & editing: RZ-T, MB, EHE

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 1971 KB - cancer_v12i1e96858_app1.docx]

Multimedia Appendix 2
Reports excluded.
[DOCX File, 540 KB - cancer_v12i1e96858_app2.docx]

Multimedia Appendix 3
Certainty of the evidence.
[DOCX File, 425 KB - cancer_v12i1e96858_app3.docx]

Checklist 1
PRISMA checklist.
[DOCX File, 2460 KB - cancer_v12i1e96858_app4.docx]

References

1. Breast cancer. World Health Organization. 2024. URL: https://www.who.int/es/news-room/fact-sheets/detail/breast-cancer [accessed 2026-07-25]
2. Bray F, Laversanne M, Sung H, et al. Global Cancer Statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
3. Breast cancer cases projected to rise by nearly 40 per cent by 2050, WHO warns. United Nations. 2025. URL: https://news.un.org/en/story/2025/02/1160391 [accessed 2026-07-25]
4. Iddrisu M, Aziato L, Dedey F. Psychological and physical effects of breast cancer diagnosis and treatment on young Ghanaian women: a qualitative study. BMC Psychiatry 2020 Jul 6;20(1):353. [doi: 10.1186/s12888-020-02760-4] [Medline: 32631279]
5. Pritzl SL, Vierkant RA, Ruddy KJ, et al. Trajectory of global mental health and global physical health in breast cancer survivorship. J Natl Compr Canc Netw 2025 Apr 14;23(5):141-146. [doi: 10.6004/jnccn.2024.7353] [Medline: 40228538]
6. Hendy A, Ibrahim RK, Darwish A, et al. Sleep disturbance, cancer-related fatigue, and depression as determinants of quality of life among breast cancer patients undergoing chemotherapy: a cross-sectional study. BMC Cancer 2025 Jul 1;25(1):1122. [doi: 10.1186/s12885-025-14538-6] [Medline: 40597987]
7. Chen L, Zheng Y, Zheng D, et al. Research trends on lymphedema after mastectomy for breast cancer patients from 2000 to 2023: a scientometric analysis. Front Oncol 2025;15:1440966. [doi: 10.3389/fonc.2025.1440966] [Medline: 39968076]
8. Rodrigues-Machado N, Bonfill-Cosp X, Quintana MJ, Santero M, Bártolo A, Olid AS. Sexual dysfunction in women with breast cancer: a systematic review. Support Care Cancer 2025 Mar 31;33(4):332. [doi: 10.1007/s00520-025-09352-6] [Medline: 40163251]
9. Yang Y, Zhao Y, Yao K, Li W, Sun H. The influence of cancer treatments on long-term psychological outcomes and quality of life in breast cancer patients. BMC Cancer 2025;25(1):1403. [doi: 10.1186/s12885-025-14667-y]
10. Jiang Y, Li H, Xiong Y, et al. Association between fear of cancer recurrence and emotional distress in breast cancer: a latent profile and moderation analysis. Front Psychiatry 2025;16:1521555. [doi: 10.3389/fpsyt.2025.1521555] [Medline: 40212837]
11. Lirola MJ, Hernández-Rodríguez AI, Cuenca-Piqueras C, Prados-Megías ME. Resilience and renewal: the personal impact of physical activity in breast cancer survivors. BMC Womens Health 2025 May 27;25(1):257. [doi: 10.1186/s12905-025-03791-3] [Medline: 40420095]
12. Paslaru AM, Plesea-Condratovici A, Moroianu LA, et al. Mind over malignancy: a systematic review and meta-analysis of psychological distress, coping, and therapeutic interventions in oncology. Medicina (Kaunas) 2025 Jun 13;61(6):1086. [doi: 10.3390/medicina61061086] [Medline: 40572774]

XSL•FO
RenderX

13. Merluzzi TV, Pustejovsky JE, Philip EJ, Sohl SJ, Berendsen M, Salsman JM. Interventions to enhance self-efficacy in cancer patients: a meta-analysis of randomized controlled trials. Psychooncology 2019 Sep;28(9):1781-1790. [doi: 10.1002/pon.5148] [Medline: 31206917]
14. Global strategy on digital health 2020-2027. : World Health Organization; 2025 URL: https://www.who.int/publications/i/item/9789240116870 [accessed 2026-08-07]
15. Shaffer KM, Turner KL, Siwik C, et al. Digital health and telehealth in cancer care: a scoping review of reviews. Lancet Digit Health 2023 May;5(5):e316-e327. [doi: 10.1016/S2589-7500(23)00049-3] [Medline: 37100545]
16. Zhang Y, Flannery M, Zhang Z, et al. Digital health psychosocial intervention in adult patients with cancer and their families: systematic review and meta-analysis. JMIR Cancer 2024 Feb 5;10:e46116. [doi: 10.2196/46116] [Medline: 38315546]
17. Marthick M, McGregor D, Alison J, Cheema B, Dhillon H, Shaw T. Supportive care interventions for people with cancer assisted by digital technology: systematic review. J Med Internet Res 2021 Oct 29;23(10):e24722. [doi: 10.2196/24722] [Medline: 34714246]
18. Pimentel-Parra GA, Soto-Ruiz MN, San Martín-Rodríguez L, Escalada-Hernández P, García-Vivar C. Effectiveness of digital health on the quality of life of long-term breast cancer survivors: a systematic review. Semin Oncol Nurs 2023 Aug;39(4):151418. [doi: 10.1016/j.soncn.2023.151418] [Medline: 37045645]
19. Chia JLL, He GS, Ngiam KY, Hartman M, Ng QX, Goh SSN. Harnessing artificial intelligence to enhance global breast cancer care: a scoping review of applications, outcomes, and challenges. Cancers (Basel) 2025 Jan 9;17(2):197. [doi: 10.3390/cancers17020197] [Medline: 39857979]
20. Zúñiga-Tapia R, Boixadós Anglès M, Hernández Encuentra E. Effectiveness of digital health interventions on quality of life in woman breast cancer patients: a systematic review. PROSPERO. 2025. URL: https://www.crd.york.ac.uk/PROSPERO/view/CRD420250578307 [accessed 2026-07-25]
21. Zúñiga-Tapia R, Boixadós-Anglès M, Hernández-Encuentra E. Digital health interventions and psychosocial outcomes in women with breast cancer: a systematic review protocol. ARS MEDICA Revista de Ciencias Médicas 2026;51(2):44-50. [doi: 10.11565/arsmed.v51i2.2229]
22. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. J Clin Epidemiol 2021 Jun;134:178-189. [doi: 10.1016/j.jclinepi.2021.03.001] [Medline: 33789819]
23. Petticrew M, Roberts H. Systematic Reviews in the Social Sciences: A Practical Guide, 1st edition: Wiley; 2006. [doi: 10.1002/9780470754887]
24. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
25. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
26. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011 Apr;64(4):401-406. [doi: 10.1016/j.jclinepi.2010.07.015] [Medline: 21208779]
27. Murad MH, Mustafa RA, Schünemann HJ, Sultan S, Santesso N. Rating the certainty in evidence in the absence of a single estimate of effect. Evid Based Med 2017 Jun;22(3):85-87. [doi: 10.1136/ebmed-2017-110668] [Medline: 28320705]
28. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
29. Lombard M, Snyder-Duch J, Bracken CC. Content analysis in mass communication: assessment and reporting of intercoder reliability. Human Comm Res 2002 Oct;28(4):587-604. [doi: 10.1111/j.1468-2958.2002.tb00826.x]
30. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368:l6890. [doi: 10.1136/bmj.l6890] [Medline: 31948937]
31. McKenzie JE, Brennan SE. Chapter 12: synthesizing and presenting findings using other methods. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 6.5 2024. URL: https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-12 [accessed 2026-08-03]
32. Salzer MS, Palmer SC, Kaplan K, et al. A randomized, controlled study of internet peer-to-peer interactions among women newly diagnosed with breast cancer. Psychooncology 2010 Apr;19(4):441-446. [doi: 10.1002/pon.1586] [Medline: 19484712]
33. Ryhänen AM, Rankinen S, Siekkinen M, Saarinen M, Korvenranta H, Leino-Kilpi H. The impact of an empowering Internet-based Breast Cancer Patient Pathway program on breast cancer patients' clinical outcomes: a randomised controlled trial. J Clin Nurs 2013 Apr;22(7-8):1016-1025. [doi: 10.1111/jocn.12007] [Medline: 23480498]
34. White V, Farrelly A, Pitcher M, Hill D. Does access to an information-based, breast cancer specific website help to reduce distress in young women with breast cancer? Results from a randomised trial. Eur J Cancer Care (Engl) 2018 Nov;27(6):e12897. [doi: 10.1111/ecc.12897] [Medline: 30137657]
35. Zhu J, Ebert L, Liu X, Wei D, Chan SWC. Mobile Breast Cancer e-Support Program for Chinese women with breast cancer undergoing chemotherapy (Part 2): multicenter randomized controlled trial. JMIR Mhealth Uhealth 2018 Apr 30;6(4):e104. [doi: 10.2196/mhealth.9438] [Medline: 29712622]

XSL•FO
RenderX

36. Wang L, Chen X, Peng Y, et al. Effect of a 4-week internet-delivered mindfulness-based cancer recovery intervention on the symptom burden and quality of life of patients with breast cancer: randomized controlled trial. J Med Internet Res 2022 Nov 22;24(11):e40059. [doi: 10.2196/40059] [Medline: 36413385]
37. Zhang X, Yao S, Wang M, et al. The impact of VR-CALM intervention based on VR on psychological distress and symptom management in breast cancer survivors. J Oncol 2022;2022:1012813. [doi: 10.1155/2022/1012813] [Medline: 35712124]
38. Singleton AC, Raeside R, Partridge SR, et al. Supporting women's health outcomes after breast cancer treatment comparing a text message intervention to usual care: the EMPOWER-SMS randomised clinical trial. J Cancer Surviv 2023 Dec;17(6):1533-1545. [doi: 10.1007/s11764-022-01209-9] [Medline: 35460441]
39. Bahçaci U, Uysal SA, Namal E. Relaxation training via tele-rehabilitation program in patients with breast cancer receiving chemotherapy during COVID-19. Digit Health 2024;10:20552076241261909. [doi: 10.1177/20552076241261909] [Medline: 39610824]
40. de Aviz LDN, Alves CF, Fonte CD, et al. Comparison of effects between telerehabilitation and in-person rehabilitation after breast cancer surgery: a randomized controlled study. Integr Cancer Ther 2024;23:15347354241256314. [doi: 10.1177/15347354241256314] [Medline: 39223789]
41. Graetz I, Hu X, Kocak M, et al. Remote monitoring app for endocrine therapy adherence among patients with early-stage breast cancer: a randomized clinical trial. JAMA Netw Open 2024 Jun 3;7(6):e2417873. [doi: 10.1001/jamanetworkopen.2024.17873] [Medline: 38935379]
42. Okuyama H, Takada F, Taira N, Nakamura S. A randomized trial of the impact of symptom monitoring using an electronic patient-reported outcome app on health-related quality of life in postmenopausal breast cancer patients receiving adjuvant endocrine therapy. Breast Cancer 2024 Sep;31(5):787-797. [doi: 10.1007/s12282-024-01592-4] [Medline: 38796818]
43. Rigg A, Kemp E, Koczwara B, et al. Feasibility, acceptability, and preliminary efficacy of a self-directed online psychosocial intervention for women with metastatic breast cancer: Finding My Way-Advanced. Support Care Cancer 2024 Oct 22;32(11):744. [doi: 10.1007/s00520-024-08924-2] [Medline: 39438337]
44. Salman Saraç F, Erkal İlhan S, Kutun S, Kutlutürkan S. The effect of informative mobile app use on anxiety, distress, and quality of life of women with breast cancer. Eur J Breast Health 2024 Jul 1;20(3):207-214. [doi: 10.4274/ejbh.galenos.2024.2024-3-9] [Medline: 39257013]
45. Pang Y, He H, Ng RP, et al. Effectiveness of an innovative mobile-based perioperative care program for women undergoing breast cancer surgery (iCareBreast): randomized controlled trial. J Med Internet Res 2025 Apr 21;27:e71684. [doi: 10.2196/71684] [Medline: 40258266]
46. Geng Z, Wang J, Lyu W, et al. Effectiveness of a theory-based tailored mHealth physical activity intervention for women undergoing chemotherapy for breast cancer: a quasi-experimental study. Int J Nurs Sci 2024 Sep;11(4):405-413. [doi: 10.1016/j.ijnss.2024.08.006] [Medline: 39830917]
47. Karyotaki E, Efthimiou O, Miguel C, et al. Internet-based cognitive behavioral therapy for depression: a systematic review and individual patient data network meta-analysis. JAMA Psychiatry 2021 Apr 1;78(4):361-371. [doi: 10.1001/jamapsychiatry.2020.4364] [Medline: 33471111]
48. Akdemir A, Smith AB, Wu VS, et al. Guided versus non-guided digital psychological interventions for cancer patients: a systematic review and meta-analysis of engagement and efficacy. Psychooncology 2024 Jan;33(1):e6290. [doi: 10.1002/pon.6290] [Medline: 38282223]
49. Zandieh S, Abdollahzadeh SM, Sadeghirad B, et al. Therapist-guided remote versus in-person cognitive behavioural therapy: a systematic review and meta-analysis of randomized controlled trials. CMAJ 2024 Mar 17;196(10):E327-E340. [doi: 10.1503/cmaj.230274] [Medline: 38499303]
50. Kerr D, Ahn D, Waki K, Wang J, Breznen B, Klonoff DC. Digital interventions for self-management of type 2 diabetes mellitus: systematic literature review and meta-analysis. J Med Internet Res 2024 Jul 22;26:e55757. [doi: 10.2196/55757] [Medline: 39037772]
51. Riadi I, Kervin L, Dhillon S, et al. Digital interventions for depression and anxiety in older adults: a systematic review of randomised controlled trials. Lancet Healthy Longev 2022 Aug;3(8):e558-e571. [doi: 10.1016/S2666-7568(22)00121-0] [Medline: 36102765]
52. Ali SH, Thu H. App fatigue in mHealth: beyond improving apps, advance equity by meeting people where they are. PLOS Digit Health 2025 Nov;4(11):e0001107. [doi: 10.1371/journal.pdig.0001107] [Medline: 41270016]
53. Heizmann AN, Chapelle C, Laporte S, Roche F, Hupin D, Le Hello C. Impact of wearable device-based interventions with feedback for increasing daily walking activity and physical capacities in cardiovascular patients: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2023 Jul 11;13(7):e069966. [doi: 10.1136/bmjopen-2022-069966] [Medline: 37433730]
54. Ueda K, Sekoguchi T, Yanagisawa H. How predictability affects habituation to novelty. PLoS One 2021;16(6):e0237278. [doi: 10.1371/journal.pone.0237278] [Medline: 34061853]
55. Luo X, Chen Y, Chen J, et al. Effectiveness of mobile health-based self-management interventions in breast cancer patients: a meta-analysis. Support Care Cancer 2022 Mar;30(3):2853-2876. [doi: 10.1007/s00520-021-06568-0] [Medline: 34561732]

XSL•FO
RenderX

56. Wildeboer G, Kelders SM, van Gemert-Pijnen JEWC. The relationship between persuasive technology principles, adherence and effect of web-Based interventions for mental health: a meta-analysis. Int J Med Inform 2016 Dec;96:71-85. [doi: 10.1016/j.ijmedinf.2016.04.005] [Medline: 27117057]
57. Bandura A, Freeman WH, Lightsey R. Self-efficacy: the exercise of control. J Cogn Psychother 1999 Jan;13(2):158-166. [doi: 10.1891/0889-8391.13.2.158]
58. Riva G, Wiederhold BK, Mantovani F. Surviving COVID-19: the neuroscience of smart working and distance learning. Cyberpsychol Behav Soc Netw 2021 Feb;24(2):79-85. [doi: 10.1089/cyber.2021.0009] [Medline: 33577414]
59. Buche H, Michel A, Piccoli C, Blanc N. Contemplating or acting? Which immersive modes should be favored in virtual reality during physiotherapy for breast cancer rehabilitation. Front Psychol 2021;12:631186. [doi: 10.3389/fpsyg.2021.631186] [Medline: 33897539]
60. Mayer RE, Fiorella L, editors. The Cambridge Handbook of Multimedia Learning, 3rd edition: Cambridge University Press; 2021. [doi: 10.1017/9781108894333.003]
61. Cuijpers P, Kleiboer A, Karyotaki E, Riper H. Internet and mobile interventions for depression: opportunities and challenges. Depress Anxiety 2017 Jul;34(7):596-602. [doi: 10.1002/da.22641] [Medline: 28471479]
62. Sommers-Spijkerman M, Austin J, Bohlmeijer E, Pots W. New evidence in the booming field of online mindfulness: an updated meta-analysis of randomized controlled trials. JMIR Ment Health 2021 Jul 19;8(7):e28168. [doi: 10.2196/28168] [Medline: 34279240]
63. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — a systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
64. Werntz A, Amado S, Jasman M, Ervin A, Rhodes JE. Providing human support for the use of digital mental health interventions: systematic meta-review. J Med Internet Res 2023 Feb 6;25:e42864. [doi: 10.2196/42864] [Medline: 36745497]

Abbreviations

CBT: cognitive behavioral therapy
DASS-21: Depression Anxiety and Stress Scale
EORTC QLQ-BR23: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Breast Cancer Module 23
EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Core 30
GRADE: Grading of Recommendations Assessment, Development and Evaluation
GSES: General Self-Efficacy Scale
HSCL-25: Hopkins Symptom Checklist-25
mHealth: mobile health
PICOC: Population, Intervention, Comparison, Outcomes, and Context
POMS: Profile of Mood States
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROMIS: Patient-Reported Outcomes Measurement Information System
QoL: quality of life
RCT: randomized controlled trial
RoB 2: Cochrane Risk of Bias tool
ROBINS-I: Risk of Bias in Non-randomized Studies of Interventions
SAS: Self-Rating Anxiety Scale
SCT: social cognitive theory
SEMCD: Self-Efficacy for Managing Chronic Disease Scale
SEMS: Self-Efficacy for Managing Symptoms
SESES-C: Stanford Emotional Self-Efficacy Scale – Cancer
SF-12: Short Form 12-Item Health Survey
SICPA: Stanford Inventory of Cancer Patient Adjustment
STAI: State-Trait Anxiety Inventory
SWiM: Synthesis Without Meta-analysis
TPB: theory of planned behavior
VR: virtual reality
WHO: World Health Organization

XSL•FO
RenderX

Edited by M Balcarras; submitted 01.Apr.2026; peer-reviewed by M Tusconi, Q Ng; revised version received 16.Jun.2026; accepted 10.Jul.2026; published 21.Aug.2026.

Please cite as:
Zúñiga-Tapia R, Boixadós M, Hernández Encuentra E
Digital Health Interventions and Psychosocial Outcomes in Women With Breast Cancer: Systematic Review
JMIR Cancer 2026;12:e96858
URL: https://cancer.jmir.org/2026/1/e96858
doi:10.2196/96858

© Rocío Zúñiga-Tapia, Mercè Boixadós, Eulàlia Hernández Encuentra. Originally published in JMIR Cancer (https://cancer.jmir.org), 21.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Use of AI to Predict and Support Medication Adherence in Patients With Breast Cancer: Systematic Review

Massimo Pezzolato[1,2*], PhD; Viktorya Voskanyan[1,2*], MSc; Ilaria Cutica[2], PhD; Chiara Marzorati[1,2], PhD; Gabriella Pravettoni[1,2], PhD

[1]Applied Research Division for Cognitive and Psychological Science, European Institute of Oncology, Via Ripamonti, 435, Milan, Italy

[2]Department of Oncology and Hemato-oncology, University of Milan, Milan, Italy

[*]these authors contributed equally

Corresponding Author:
Massimo Pezzolato, PhD
Applied Research Division for Cognitive and Psychological Science, European Institute of Oncology, Via Ripamonti, 435, Milan, Italy

Abstract

Background: Oral medications are commonly used in the treatment of breast cancer (BC), despite high rates of nonadherence. As adherence is fundamental for optimal treatment, finding ways to effectively improve it is important. Artificial intelligence (AI) is being widely applied to health care.

Objective: This review aims to offer an overview of the contribution of AI to medication nonadherence among patients with BC, suggesting future research directions, and highlighting existing gaps.

Methods: Four databases (PubMed, Embase, Scopus, and Web of Science) were searched. Studies were eligible if they used AI to predict, monitor, or support medication adherence in patients with BC, were published in English, and had full text available. Reviews, conference abstracts, editorials, and studies with mixed samples were excluded. This review was conducted following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines and was registered in PROSPERO (International Prospective Register of Systematic Reviews) database (CRD42024587020). The risk of bias was evaluated using the PROBAST (Prediction Model Risk of Bias Assessment Tool) and the Downs and Black's methodological quality scale.

Results: Ten studies were included in this review. Most of them (k=9) focused on developing machine learning models to predict medication nonadherence. These studies used a range of techniques, including logistic regression, artificial neural networks, and random forests. Model performance varied widely, with area under the receiver operating characteristic curve values ranging from 0.61 to 1.00. Predictors of nonadherence were clustered into 4 groups: clinical, disease, and treatment-related factors (eg, side effects and comorbidities); behavioral factors (eg, prior nonadherence); psychosocial factors (eg, quality of life and self-efficacy); and sociodemographic factors (eg, age and income). The only intervention study identified evaluated an AI-based chatbot and reported promising results, showing a 20% increase in adherence among participants who engaged with its reminder feature. All included studies were at high risk of bias, mainly due to the absence of model calibration or insufficient reporting, and their findings should therefore be interpreted with caution. A further limitation was the lack of attention to implementation: predictive accuracy alone is insufficient, and future work must also address actionability, safety, and cost-effectiveness to enable real clinical use. Progress in this area will require coordinated efforts among researchers, developers, clinicians, and policymakers to support the responsible development and implementation of these tools into routine care.

Conclusions: To our knowledge, this is the first systematic review of AI applications specifically targeting medication adherence in BC. It focuses on both predictive and interventional studies, mapping current AI applications within this specific clinical context. The findings highlight gaps in the implementation phase and emphasize the need for future research integrating a coordinated, multidisciplinary approach involving researchers, AI specialists, policymakers, and health care teams.

Trial Registration: PROSPERO CRD42024587020; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251123188

(JMIR Cancer 2026;12:e80128) doi:10.2196/80128

KEYWORDS

breast cancer; medication adherence; artificial intelligence; predictive modelling; endocrine therapy; oral anticancer medications

XSL•FO
RenderX

Introduction

The use of oral antineoplastic agents has rapidly increased in cancer management over the last few decades, improving overall survival rates and offering the convenience of self-administered treatment [1-3]. While oral agents can enhance cancer care by giving patients greater autonomy and control over their treatment, they also demand active participation, adding complexity to an already demanding routine. Following a cancer diagnosis, patients must manage frequent hospital visits, lifestyle adjustments such as following a healthy diet and staying physically active, and now also adhering rigorously to an oral medication regimen, introducing an additional layer of responsibility [4,5]. Often, this occurs at a time when patients are psychologically and physically drained, having to deal with fear, worries, anxiety, and treatment side effects, all of which contribute to reduced quality of life [6] and further complicate the challenge of optimal medication adherence.

Medication adherence refers to the extent to which patients take their medications as prescribed by the health care provider, including the correct timing, frequency, and dosage [7]. The phenomenon of nonadherence is common across different patient populations, with estimates suggesting that approximately 50% of adult patients fail to achieve optimal adherence [8,9].

Adherence among patients with breast cancer (BC) has been reported to range from 33.3% to 88.6%, with an average decrease of 25% from the first to the fifth year of treatment [10]. Although these estimates span a wide range, they highlight that nonadherence is a significant concern in this clinical population. This is particularly alarming given its association with higher mortality rates, poorer clinical outcomes, and increased health care costs [11,12].

A wide range of factors may influence patients' medication-taking behaviors and contribute to nonadherence, including clinical factors (related to the disease and its treatment), psychosocial and behavioral aspects, sociodemographic characteristics (eg, age, geographic area, and income), and health care system factors (eg, the quality of the patient-provider relationship) [9,13]. Among these factors, commonly reported influences on treatment adherence in patients with BC include the presence of side effects associated with oral antineoplastic agents (eg, joint pain, hot flushes, depression, and weight gain), patients' perceptions of the necessity of treatment or concerns about it, poor quality of the patient-provider relationship, and a lack of social, practical, or emotional support [13]. This overview highlights that adherence is a complex phenomenon, influenced by a diverse set of interrelated and multifaceted factors.

Given the scale of the problem and its negative consequences, improving adherence is essential. However, the various interventions proposed so far have shown only limited effectiveness, and the need for innovative solutions to predict and support medication adherence is still pressing [5,14-17]. Artificial intelligence (AI)–driven technologies, considering their significant advancements and efficiency across various domains [18,19], have the potential to play a key role in bridging this gap.

AI is an umbrella term commonly used to refer to those technologies that simulate complex human skills such as learning, creativity, decision-making, and problem-solving [20-22]. AI can be further characterized into subsets such as machine learning (ML), natural language processing (NLP), and generative artificial intelligence (GenAI) [20].

ML refers to a class of algorithms that are trained on data to automatically learn and improve their performance over time. These algorithms are typically easy to interpret, perform effectively on small to medium-sized datasets, and require relatively short implementation times. Deep learning (DL), a subset of ML, provides more advanced capacities to handle large and complex datasets. However, DL models typically require greater computational resources and involve longer training times due to their multilayered and complex architectures. Moreover, these models are more challenging to interpret compared to traditional ML approaches. Despite these constraints, DL has demonstrated notable results in tasks such as image recognition, cancer subtype classification, and clinical decision-making [23,24]. A particularly relevant application of ML and DL is predictive modeling, a process in which computational analyses are used to generate predictions about the occurrence of specific events or outcomes. ML-powered predictive models have been used in medicine and health care with promising results in areas such as clinical decision-making, disease progression prediction, and the identification of high-risk patients for early intervention [25].

NLP comprises a set of techniques used to process, understand, and sometimes generate human language [20]. NLP models are widely used in health care for tasks involving the processing of speech or text, such as sentiment analysis, screening of electronic health records, computer-aided diagnosis, and support for clinical decision-making [20,26,27].

GenAI is a subset of AI that leverages ML models and NLP techniques to generate content in various formats, such as text, speech, and images. Conversational agents (also known as chatbots) are a widely adopted application of GenAI, leveraging ML and NLP to engage in dialogue, answer questions, provide suggestions, and perform a variety of tasks. In the health care domain, they have been used to deliver patient education, support treatment and monitoring, promote medication adherence, facilitate behavior change interventions, and assist in diagnosis [20,28].

Given its significant impact across various health care domains, such as screening, diagnostics, symptom management, personalized risk prediction, clinical decision-making, and patient engagement [29-32], AI may also represent a promising tool for addressing medication adherence in patients with BC. In this context, AI can contribute to 2 distinct yet interconnected aspects of the problem. First, it may directly support patients in sustaining adequate adherence, for example, when embedded in tools such as chatbots or eHealth platforms that can provide information, reminders, and tailored interventions. Second, AI can analyze large and complex datasets more efficiently than traditional statistical approaches [33], enabling the development of predictive models that identify patients at higher risk of nonadherence who may require timely, targeted support.

XSL•FO
RenderX

Although this second contribution does not act directly on patients' behavior, it remains highly relevant. Many health care systems operate under significant resource constraints, making it unrealistic to provide adherence-supporting interventions to all patients with BC. Prioritizing those most likely to benefit is therefore crucial. Thus, this systematic review aims to offer a comprehensive overview of AI-based approaches used to address medication nonadherence among patients with BC, either by predicting the risk of nonadherence or by monitoring and directly supporting adherence. To our knowledge, this is the first systematic review covering the wide range of AI capabilities addressing nonadherence in this population. Furthermore, the review offers suggestions for future research directions, highlighting open challenges and ethical considerations.

Methods

Study Design

This systematic review was conducted following the PRISMA (Preferred Reporting Items for Systematic Literature Reviews and Meta - Analysis) guidelines [34]. The PRISMA checklist is presented in Checklist 1. This review was registered in the PROSPERO (International Prospective Register of Systematic Reviews; CRD42024587020) after conducting the initial database search. A protocol for the review was not prepared.

A meta-analysis was not performed due to substantial heterogeneity, specifically a high degree of clinical and methodological diversity across studies, including differences in study design, AI techniques, and in the operationalization and measurement of medication adherence. All these factors made the statistical pooling inappropriate and potentially misleading.

Search Strategy

Overview

A search strategy was developed and reported in accordance with the PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses literature search extension) guidelines [35]. The search was conducted by 2 authors (MP and VV) in the following databases: Scopus (Elsevier), PubMed (National Library of Medicine), Embase (Elsevier), and Web of Science Core Collection (Clarivate). No language, publication date, or other search restrictions were applied.

The keywords revolved around three core review terms: (1) adherence, (2) BC, and (3) AI (see Multimedia Appendix 1 for the complete search strategy). Studies published until February 4, 2026, were captured.

Inclusion and Exclusion Criteria

Studies were considered eligible when (1) they reported on any AI technology (eg, ML, DL, NLP, and GenAI) being used to predict, monitor, or support oral medication adherence, (2) they involved a patient with BC sample, (3) the paper was written in English, and (4) the full text was available online. Reviews, conference abstracts, and editorials were excluded. Additionally, studies with mixed populations where it was impossible to separate the results of patients with BC were excluded.

Screening

The screening was organized in Rayyan software: a web and mobile tool for reviews [36], which was also used to identify and remove duplicate records before screening. Both title or abstract and full-text screening were conducted by 2 researchers (MP and VV) in "blind mode" (ie, researchers were unaware of each other's decisions). After the initial screening (title or abstract), MP and VV read the full text of the retrieved papers that met the inclusion criteria to confirm their relevance. Discrepancies in the judgments during both title or abstract and full-text screening were sent to a third author (IC), and disagreements were settled by discussion until reaching a consensus.

Data Extraction and Summary

Data were extracted using a shared, standardized table developed jointly by MP and VV at the outset of the review, which was applied consistently across all studies. For each included paper, relevant characteristics and results were extracted and inserted into the spreadsheet. The extracted data were authors, title, year of publication, country, study design, aim, sample size, AI subtypes and algorithms, adherence measure, and results.

For studies describing the development of ML predictive models, information on predictors, model performance metrics, and types of validation was also extracted. Performance was summarized using discrimination measures, including the area under the receiver operating characteristic curve (AUC), F_1-score, concordance index (C-index), and accuracy.

In this paper, because the 2 types of evidence differ substantially, separate Results and Discussion sections were developed to present findings from studies describing the development of ML predictive models and from the single intervention study identified.

The data were extracted by 2 authors (MP and VV) and checked by 2 researchers (IC and CM). Formal interrater agreement was not assessed.

Risk of Bias Assessment

For studies describing the development of ML predictive models, the risk of bias was assessed using the PROBAST (Prediction Model Risk of Bias Assessment Tool) [37]. PROBAST evaluates the risk of bias in predictive modeling studies through 20 signaling questions distributed across 4 domains: participants, predictors, outcome, and analysis. In addition, it assesses concerns regarding the applicability of the predictive models in the domains of participants, predictors, and outcome.

Conversely, the risk of bias in the study describing the implementation of an intervention to support patients with BC was evaluated with Downs and Black's methodological quality scale [38]. This scale comprises 27 items and provides a score on 5 different methodological quality dimensions: reporting bias, external validity, internal validity, confounding - selection bias, and statistical power.

XSL•FO
RenderX

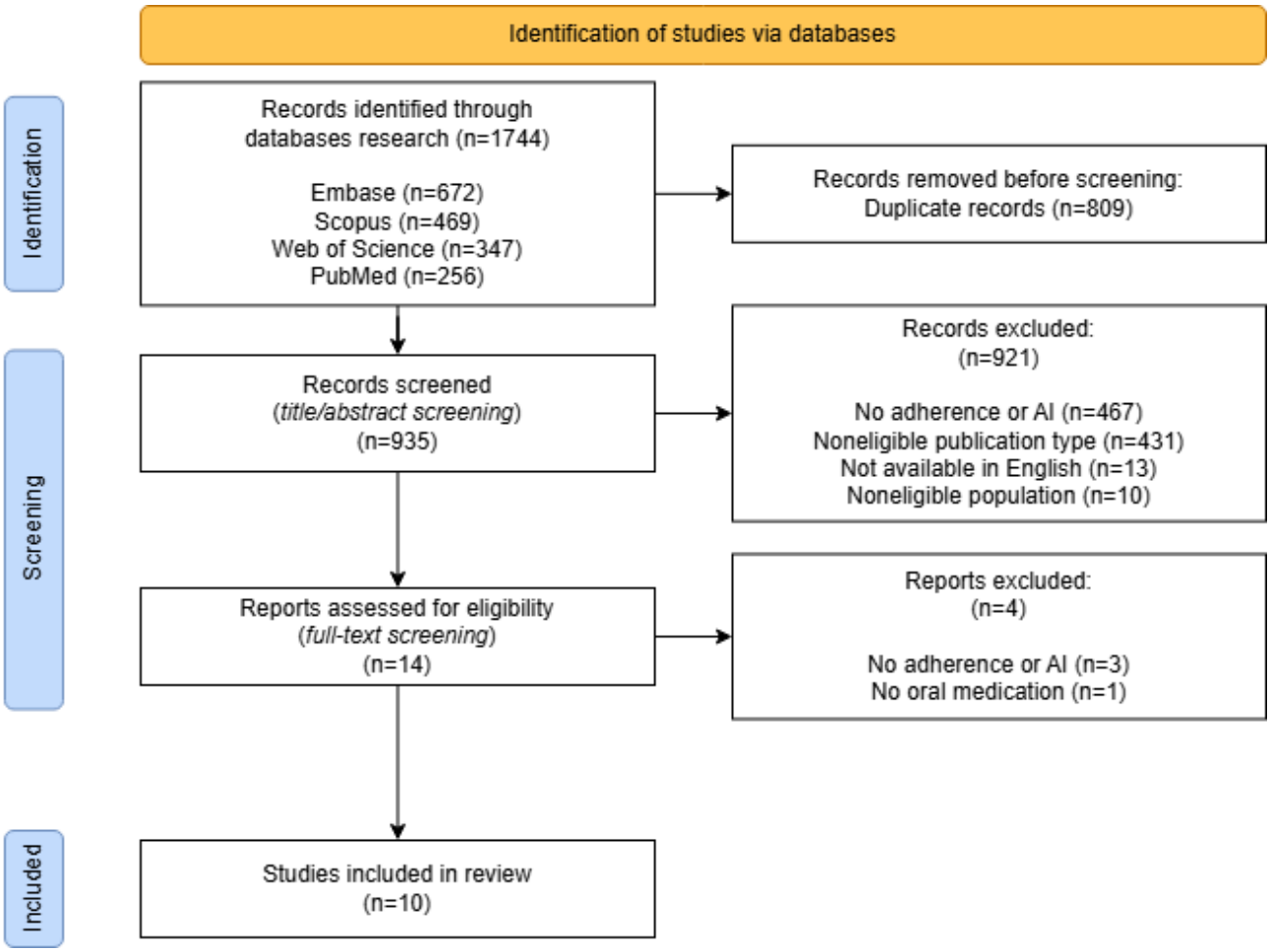
The risk of bias assessment was conducted by 2 authors (MP and VV), and disagreements were resolved by discussion until reaching a consensus.

Results

Summary of Study Characteristics

The search returned a total of 1744 papers. After duplicate removal and screening of titles and abstracts, 14 studies were considered eligible. After screening the full text, 10 of them met our inclusion criteria (see Figure 1 for a detailed flowchart of the screening process).

Figure 1. PRISMA flowchart illustrating the selection of studies included in this review. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



The included studies were published between 2019 and 2025. Most of them were conducted in the United States (k=5) and Europe (France: k=3; Italy: k=1), with 1 study carried out in Taiwan. Both early-stage BC stages (k=5) and metastatic patients (k=2) were represented in different studies, and 3 studies did not report patients' cancer stages. The sample sizes of the studies ranged from 18 to 229,695.

Adherence was measured using administrative health data or medical records (k=4), self-report measures (k=3), and medication event monitoring systems (k=3).

A total of 9 of the included studies described the development of ML predictive models (7 retrospective cohort studies, 1 prospective study, and 1 randomized controlled trial protocol), while only 1 prospective study described the implementation of an AI-powered intervention to support oral medication adherence.

The key characteristics of the included studies are summarized in Table 1, while Table 2 provides an overview of the AI-based predictive models and adherence support intervention.

XSL•FO
RenderX

Table . Key characteristics of the included studies, including author, year, title, country, design, study aim, sample size, and adherence measures.

Author	Year	Title	Country	Design	Aim	Sample size	Adherence measure
Balazard et al [39]	2023	Adjuvant endocrine therapy uptake, toxicity, quality of life, and prediction of early discontinuation.	France	Retrospective cohort study	To develop a prediction model of early ET[a] discontinuation.	BC[b], stage I-III (descriptive analyses, n=6488; model development, n=5282; training set, n=4225; validation set, n=1057)	Patients' self-declarations recorded in electronic case report forms.
Chaix et al [40]	2019	When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot.	France	Prospective study	To evaluate 1 year of conversations between patients with BC and a chatbot.	BC, stage N/A[c] (n=958; adherence measurement, n=33)	Answers to the chatbot's reminder messages were used to measure adherence rates.
Kaur et al [41]	2025	A computational framework for longitudinal medication adherence prediction in breast cancer survivors: a social cognitive theory based approach.	United States	Retrospective cohort study	To predict daily and weekly adherence by integrating static and dynamic factors.	BC, stage I-III (n=32)	Medication event monitoring system.
Kaur et al [42]	2021	Theory-guided randomized neural networks for decoding medication-taking behavior.	United States	Retrospective cohort study	To predict daily medication-taking behavior over 3 time points (baseline, 4 months, and 8 months).	BC, stage I-III (n=32)	Medication event monitoring system.
Kuo et al [43]	2022	Using data mining technology to predict medication-taking behaviour in women with breast cancer: a retrospective study.	Taiwan	Retrospective cohort study	To predict long-term AHT[d] adherence and persistence.	BC, stage n.a. (n=385)	Clinical registry and medical records (MPR[e]>80%).
Masiero et al [44]	2023	A machine learning model to predict patients' adherence behavior and a decision support system for patients with metastatic breast cancer: protocol for a randomized controlled trial.	Italy	RCT[f] protocol	To evaluate a decision support system and an ML[g] predictive model, and to refine it with a set of new variables.	mBC[h] (RCT, n=100; retrospective data for the predictive model, n=2750)	Weekly medication diaries and self-report measures.

XSL•FO
RenderX

Author	Year	Title	Country	Design	Aim	Sample size	Adherence measure
Rattsev et al [45]	2024	Incorporation of emergent symptoms and genetic covariates improves prediction of aromatase inhibitor therapy discontinuation.	United States	Retrospective cohort study	To build a survival ML model for the prediction of aromatase inhibitor discontinuation and to explore how the predictive value of various risk factors changes over time.	BC, stage 0-III (n=181)	Medical records (therapy discontinuation).
Rinder et al [46]	2024	Dynamic projection of medication nonpersistence and nonadherence among patients with early breast cancer.	France	Retrospective cohort study	To model persistence and adherence to oral anticancer treatment.	eBC[i] (n=229,695)	Administrative health data (MPR>80%).
Yerrapragada et al [47]	2021	Machine learning to predict tamoxifen nonadherence among US commercially insured patients with metastatic breast cancer.	United States	Retrospective cohort study	To predict tamoxifen nonadherence in the first year after treatment initiation using ML algorithms to identify predictive factors, and to evaluate administrative data for the development of a nonadherence screening tool.	mBC (n=3022)	Administrative health data (MPR>80%).
Yuan et al [48]	2025	Multimodal sensing and modeling of endocrine therapy adherence in breast cancer survivors.	United States	Prospective study	To predict daily medication adherence using multimodal data.	BC, stage I-III (descriptive statistics, n=20; analysis, n=18)	Medication event monitoring system.

[a]ET: endocrine therapy.

[b]BC: breast cancer.

[c]N/A.: not available.

[d]AHT: adjuvant hormonal therapy.

[e]MPR: medication possession ratio.

[f]RCT: randomized controlled trial.

[g]ML: machine learning.

[h]mBC: metastatic breast cancer.

[i]eBC: early-stage breast cancer.

XSL•FO
RenderX

Table . Summary of AI[a]-based predictive models and adherence-support interventions, detailing technologies and algorithms used, predictors, performance metrics, validation approaches, intervention features, and main results. Higher values of AUC[b], accuracy, and C-index[c] indicate better predictive performance.

Author	Year	AI technologies/algorithms	Predictors	Performance metrics	Validation	Intervention	Results
Balazard et al [39]	2023	ML[d]/L2-penalized Cox regression, gradient boosted trees with a Cox partial likelihood	Discontinuation of ET[e], QoL[f], physical functioning, role functioning, social functioning, emotional functioning, cognitive functioning scores, mastectomy and radiotherapy, fatigue, pain, dyspnea, insomnia, systemic therapy side effects, breast symptoms, and arm symptoms.	C-index (0.60 - 0.64).	Internal (nested cross-validation and held-out validation set)	N/A[g].	Patients who have already discontinued an ET after primary treatment were more likely to discontinue ET definitively; higher QoL, physical functioning, role functioning, social functioning, emotional functioning, cognitive functioning scores, and receiving mastectomy and radiotherapy were associated with lower rates of early ET discontinuation; fatigue, pain, dyspnea, insomnia, systemic therapy side effects, breast symptoms, and arm symptoms were associated with a higher rate of early ET discontinuation.
Chaix et al [40]	2019	ML/N/A	N/A.	N/A.	N/A	A chatbot that interacts with users by simulating a human conversation through text was designed to empower patients and answer their questions with personalized insights.	The overall satisfaction was 93.95% (900/958). The average compliance of patients using the medication reminder feature improved by more than 20%.
Kaur et al [41]	2025	DL[h]/long short-term memory, feed-forward neural networks, SHAP[i]	Medication-taking behavioral patterns, symptom burden, perceived susceptibility, decision regret, and medication-taking self-efficacy.	Accuracy (87.25; 76.04), precision (92.04; 75.83), recall (93.15; 85.8), and specificity (77.5; 72.3).	Internal (nested cross-validation)	N/A.	The models showed strong predictive performance, outperforming traditional ML models across all evaluated performance metrics in the daily model, and in terms of accuracy, recall, and specificity in the weekly model.

XSL•FO
RenderX

Author	Year	AI technologies/algorithms	Predictors	Performance metrics	Validation	Intervention	Results
Kaur et al [42]	2021	ML/randomized neural networks	Survey data regarding BC[j] survivorship knowledge, medication attitudes, medication self-efficacy, QoL, and other factors.	Overall accuracy and individual accuracy (0.92<OA[k]<0.97).	N/A	N/A.	The model outperforms existing computational models in terms of prediction accuracy under conditions of randomness. It was able to predict the randomness of the subjective values and decision rules that contribute to the dynamics of patients' medication-taking behavior.
Kuo et al [43]	2022	ML/decision tree, artificial neural networks	The top 5 influencing factors were duration of AHT[l] discontinuation, duration of AHT use, age at diagnosis, BMI, and receipt of radiotherapy.	Overall accuracy, specificity, sensitivity, and AUC (0.90 - 1.00).	N/A	N/A.	All 3 models achieved high classification performance. Multiple logistic regression was the most effective approach (accuracy: 96.37%; specificity: 96.75%; sensitivity: 96.12%).
Masiero et al [44]	2023	ML, NLP[m]/Shapley values	Sociodemographic variables, diagnosis, biochemical and medical tests, procedures and medical history, treatment programs, treatment side effects, comorbidities, and familiarities, initiation of the treatment interruption, and skipped treatment doses. The RCT[n] will allow the addition of the following variables as predictors to the predictive model: personality traits, self-efficacy for coping with cancer, sense of coherence, pain, anxiety, depression, risk perception, and QoL.	AUC, precision, recall, sensitivity, specificity, κ, and positive and negative predictive values.	N/A	Participants in the intervention group will have access to a web-based decision support system, and an ML-based application will be used in shared decision-making sessions.	N/A.

XSL•FO
RenderX

Author	Year	AI technologies/algorithms	Predictors	Performance metrics	Validation	Intervention	Results
Rattsev et al [45]	2024	ML/Cox proportional hazards, random survival forest, penalized Cox proportional hazards, gradient boosted models, and SHAP	Genetic markers, disease stage, sleep disturbance, anxiety, depression, previous taxane therapy, fatigue, age, self-reported adherence, endocrine symptoms, physical function, BMI, and joint pain.	Mean AUC (0.57 - 0.66), integrated Brier score, time-dependent AUC, sensitivity, specificity, and positive predictive values.	Internal (cross-validation)	N/A.	The results indicated that integrating genomic and longitudinal features enhances the model's predictive performance, thereby potentially facilitating the identification of patients at risk of early discontinuation.
Rinder et al [46]	2024	DL/gated-recurrent unit network, feed-forward neural network, and SHAP	Age (>70 y), past nonadherence, taking more than 1 treatment in the previous 3 months, and low income were predictors of nonpersistence. The predictors of nonadherence were similar, adding age (<50 y) and irregular intervals in treatment purchases.	AUC (0.71 - 0.73).	Internal (held-out validation set)	N/A.	The model was able to estimate the risk of nonpersistence and nonadherence among French female patients with localized BC and identified that some factors (age, past behavior, and number of treatments) may be associated with the risk of nonpersistence and nonadherence.
Yerrapragada et al [47]	2021	ML/logistic regression, boosted logistic regression, random forest, feed-forward neural network models	Patient features (55 y or older at treatment start, receiving care in the South, being a surviving spouse), pretreatment procedures (lymphatic nuclear medicine, radiation oncology, arterial surgery, microbiology, and imaging) or therapy (β-blockers, antidepressants, and stimulants), and baseline comorbid diagnoses (upper respiratory disease, breathing abnormality, dorsopathy, neurotic disorders, cerebrovascular disease, abdominal pain, and genital disorders).	AUC (0.61 - 0.64), F_1-score (0.59 - 0.62), sensitivity, and specificity.	Internal (cross-validation)	N/A.	All models had poor predictive accuracy. Logistic regression (AUC=0.64) was interpreted with 94% sensitivity (95% CI 89 to 92) and 0.31 specificity (95% CI 29 to 33). The model accurately classified adherence (negative predictive value 89%) but failed in discriminating nonadherence (positive predictive value 48%).

XSL•FO
RenderX

Author	Year	AI technologies/algorithms	Predictors	Performance metrics	Validation	Intervention	Results
Yuan et al [48]	2025	DL/random forest, XGBoost[o], light gradient boosting machine, long short-term memory, SHAP	Recent adherence history, subjective reports of mood, physical side effects, social support, context of medication-taking behaviors, physical activity, sleep, cardiovascular function, demographic data, and static psychosocial traits.	Macro-averaged balanced accuracy (0.84), macro-averaged F_1-score (0.88).	Internal (nested cross-validation)	N/A.	Results indicated the feasibility of using multimodal sensing data to predict daily adherence with moderate accuracy. Models that integrated multimodal data consistently outperformed those relying on a single modality.

[a]AI: artificial intelligence.
[b]AUC: area under the receiver operating characteristic curve.
[c]C-index: concordance index.
[d]ML: machine learning.
[e]ET: endocrine therapy.
[f]QoL: quality of life.
[g]N/A: not available.
[h]DL: deep learning.
[i]SHAP: Shapley Additive Explanations.
[j]BC: breast cancer.
[k]OA: overall accuracy.
[l]AHT: adjuvant hormonal therapy.
[m]NLP: natural language processing.
[n]RCT: randomized controlled trial.
[o]XGBoost: Extreme Gradient Boosting.

Predictive Models

Most of the included studies (n=9) described the development of ML predictive models [39,41-48]. By feeding ML algorithms with diverse types of patient data, these models aim to offer predictions about which patients are more likely to become nonadherent or to face issues related to medication-taking along their care journey.

Only 2 studies [39,45] developed regression models, using a numerical outcome of interest: the time from treatment initiation to discontinuation. All other studies focused on classification models, defining adherence as a categorical variable and aiming to predict which patients may experience difficulties with medication adherence.

Overall, the performance metrics (ie, AUC, C-index, F_1-score, and accuracy) ranged from poor to excellent (see Table 3 for a summary of the model performance, excluding post hoc analyses).

XSL•FO
RenderX

Table . Performance metrics of the predictive models evaluated in this study, excluding post hoc analyses.

First author, year	Performance metrics
Balazard et al [39], 2023	C[a]=0.60 - 0.64
Kaur et al [41], 2025	A[b]=87.25 (daily model); A=76.04 (weekly model)
Kaur et al [42], 2021	OA[c]=0.92 - 0.97
Kuo et al [43], 2022	AUC[d]=0.90 - 1.00
Masiero et al [44], 2023	N.a.[e]
Rattsev et al [45], 2024	AUC=0.71(6-mo); AUC=0.67 (12 mo; mean time-dependent AUC=0.65)
Rinder et al [46], 2024	AUC=0.71 - 0.73
Yerrapragada et al [47], 2021	AUC=0.61 - 0.64; F[f]=0.59 - 0.62
Yuan et al [48], 2025	BA[g]=0.84, F=0.88[h]

[a]C: concordance index.
[b]A: accuracy.
[c]OA: overall accuracy.
[d]AUC: area under the receiver operating characteristic curve.
[e]N.a.: not available.
[f]F: F_1-score.
[g]BA: macro-averaged balanced accuracy.
[h]Macroaveraged F_1-score.

Several diverse ML algorithms were implemented. Balazard et al [39] evaluated the performance of 2 complementary ML models: L2-penalized Cox regression and gradient boosted trees with a Cox partial likelihood. The models showed similar performance with C-indexes ranging from 0.60 to 0.64. They were first validated internally using nested cross-validation on the training set, and the best-performing model was then evaluated on the held-out validation set.

Yerrapragada et al [47] trained and evaluated 4 models: logistic regression, boosted logistic regression, random forest, and feed-forward neural network. The models were internally validated using 10-fold cross-validation and showed poor performance with AUC ranging from 0.61 to 0.64 and F_1-scores ranging from 0.59 to 0.62 [49]. The authors then applied a post hoc balancing technique, the synthetic minority oversampling technique; the evaluation after the refitting showed an improvement in the performance of the random forest model (AUC 0.93, F_1-score 0.78) and the feed-forward neural network model (AUC 0.79, F_1-score 0.71).

Kuo et al [43] developed an ensemble learning scheme by applying 3 ML models: multiple logistic regression, decision trees, and artificial neural networks. All models demonstrated high performance, with AUC values ranging from 0.90 to 1.00 and accuracy between 92.29% and 96.37%. However, the authors did not specify whether any validation method was applied to the models.

Kaur et al [42] designed a randomized neural network and compared its prediction performance (overall accuracy and individual accuracy) with that of a traditional neural network. The randomized neural network model outperformed the traditional one, achieving high accuracy levels (>0.92) both overall and at the individual level. Although the paper did not report any internal or external validation, the authors indicated their intention to perform leave-one-subject-out cross-validation to properly evaluate the developed model.

The same research group subsequently developed other DL models combining long short-term memory and feed-forward neural networks [41]. Specifically, they built a daily model to predict patients' medication-taking behavior for the following day and a weekly model to predict adherence over the following week. When compared with traditional ML approaches (ie, random forest, support vector machine, XGBoost [Extreme Gradient Boosting], and logistic regression), the proposed models demonstrated superior performance. For the daily and weekly models, respectively, accuracy reached 87.25 and 76.04, precision 92.04 and 75.83, recall 93.15 and 85.8, and specificity 77.5 and 72.3. With regard to internal validation, the study adopted a nested 5-fold cross-validation approach. In addition, the contribution of individual predictors was examined using SHAP (Shapley Additive Explanations) values.

Rinder et al [46] developed a DL predictive model by combining a recurrent neural network and a feed-forward neural network. The model demonstrated satisfactory predictive performance, with an AUC of 0.71 for persistence and 0.73 for adherence. Internal validation was performed using a 60:20:20 train-validation-test split. Once again, SHAP values were analyzed to gain insights into the specific factors' contributions to the model's outcomes.

Masiero et al [44] developed ML predictive models currently undergoing evaluation in a randomized controlled trial involving patients with metastatic BC. The predictive models were developed by applying NLP techniques to an electronic health

XSL•FO
RenderX

record dataset, and they will be further optimized by feeding newly collected data during the randomized controlled trial. The models' predictive performance was evaluated with several metrics (AUC, precision, recall, sensitivity, specificity, κ, and positive and negative predictive values), but the results of the evaluation are not reported. The predictive models use Shapley values for interpretability, allowing the evaluation of how much each specific risk factor contributes to the given prediction.

Yuan et al [48] developed a 2-tier DL ensemble to predict daily adherence. At tier 1, long short-term memory models and tree-based regressors (ie, random forest, XGBoost, and LightGBM [light gradient boosting machine]) were trained. At tier 2, the outputs of these models were integrated using an optimized soft-voting strategy. The resulting multimodal model predicted adherence more accurately than single-modality models, achieving a balanced accuracy of 0.84 and an F_1-score of 0.88. The model was validated using leave-one-participant-out cross-validation, and its predictors were interpreted using SHAP analysis.

Finally, Rattsev et al [45] developed a survival ML model to predict time-to-discontinuation of aromatase inhibitors. The predictive framework was based on 4 survival ML algorithms: Cox proportional hazards, random survival forest, penalized Cox proportional hazards, and gradient boosted models. The results indicated that incorporating genetic markers and follow-up data improved model performance; however, overall discrimination remained modest, with a mean time-dependent AUC of 0.65 and AUC values of 0.71 and 0.67 at 6 and 12 months, respectively. As in the previous studies, SHAP values were used to examine the relative importance of individual features.

Predictors of Nonadherence

The studies describing the development of predictive models using ML identified several predictors of nonadherence. These can be categorized into four groups: (1) clinical, disease-, and treatment-related factors; (2) behavioral factors; (3) psychosocial factors; and (4) sociodemographic factors.

Among clinical, disease-related, and treatment-related factors, the presence of side effects or symptoms [39,41,45,48], complex drug regimens [46,47], comorbidities [47], and low BMI [43] were identified as predictors of nonadherence. Likewise, an increase in BMI was associated with decreased risk of discontinuation [45]. In addition, prior taxane therapy was associated with a higher risk of discontinuation, whereas specific genetic markers and more advanced disease stages were associated with a lower risk [45]. Contradictory results were identified regarding some treatment-related variables. One study found that a history of pretreatment radiotherapy or diagnostic imaging procedures was associated with higher nonadherence [47], whereas another study reported that a recent mastectomy or radiotherapy was associated with increased adherence [39].

Behavioral factors associated with nonadherence included a history of previous nonadherence [39,46], a longer time since treatment initiation, and longer durations of treatment discontinuation episodes [43]. In line with these findings, higher levels of self-reported adherence were associated with a lower risk of discontinuation [45]. Moreover, recent medication-taking patterns were also relevant: adherence on the previous days or during the preceding weeks, as well as consistently taking the medication at the same time each day over the course of a week, was predictive of subsequent adherence [41,45,48].

Psychosocial factors associated with higher adherence included better quality of life [39,42], improved physical, role, social, emotional, and cognitive functioning [39], and greater social support Yuan et al [48]. Moreover, patients' knowledge, attitudes toward medicines [42], self-efficacy in managing medication [41,42], perceived susceptibility to cancer recurrence, and decision regret were identified as predictors of medication-taking behavior [41]. In contrast, baseline depression and sleep disturbance [45], as well as self-reported sadness [48], were associated with a higher predicted risk of nonadherence. Findings regarding anxiety were inconsistent: Rattsev et al [45] reported that higher baseline anxiety was associated with a lower risk of discontinuation, whereas Yuan et al [48] found that higher anxiety was associated with an increased risk of nonadherence.

Finally, among sociodemographic factors, age was the most commonly reported predictor of adherence. Nonadherence was found to be more prevalent among patients at the extremes of the age spectrum, affecting both younger (age <50 y) [43,46] and older individuals (age >65-70 y) [43,45-47]. In addition, lower income [46], being a surviving spouse, and receiving treatment in certain geographic areas [47] were also associated with nonadherence.

Intervention Supporting Medication Adherence

Only one of the included studies described the implementation of an intervention that specifically aimed at supporting adherence by patients with BC to oral medication [40]. The authors developed an ML-based chatbot that provides personalized information via text and supports treatment management. Patients found the chatbot helpful and supportive (with an overall satisfaction rate of 93.95%) and engaged with it frequently. One of its features was a reminder function that patients could easily activate. Although 958 patients participated in the study, only 33 (3.44%) activated the reminder function for a sufficient duration to allow adherence analysis. The authors reported that these patients showed a significant increase in adherence, with a 20% increase in compliance (P=.04) among those who used the reminder function.

Risk of Bias

A summary of the risk of bias and applicability assessment is presented in Table 4.

XSL•FO
RenderX

Table . Risk of bias assessment of the included prediction models using the PROBAST[a].

First author, year	Applicability concerns							Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB[b]	Applicability
Balazard et al [39], 2023	Low	Low	Low	High	Low	High	Low	High	High
Kaur et al [41], 2025	Low	Low	Low	High	Low	High	Low	High	High
Kaur et al [42], 2021	Unclear	Unclear	Low	High	Low	Unclear	Low	High	Low
Kuo et al [43], 2022	Low	Low	Low	High	Low	High	Low	High	High
Masiero et al [44], 2023	Low	Unclear	Unclear	High	Low	Low	High	High	High
Rattsev et al [45], 2024	Low	Low	Low	High	Low	High	Low	High	High
Rinder et al [46], 2024	Low	Low	Low	High	Low	High	Low	High	High
Yerrapragada et al [47], 2021	Low	Low	Low	High	Low	Low	Low	High	Low
Yuan et al [48], 2025	Low	Low	Low	High	Low	High	Low	High	High

[a]PROBAST: Prediction Model Risk of Bias Assessment Tool.

[b]RoB: Risk of Bias.

All included studies reporting on ML predictive models demonstrated a high overall risk of bias. This classification followed the PROBAST assessment criteria, according to which a high risk in any single domain results in an overall high-risk rating. In this case, the overall high-risk classification was driven by the fact that all studies showed a high risk of bias in the analysis domain. This domain includes the assessment of analytical methods and statistical considerations, as inaccuracies in applying these components may lead to bias. Specifically, the included studies do not meet the prerequisites for a comprehensive evaluation of model performance, as they do not report calibration plots or tables, nor do they assess calibration using standard methods such as the Hosmer-Lemeshow test. According to PROBAST, both discrimination and calibration must be reported when assessing a predictive model. Discrimination reflects how well a model distinguishes between cases, but without calibration, the predicted probabilities may be unreliable, potentially resulting in inappropriate clinical decisions [37,50]. Rattsev et al [45] represent an exception, as they assessed model calibration using the Integrated Brier Score. Nevertheless, this study was also judged to be at high risk of bias in the analysis domain, given the low number of events per variable, which increases the risk of overfitting.

Other studies raised applicability concerns. It is important to clarify that the applicability section of PROBAST does not refer to the general applicability of the models, but rather to their relevance to the review question defined before conducting the PROBAST assessment [37]. In this case, the review question underlying the assessment concerns ML predictive models developed to predict adherence to oral medication in patients with BC. The models developed by Masiero et al [44] raised concerns in the outcome domain, as they do not directly predict adherence. Instead, they focus on two outcomes that are reportedly associated with adherence and are used to infer potential adherence issues: (1) short- and long-term side effects, and (2) physical status and comorbid conditions. A total of 6 additional studies raised concerns in the predictors domain [39,41,43,45,46,48]. All of these included some form of prior medication-taking behavior as a predictor in their models. However, by definition, such variables are not available at the start of treatment. This raises applicability concerns if the model is intended to be used at treatment initiation, as its predictions would rely heavily on predictors that are not yet available.

Finally, it is important to note that several specific items, and consequently some domains, were rated as “unclear” due to insufficient reporting of relevant information.

The methodological quality of the only included study describing an intervention to support medication adherence [40] was assessed using Downs and Black’s methodological quality scale [38], as PROBAST is not designed for evaluating studies that do not develop or validate predictive models. As reported in Multimedia Appendix 2, the prospective study scored low across all methodological quality dimensions of the scale, indicating a high risk of bias. In particular, the low score in the “reporting” domain (5/11) was mainly due to the inadequate reporting of patient characteristics, the distribution of key confounders, and some of the study findings. The lack of information about the participants resulted in a low score (0/3)

XSL•FO
RenderX

for external validity, as it was impossible to answer the scale's items. The internal validity score (3/7) was penalized due to the inherent nature of the intervention, which made it impossible to blind participants or researchers, and because the component related to medication adherence showed very low compliance. Regarding selection bias, the low score (1/6) was primarily attributed to the absence of randomization in the study design, and once again, to the lack of relevant information about the patients and potential confounders. Finally, the power of the analysis concerning patient adherence was deemed inadequate (score: 0/5), given the small sample of participants for whom adherence could be assessed.

Discussion

Principal Findings

This systematic review provides a comprehensive overview of how AI-based technologies are currently being used to address medication adherence among patients with BC. Across the 10 included studies, the predominant application of AI was the development of predictive models, as all but 1 study focused on this purpose. By synthesizing their findings, we were able to identify the main groups of predictors associated with nonadherence and to report the key characteristics and performance of each model. The only study that described an intervention was also examined, and its results were summarized. This process made it possible to highlight the limitations affecting current AI applications in this area, especially in light of the thorough risk of bias assessment performed.

Drawing on the findings synthesized, several observations regarding the current landscape and suggestions for desirable future developments can be made.

To begin with, the small number of publications identified (k=10), all published between 2019 and 2025, suggests that the application of AI technologies to predict or support medication adherence is both limited and relatively recent. Nonetheless, AI-driven models appear to hold promise in predicting medication-taking behaviors, identifying key determinants of adherence, and estimating the risk of nonadherence with increasing accuracy [41-43,46,48]. However, the analysis of the findings highlights persisting limitations and unresolved challenges, underscoring the need for further research to fully unlock the potential and enhance the impact of these technologies.

Predictive Models

We will begin by discussing the insights and observations emerging from the analyses of ML-based predictive models. First of all, the sample sizes of the included studies, except for the study by Rinder et al [46] using a sample size of 229,695, are similar to those used in predictive models using traditional statistical methods. This is a prominent point showing that AI's potential is not yet fully realized. The ability to process large datasets is one of AI's core strengths and could significantly enhance both the accuracy and robustness of predictive models [51,52]. A possible explanation for the sample sizes observed in the retrieved studies may lie in the need to strike an appropriate balance between sample size and the quality and relevance of the data available. Indeed, predictive models based on large datasets that include mainly irrelevant variables may still yield statistically significant results due to their size, yet prove clinically uninformative. It is crucial, therefore, not to conflate statistical significance with clinical relevance [53].

Furthermore, many of the predictors of medication nonadherence identified by AI-based models had already emerged in previous research using traditional statistical methods or qualitative approaches [10,54,55]. This convergence is to be expected, as the key contribution of AI does not lie in the discovery of new factors, but rather in enhancing analytical precision and uncovering complex interactions among multiple variables.

The predictive performance of AI models has been assessed using a range of metrics, including accuracy, sensitivity, specificity, AUC, the C-index, and the F_1-score. However, performance metrics alone may not be particularly informative and should be considered alongside additional parameters to ensure the accuracy of the predictions and guarantee their clinical usefulness [56,57]. For example, 4 studies that demonstrated "excellent" performance [41-43,48] had relatively small sample sizes (ie, n=18, n=32, n=32, and n=385), which could make the models less reliable. Moreover, even the most accurate predictive model holds little value if it lacks clinical actionability, namely, if its predictions are not translated into concrete interventions, or if its implementation is not cost-effective [56]. These crucial aspects were scarcely, if at all, addressed in the studies reviewed. Considering these aspects, the performance of the models should be interpreted with caution.

According to the PROBAST evaluation, all included studies exhibited a high overall risk of bias; this result was primarily driven by a high risk of bias in the analysis domain. In fact, although the models' discrimination performance was reported, the lack of calibration assessment exposes the predictive models to potential bias [37]. Discrimination indicates how well a model differentiates between cases, while calibration reflects how closely the predicted probabilities match the actual observed outcomes. For example, a poorly calibrated model might overestimate risk, leading to unnecessary interventions, or underestimate it, thereby missing high-risk patients who require support. This represents a major limitation, as models that are not calibrated may produce unreliable probability estimates, potentially leading to inaccurate clinical decisions when implemented in practice [50]. Underreporting of calibration in predictive model studies is a known phenomenon and has been observed across various research areas. For instance, a systematic review on predictive models for cardiovascular diseases identified that only 36% (259/717) of the studies reported calibration [58].

Another key limitation in the predictive model studies is the absence of external validation, undermining the models' generalizability to other clinical settings, where differences in patients' characteristics, care pathways, or data quality may significantly affect model performance. While 6 studies reported internal validation procedures, such as cross-validation or the use of a held-out validation set [39,41,45-48], none incorporated

XSL•FO
RenderX

external validation. External validation involves testing the performance of the predictive model on an independent dataset to assess whether its performance is consistent and generalizable to different populations. Although this procedure could be unnecessary when large, representative samples are used, it is fundamental when models are built on smaller samples, as was the case in some of the included studies [59]. Collins et al [59] recommend that during the development of a predictive model, all available data should be used, alongside appropriate internal validation procedures. They also suggest that external validation should be conducted in subsequent studies. We were unable to identify any external validation studies for the included predictive models; however, we cannot exclude the possibility that such studies are ongoing.

Another important aspect not addressed in the included predictive model studies is implementation. The absence of reports on clinical implementation highlights a significant gap between model development and real-world application. This phenomenon is not new; the implementation gap of ML in health care has been widely acknowledged and discussed in the scientific literature [56,60]. Possible strategies to bridge this gap will be further discussed in the Clinical Implications section.

In summary, the limitations observed in the included studies describing predictive models, such as the lack of external validation, assessment of calibration, and implementation strategies, may compromise their reliability and real-world applicability. To fully realize the potential of AI in predicting medication adherence, further research is needed, with a stronger focus on methodological robustness and clinical integration.

AI-Powered Interventions

This systematic review was designed to capture not only the role of AI in predicting patients' medication adherence, but also its use in supporting patients and helping them develop or maintain optimal adherence. Notably, only 1 study describing an AI-powered intervention to support adherence was identified [40]. This finding is informative in itself, potentially highlighting a research area that remains underexplored and could benefit from further investigation.

Moreover, the included intervention study gives rise to several observations. Although supporting medication adherence was not the sole aim of the chatbot developed by the authors, the study reported positive results specifically related to this outcome, suggesting that AI-powered chatbots can effectively support patients with BC by helping them improve adherence. However, this result should be interpreted with caution, as only a small percentage of participants (33/958, 3.44%) used the specific reminder feature. This limited uptake raises important questions about its practical effectiveness in clinical settings and its real potential to engage patients in behavioral change. The low scores obtained in the risk of bias assessment, particularly regarding the reporting of participant characteristics, potential confounders, and the clear presentation of study findings, underscore the need to interpret the results with caution. Further research is needed to confirm the chatbot's effectiveness, with a stronger focus on patient engagement, especially concerning the adherence support component. Future studies should ideally include a control group to better isolate the intervention's impact on outcomes.

A final, important point to consider is the safety of this AI tool. The chatbot described in the study also provides information to patients and answers their questions. It is therefore crucial that the information delivered is accurate and reliable, as misinformation could lead to potentially dangerous and harmful consequences. Interestingly, the same chatbot was evaluated in a blinded, randomized controlled noninferiority trial, comparing the quality of the information it provided with that of physicians. The results indicated that the quality of the information provided by the chatbot was noninferior to that provided by physicians [61].

Ethical and Safety Considerations

Ensuring safety remains a central concern in the development and use of AI-driven tools and predictive models [56]. The limited number of implementation studies makes it difficult to evaluate whether these models perform safely and reliably in real-world settings, especially when used to inform clinical decisions. This gap in evidence represents a major barrier to assessing potential risks and unintended consequences.

Privacy and data security represent equally critical issues, given that ML models rely on sensitive health information that is vulnerable to breaches and cyberattacks [62]. Recent frameworks, such as the integrated security and ethics model, have attempted to offer structured guidance by integrating principles of transparency, accountability, fairness, and privacy protection [63].

Another key challenge is algorithmic bias, often stemming from unrepresentative training data or inappropriate model generalization [20,64]. National agencies have proposed guiding principles to help stakeholders mitigate these risks, emphasizing equity, transparency, patient engagement, and accountability across the entire algorithm lifecycle [20,65].

Finally, concerns about transparency persist, as many ML models function as opaque "black boxes" [62,66]. While interpretability tools, such as Shapley values, can offer partial insight into model behavior, their use remains limited and their explanations approximate.

In summary, the ethical and safety considerations surrounding AI in health care highlight the need for responsible development and careful implementation. Adhering to established frameworks, prioritizing transparency and equity, and rigorously evaluating AI tools in real-world settings are essential steps for fostering trust and ensuring that these technologies deliver meaningful benefits in clinical practice.

Study Limitations

This review has some limitations worth mentioning. First, although a comprehensive search strategy was used, it should be mentioned that an information scientist was not consulted during the development of the search string, and it is possible that some relevant studies were not identified. Additionally, the inclusion of only English-language publications may have resulted in the exclusion of pertinent research published in other languages.

XSL•FO
RenderX

Moreover, the recency of some included studies implies that external validation or implementation studies may still be ongoing. Their absence could limit the completeness of our evaluation, which should be revised as new evidence emerges. Nonetheless, the lack of external validation substantially impacts the reliability of the reported model performance.

In addition, the great heterogeneity of the included studies in terms of adherence measurement methods, performance metrics, variables considered, study aims, and sample sizes poses a challenge to the comparison and synthesis of their findings. Nonetheless, this variability is informative in itself, as it reflects the current state of research in this evolving field.

Another limitation to consider is the potential impact of the cancer stage on medication adherence; for example, individuals with early-stage cancer might demonstrate different motivational and psychological states compared to patients in the metastatic phase, which can influence their adherence behaviors. The heterogeneity of the population with BC included in this review (both early stage and metastatic) might limit the ability to generalize the findings across all cancer stages.

Finally, given the fast-evolving nature of the topic, this systematic review may become outdated rapidly, as AI technologies are continually advancing in power and capabilities, and new studies are constantly being conducted.

Clinical Implications

From the comprehensive examination of the included studies, some key clinical implications can be drawn. Notably, the issue of implementation recurs as a central challenge throughout the discussion of the findings.

Investing efforts in planning, conducting, and evaluating model implementation is of paramount importance, as only through their integration into clinical practice can these models support health care professionals in identifying at-risk patients early, thereby enabling tailored interventions at an earlier stage.

It has been suggested that, in order to bridge the implementation gap, it is necessary to move beyond the excessive reliance on performance metrics such as sensitivity, specificity, and AUC. While these metrics quantify how closely predictions align with actual outcomes, they do not necessarily reflect the model's practical utility in clinical settings. Indeed, a predictive model with high performance according to these metrics may not lead to tangible improvements in patient care [56,57]. Seneviratne et al [56] proposed 3 practical aspects that should be considered when developing a predictive model or assessing the applicability of an existing one: actionability, safety, and utility. Actionability refers to the extent to which a predictive model can inform or prompt a concrete action by the patient or the health care professional. In our context, a predictive model could be integrated into hospital clinical workflows so that, following routine assessments and screenings, if a patient is identified as being at high risk of nonadherence, an alert would automatically be triggered within their electronic medical record. This alert could then prompt appropriate interventions, such as educational initiatives, consultations with health care professionals, or referral to eHealth support programs. Safety is a key consideration in the clinical application of ML models, and it has been addressed in a previous section (see the Ethical and Safety Considerations section). Finally, utility should be carefully assessed by comparing costs and clinical outcomes with and without the implementation of the predictive model. In our context, does the use of the model lead to improved medication adherence among hospital patients? Does it reduce the costs associated with nonadherence? Is its implementation cost-effective and sustainable? These are highly relevant questions, and it is essential that clinicians and researchers consider them from the development phase and actively commit to evaluating their models against these criteria in later stages, to ensure that the potential clinical impact of the models can be fully realized.

Regarding AI-powered chatbots, further research is needed to ascertain their effectiveness in supporting medication adherence. The ethical and safety considerations discussed in the dedicated section are equally relevant to these tools, particularly in relation to patient privacy and the transparency of their underlying mechanisms [67,68]. Keeping in mind the crucial importance of meeting these prerequisites, advancements in AI technologies now make it feasible to develop AI tools that monitor and support patients throughout their cancer journey. These tools could assist patients in tracking medication-taking behaviors, deliver reminders, and guide patients in managing treatment-related side effects. Moreover, AI-powered chatbots, trained on verified clinical guidelines and scientific knowledge, can provide patients with instant support and information, reducing the burden on health care professionals [40,69]. However, for ensuring high quality of these AI models and their wide usage, it is crucial for scientific researchers, technology professionals, clinicians, and policymakers to work collaboratively to design, develop, and integrate these technologies into health care.

Conclusions

To our knowledge, this is the first systematic review to provide a comprehensive and up-to-date overview of the current applications of AI in predicting and supporting medication adherence among patients with BC. The review encompasses both predictive and interventional studies, thereby mapping the existing applications and highlighting critical gaps. Although these technologies show potential, particularly in identifying patients at risk of nonadherence, the evidence base remains limited, and the high risk of bias observed calls for caution when interpreting current findings.

At this stage, the clinical contribution of AI remains preliminary. Meaningful progress will require stronger attention to implementation, clearer plans for real-world integration, and approaches that actively involve patients throughout development and use. Ethical and practical considerations will also need to guide each step, from development to clinical deployment. With these conditions in place, AI-based tools may, over time, evolve into a useful complement to existing adherence management strategies, but further rigorous work is needed before they can be considered ready for routine practice.

Our findings provide a foundation for future investigation and underscore the importance of a coordinated, multidisciplinary

XSL•FO
RenderX

approach to enhance the accuracy, reliability, and safety of AI tools while addressing critical ethical challenges.

Acknowledgments

MP and VV are PhD students in Medical Humanities within the European School of Molecular Medicine (SEMM) at the University of Milan, Italy. The authors confirm that, apart from routine grammar checks during the revision process, no generative artificial intelligence tools were used in producing the text, figures, or any other informational content of this paper.

Funding

No funding was received for this work.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: MP, VV, IC, CM
Data curation: MP, VV
Formal analysis: MP, VV
Investigation: MP, VV, IC, CM
Project administration: GP
Resources: IC
Supervision: GP
Visualization: MP, VV
Writing – original draft: MP, VV
Writing – review & editing: IC, CM

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 230 KB - cancer_v12i1e80128_app1.docx]

Multimedia Appendix 2
Table for information on the prospective study.
[DOCX File, 15 KB - cancer_v12i1e80128_app2.docx]

Checklist 1
PRISMA 2020 checklist.
[DOCX File, 11145 KB - cancer_v12i1e80128_app3.docx]

References

1. Greer JA, Amoyal N, Nisotel L, et al. A systematic review of adherence to oral antineoplastic therapies. Oncologist 2016 Mar;21(3):354-376. [doi: 10.1634/theoncologist.2015-0405] [Medline: 26921292]
2. Mathes T, Antoine SL, Pieper D, Eikermann M. Adherence enhancing interventions for oral anticancer agents: a systematic review. Cancer Treat Rev 2014 Feb;40(1):102-108. [doi: 10.1016/j.ctrv.2013.07.004] [Medline: 23910455]
3. Rosenberg SM, Petrie KJ, Stanton AL, Ngo L, Finnerty E, Partridge AH. Interventions to enhance adherence to oral antineoplastic agents: a scoping review. J Natl Cancer Inst 2020 May 1;112(5):443-465. [doi: 10.1093/jnci/djz244] [Medline: 31899790]
4. Hamer J, Warner E. Lifestyle modifications for patients with breast cancer to improve prognosis and optimize overall health. CMAJ 2017 Feb 21;189(7):E268-E274. [doi: 10.1503/cmaj.160464] [Medline: 28246240]
5. Pezzolato M, Marzorati C, Lanzoni L, et al. Interventions to increase adherence to oral therapies in breast cancer patients: a systematic review based on the behavior change technique taxonomy. Psychooncology 2023 Oct;32(10):1481-1502. [doi: 10.1002/pon.6203] [Medline: 37571974]

XSL•FO
RenderX

6. Marzorati C, Voskanyan V, Sala D, et al. Psychosocial factors associated with quality of life in cancer patients undergoing treatment: an umbrella review. Health Qual Life Outcomes 2025 Apr 5;23(1):31. [doi: 10.1186/s12955-025-02357-z] [Medline: 40188134]
7. Cramer JA, Roy A, Burrell A, et al. Medication compliance and persistence: terminology and definitions. Value Health 2008;11(1):44-47. [doi: 10.1111/j.1524-4733.2007.00213.x] [Medline: 18237359]
8. Brown MT, Bussell JK. Medication adherence: WHO cares? Mayo Clin Proc 2011 Apr;86(4):304-314. [doi: 10.4065/mcp.2010.0575] [Medline: 21389250]
9. Adherence to long-term therapies: evidence for action [internet]. World Health Organization. 2003. URL: https://iris.who.int/handle/10665/42682 [accessed 2026-03-30]
10. Yussof I, Mohd Tahir NA, Hatah E, Mohamed Shah N. Factors influencing five-year adherence to adjuvant endocrine therapy in breast cancer patients: a systematic review. Breast 2022 Apr;62:22-35. [doi: 10.1016/j.breast.2022.01.012] [Medline: 35121501]
11. Eliassen FM, Blåfjelldal V, Helland T, et al. Importance of endocrine treatment adherence and persistence in breast cancer survivorship: a systematic review. BMC Cancer 2023 Jul 4;23(1):625. [doi: 10.1186/s12885-023-11122-8] [Medline: 37403065]
12. Makubate B, Donnan PT, Dewar JA, Thompson AM, McCowan C. Cohort study of adherence to adjuvant endocrine therapy, breast cancer recurrence and mortality. Br J Cancer 2013 Apr 16;108(7):1515-1524. [doi: 10.1038/bjc.2013.116] [Medline: 23519057]
13. Todd A, Waldron C, McGeagh L, et al. Identifying determinants of adherence to adjuvant endocrine therapy following breast cancer: a systematic review of reviews. Cancer Med 2024 Feb;13(3):e6937. [doi: 10.1002/cam4.6937] [Medline: 38240343]
14. Finitsis DJ, Vose BA, Mahalak JG, Salner AL. Interventions to promote adherence to endocrine therapy among breast cancer survivors: a meta-analysis. Psychooncology 2019 Feb;28(2):255-263. [doi: 10.1002/pon.4959] [Medline: 30511789]
15. Heiney SP, Parker PD, Felder TM, Adams SA, Omofuma OO, Hulett JM. A systematic review of interventions to improve adherence to endocrine therapy. Breast Cancer Res Treat 2019 Feb;173(3):499-510. [doi: 10.1007/s10549-018-5012-7] [Medline: 30387003]
16. Hurtado-de-Mendoza A, Cabling ML, Lobo T, Dash C, Sheppard VB. Behavioral interventions to enhance adherence to hormone therapy in breast cancer survivors: a systematic literature review. Clin Breast Cancer 2016 Aug;16(4):247-255. [doi: 10.1016/j.clbc.2016.03.006] [Medline: 27133733]
17. Nieuwlaat R, Wilczynski N, Navarro T, et al. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2014 Nov 20;2014(11):CD000011. [doi: 10.1002/14651858.CD000011.pub4] [Medline: 25412402]
18. Bekbolatova M, Mayer J, Ong CW, Toma M. Transformative potential of AI in healthcare: definitions, applications, and navigating the ethical landscape and public perspectives. Healthcare (Basel) 2024 Jan 5;12(2):2. [doi: 10.3390/healthcare12020125] [Medline: 38255014]
19. Bhattacharya S, Saleem SM, Singh A, Singh S, Tripathi S. Empowering precision medicine: regenerative AI in breast cancer. Front Oncol 2024;14:1465720. [doi: 10.3389/fonc.2024.1465720] [Medline: 39372870]
20. Arigo D, Jake-Schoffman DE, Pagoto SL. The recent history and near future of digital health in the field of behavioral medicine: an update on progress from 2019 to 2024. J Behav Med 2025 Feb;48(1):120-136. [doi: 10.1007/s10865-024-00526-x] [Medline: 39467924]
21. Stryker C, Kavlakoglu E. What is artificial intelligence (AI)? IBM. URL: https://www.ibm.com/think/topics/artificial-intelligence [accessed 2026-03-30]
22. Sheikh H, Prins C, Schrijvers E. Artificial intelligence: definition and background. In: Mission AI Research for Policy: Springer International Publishing; 2023:15-41 URL: https://link.springer.com/10.1007/978-3-031-21448-6_2 [accessed 2026-03-30]
23. Huang S, Yang J, Fong S, Zhao Q. Artificial intelligence in cancer diagnosis and prognosis: opportunities and challenges. Cancer Lett 2020 Feb 28;471:61-71. [doi: 10.1016/j.canlet.2019.12.007] [Medline: 31830558]
24. Zhang B, Shi H, Wang H. Machine learning and AI in cancer prognosis, prediction, and treatment selection: a critical approach. J Multidiscip Healthc 2023;16:1779-1791. [doi: 10.2147/JMDH.S410301] [Medline: 37398894]
25. Toma M, Wei OC. Predictive modeling in medicine. Encyclopedia 2023;3(2):590-601. [doi: 10.3390/encyclopedia3020042]
26. Jerfy A, Selden O, Balkrishnan R. The growing impact of natural language processing in healthcare and public health. Inquiry 2024;61:469580241290095. [doi: 10.1177/00469580241290095] [Medline: 39396164]
27. Li C, Zhang Y, Weng Y, Wang B, Li Z. Natural language processing applications for computer-aided diagnosis in oncology. Diagnostics (Basel) 2023 Jan 12;13(2):286. [doi: 10.3390/diagnostics13020286] [Medline: 36673096]
28. Tudor Car L, Dhinagaran DA, Kyaw BM, et al. Conversational agents in health care: scoping review and conceptual analysis. J Med Internet Res 2020 Aug 7;22(8):e17158. [doi: 10.2196/17158] [Medline: 32763886]
29. Al Kuwaiti A, Nazer K, Al-Reedy A, et al. A review of the role of artificial intelligence in healthcare. J Pers Med 2023 Jun 5;13(6):951. [doi: 10.3390/jpm13060951] [Medline: 37373940]
30. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021 Jul;8(2):e188-e194. [doi: 10.7861/fhj.2021-0095] [Medline: 34286183]

XSL•FO
RenderX

31. Durosini I, Pizzoli SFM, Strika M, Pravettoni G. Chapter 16 - artificial intelligence and medicine: a psychological perspective on AI implementation in healthcare context. In: Artificial Intelligence for Medicine [Internet]: Academic Press; 2024:231-237. [doi: 10.1016/B978-0-443-13671-9.00011-9]
32. Pizzoli SFM, Durosini I, Strika M, Pravettoni G. Chapter 5 - artificial intelligence and digital worlds: new frontiers of integration between AI and other technological tools. In: Artificial Intelligence for Medicine [Internet]: Academic Press; 2024:63-72. [doi: 10.1016/B978-0-443-13671-9.00008-9]
33. Kumar Y, Marchena J, Awlla AH, Li JJ, Abdalla HB. The AI-powered evolution of big data. Appl Sci (Basel) 2024;14(22):10176. [doi: 10.3390/app142210176]
34. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
35. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev 2021 Jan 26;10(1):39. [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
36. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
37. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. Ann Intern Med 2019 Jan 1;170(1):W1-W33. [doi: 10.7326/M18-1377] [Medline: 30596876]
38. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998 Jun;52(6):377-384. [doi: 10.1136/jech.52.6.377] [Medline: 9764259]
39. Balazard F, Bertaut A, Bordet É, et al. Adjuvant endocrine therapy uptake, toxicity, quality of life, and prediction of early discontinuation. J Natl Cancer Inst 2023 Sep 7;115(9):1099-1108. [doi: 10.1093/jnci/djad109] [Medline: 37434306]
40. Chaix B, Bibault JE, Pienkowski A, et al. When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot. JMIR Cancer 2019 May 2;5(1):e12856. [doi: 10.2196/12856] [Medline: 31045505]
41. Kaur N, Gonzales Iv M, Alcaraz CG, Gong J, Wells KJ, Barnes LE. A computational framework for longitudinal medication adherence prediction in breast cancer survivors: a social cognitive theory based approach. PLOS Digit Health 2025 Jun;4(6):e0000839. [doi: 10.1371/journal.pdig.0000839] [Medline: 40493566]
42. Kaur N, Gonzales M, Alcaraz CG, Barnes LE, Wells KJ, Gong J. Theory-guided randomized neural networks for decoding medication-taking behavior. 2021 IEEE EMBS Intl Conf Biomed Health Inf (BHI) 2021:1-4. [doi: 10.1109/BHI50953.2021.9508614]
43. Kuo CC, Wang HH, Tseng LP. Using data mining technology to predict medication-taking behaviour in women with breast cancer: a retrospective study. Nurs Open 2022 Nov;9(6):2646-2656. [doi: 10.1002/nop2.963] [Medline: 34156764]
44. Masiero M, Spada GE, Sanchini V, et al. A machine learning model to predict patients' adherence behavior and a decision support system for patients with metastatic breast cancer: protocol for a randomized controlled trial. JMIR Res Protoc 2023 Dec 14;12(1):e48852. [doi: 10.2196/48852] [Medline: 38096002]
45. Rattsev I, Stearns V, Blackford AL, et al. Incorporation of emergent symptoms and genetic covariates improves prediction of aromatase inhibitor therapy discontinuation. JAMIA Open 2024 Apr;7(1):ooae006. [doi: 10.1093/jamiaopen/ooae006] [Medline: 38250582]
46. Rinder P, Marcille T, Sinel-Boucher P, et al. Dynamic projection of medication nonpersistence and nonadherence among patients with early breast cancer. JAMA Netw Open 2024 May 1;7(5):e2411909. [doi: 10.1001/jamanetworkopen.2024.11909] [Medline: 38758553]
47. Yerrapragada G, Siadimas A, Babaeian A, Sharma V, O'Neill TJ. Machine learning to predict tamoxifen nonadherence among US commercially insured patients with metastatic breast cancer. JCO Clin Cancer Inform 2021 Aug;5(5):814-825. [doi: 10.1200/CCI.20.00102] [Medline: 34383580]
48. Yuan F, Kaur N, Wang Z, et al. Multimodal sensing and modeling of endocrine therapy adherence in breast cancer survivors. Proc ACM Interact Mob Wearable Ubiquitous Technol 2025 Dec;9(4):1-33. [doi: 10.1145/3770864] [Medline: 41415333]
49. Carter JV, Pan J, Rai SN, Galandiuk S. ROC-ing along: evaluation and interpretation of receiver operating characteristic curves. Surgery 2016 Jun;159(6):1638-1645. [doi: 10.1016/j.surg.2015.12.029] [Medline: 26962006]
50. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. JAMA 2017 Oct 10;318(14):1377-1384. [doi: 10.1001/jama.2017.12126] [Medline: 29049590]
51. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020 Mar 18;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]
52. Tsegaye B, Snell KIE, Archer L, et al. Larger sample sizes are needed when developing a clinical prediction model using machine learning in oncology: methodological systematic review. J Clin Epidemiol 2025 Apr;180:111675. [doi: 10.1016/j.jclinepi.2025.111675] [Medline: 39814217]
53. Steiner JF. Can we identify clinical predictors of medication adherence... and should we? Med Care 2010 Mar;48(3):193-195. [doi: 10.1097/MLR.0b013e3181d51ddf] [Medline: 20182266]

XSL•FO
RenderX

54. Given BA, Spoelstra SL, Grant M. The challenges of oral agents as antineoplastic treatments. Semin Oncol Nurs 2011 May;27(2):93-103. [doi: 10.1016/j.soncn.2011.02.003] [Medline: 21514479]
55. Nizet P, Touchefeu Y, Pecout S, et al. Exploring the factors influencing adherence to oral anticancer drugs in patients with digestive cancer: a qualitative study. Support Care Cancer 2022 Mar;30(3):2591-2604. [doi: 10.1007/s00520-021-06663-2] [Medline: 34812952]
56. Seneviratne MG, Shah NH, Chu L. Bridging the implementation gap of machine learning in healthcare. BMJ Innov 2020 Apr;6(2):45-47. [doi: 10.1136/bmjinnov-2019-000359]
57. Shah NH, Milstein A, Bagley PhD SC. Making machine learning models clinically useful. JAMA 2019 Oct 8;322(14):1351-1352. [doi: 10.1001/jama.2019.10306] [Medline: 31393527]
58. Wessler BS, Lai Yh L, Kramer W, et al. Clinical prediction models for cardiovascular disease: tufts predictive analytics and comparative effectiveness clinical prediction model database. Circ Cardiovasc Qual Outcomes 2015 Jul;8(4):368-375. [doi: 10.1161/CIRCOUTCOMES.115.001693] [Medline: 26152680]
59. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ 2024 Jan 8;384:e074819. [doi: 10.1136/bmj-2023-074819] [Medline: 38191193]
60. Chen JH, Asch SM. Machine learning and prediction in medicine - beyond the peak of inflated expectations. N Engl J Med 2017 Jun 29;376(26):2507-2509. [doi: 10.1056/NEJMp1702071] [Medline: 28657867]
61. Bibault JE, Chaix B, Guillemassé A, et al. A chatbot versus physicians to provide information for patients with breast cancer: blind, randomized controlled noninferiority trial. J Med Internet Res 2019 Nov 27;21(11):e15787. [doi: 10.2196/15787] [Medline: 31774408]
62. Williamson SM, Prybutok V. Balancing privacy and progress: a review of privacy challenges, systemic oversight, and patient perceptions in AI-driven healthcare. Appl Sci (Basel) 2024;14(2):675. [doi: 10.3390/app14020675]
63. Gawankar S, Nair S, Pawar V, Vhatkar A, Chavan P. Patient privacy and data security in the era of AI-driven healthcare. 2024 8th Int Conf Comput, Commun, Control Autom (ICCUBEA) 2024:1-6. [doi: 10.1109/ICCUBEA61740.2024.10775004]
64. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. J Glob Health 2019 Dec;9(2):010318. [doi: 10.7189/jogh.09.020318] [Medline: 31788229]
65. Chin MH, Afsar-Manesh N, Bierman AS, et al. Guiding principles to address the impact of algorithm bias on racial and ethnic disparities in health and health care. JAMA Netw Open 2023 Dec 1;6(12):e2345050. [doi: 10.1001/jamanetworkopen.2023.45050] [Medline: 38100101]
66. Khan N, Nauman M, Almadhor AS, Akhtar N, Alghuried A, Alhudhaif A. Guaranteeing correctness in black-box machine learning: a fusion of explainable ai and formal methods for healthcare decision-making. IEEE Access 2024;12:90299-90316. [doi: 10.1109/ACCESS.2024.3420415]
67. Basharat I, Shahid S. AI-enabled chatbots healthcare systems: an ethical perspective on trust and reliability. JHOM (forthcoming) 2024. [doi: 10.1108/JHOM-10-2023-0302]
68. May R, Denecke K. Security, privacy, and healthcare-related conversational agents: a scoping review. Inform Health Soc Care 2022 Apr 3;47(2):194-210. [doi: 10.1080/17538157.2021.1983578] [Medline: 34617857]
69. Xu L, Sanders L, Li K, Chow JCL. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR Cancer 2021 Nov 29;7(4):e27850. [doi: 10.2196/27850] [Medline: 34847056]

Abbreviations

AI: artificial intelligence
AUC: area under the receiver operating characteristic curve
BC: breast cancer
C-index: concordance index
DL: deep learning
GenAI: generative artificial intelligence
LightGBM: light gradient boosting machine
ML: machine learning
NLP: natural language processing
PRISMA : Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses literature search extension
PROBAST : Prediction Model Risk of Bias Assessment Tool
PROSPERO : International Prospective Register of Systematic Reviews
SHAP: Shapley Additive Explanations
XGBoost: Extreme Gradient Boosting

XSL•FO
RenderX

Edited by M Balcarras; submitted 04.Jul.2025; peer-reviewed by FG Rebitschek, Z Ping; revised version received 24.Feb.2026; accepted 24.Feb.2026; published 21.Apr.2026.

Please cite as:
Pezzolato M, Voskanyan V, Cutica I, Marzorati C, Pravettoni G
Use of AI to Predict and Support Medication Adherence in Patients With Breast Cancer: Systematic Review
JMIR Cancer 2026;12:e80128
URL: https://cancer.jmir.org/2026/1/e80128
doi:10.2196/80128

© Massimo Pezzolato, Viktorya Voskanyan, Ilaria Cutica, Chiara Marzorati, Gabriella Pravettoni. Originally published in JMIR Cancer (https://cancer.jmir.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Improving Long-Term Adherence to Endocrine Therapy Among Breast Cancer Survivors: Development of a Multiscale Modeling and Intervention System

Manuel Gonzales IV[1], MS; Cristian Garcia-Alcaraz[1], MA, MS; Navreet Kaur[2], PhD; Anna N Baglione[2], PhD; Sarah Livermon[2], BS; Laura E Barnes[2], PhD; Kristen J Wells[3], MPH, PhD

[1]SDSU/UCSD Joint Doctoral Program in Clinical Psychology, 5500 Campanile Drive, San Diego, CA, United States

[2]Department of Engineering Systems and Environment, University of Virginia, Charlottesville, VA, United States

[3]Department of Psychology, San Diego State University, San Diego, CA, United States

Corresponding Author:
Manuel Gonzales IV, MS
SDSU/UCSD Joint Doctoral Program in Clinical Psychology, 5500 Campanile Drive, San Diego, CA, United States

Abstract

Background: Breast cancer is a significant public health burden. Despite its critical role in preventing the recurrence of breast cancer, rates of long-term adherence to endocrine therapy (ET) remain low among certain breast cancer survivors. Using embedded sensors in smartphones and wearables, ecological momentary assessment data and health behavior theory may facilitate a richer understanding of the real-world context of medication-taking behaviors, which can aid in the development of personalized interventions.

Objective: The objective of this paper is to describe the development of a multiscale modeling intervention (MMI) system to facilitate adherence to daily oral ET for breast cancer survivors. This represents the first phase of a larger project that aims to use machine learning to predict when breast cancer survivors are most likely to miss their ET medications in order to deploy personalized interventions. The purpose of this paper was (1) to determine the acceptability of the proposed MMI system, (2) to identify modifiable predictors of ET medication adherence among breast cancer survivors, and (3) to select surveys or items measuring constructs associated with ET adherence among breast cancer survivors for inclusion in the MMI system.

Methods: Study 1 consisted of usability interviews with a cohort of breast cancer survivors (n=25) prescribed ET. For study 1, all qualitative usability interviews were conducted using a semistructured interview guide and assessed whether breast cancer survivors were willing to use various components of the MMI system. Study 2 consisted of (1) a secondary data analysis of ET adherence data from 32 breast cancer survivors using a social cognitive theory framework and (2) a review of research literature of constructs and surveys measuring constructs associated with ET adherence among breast cancer survivors using a social cognitive theory framework. The secondary data analysis included the use of randomized neural network analysis to predict factors strongly associated with medication adherence.

Results: In study 1, usability interview findings suggested that participants were willing to use an ecological momentary assessment smartphone app, a smartwatch and associated smartphone app, a smart pill bottle or smart pill box and associated smartphone app, and the entire MMI system for a 6-month study period. In study 2, the randomized neural network analysis identified 104 survey items with significant contributions to 4-week medication adherence using a threshold of the 70th percentile for feature importance. After a review of peer-reviewed studies, we abstracted modifiable constructs significantly associated with adherence to adjuvant ET and identified 42 surveys used to measure these constructs. When these findings were combined, the final survey for the MMI system consisted of 32 surveys and demographic items.

Conclusions: Our research highlights the use of social cognitive theory, data-driven models, and participant feedback to inform the development of a medication adherence monitoring system. Data from studies 1 and 2 were used to develop a prototype MMI system that will be deployed in a future longitudinal study with 20 breast cancer survivors over 6 months.

Trial Registration: ClinicalTrials.gov NCT06865755; https://clinicaltrials.gov/study/NCT06865755

(JMIR Cancer 2026;12:e68255) doi:10.2196/68255

KEYWORDS

breast cancer survivors; social cognitive theory; mobile app; mHealth; endocrine therapy; medication adherence; breast neoplasms; maintenance chemotherapy; ecological momentary assessment; digital health

XSL•FO
RenderX

Introduction

Breast cancer is one of the most commonly diagnosed cancers worldwide [1]. In 2025, an estimated 319,750 people will be diagnosed with breast cancer, and 42,680 people will die from it in the United States [2]. Additionally, there are approximately 4.1 million breast cancer survivors in the United States [3]. Based on the most recent statistics, in the United States, the 5-year relative survival rate for breast cancer from the years 2014 to 2020 was 91% [2]. Individuals diagnosed with early stage and nonmetastatic breast cancer typically undergo a treatment regimen that can include surgery (ie, lumpectomy and mastectomy), radiation, chemotherapy (including trastuzumab), endocrine therapy (ET), immunotherapy, or a combination of treatments. The selection of treatment approaches is based on the stage of the cancer, whether breast cancer cells have certain receptors, and a patient's overall health and age [4-6]. Treatment for an early stage of breast cancer aims to eliminate the cancer, prevent recurrence, and maximize quality of life.

Giaquinto et al [3] reported that 80% of breast cancers diagnosed from the years 2015 to 2019 were estrogen receptor–positive. Survivorship guidelines developed by the American Cancer Society and American Society of Clinical Oncology recommend that breast cancer survivors with estrogen receptor–positive breast cancer take daily oral ET, including tamoxifen or aromatase inhibitors, for 5 to 10 years following treatment with surgery, chemotherapy, and radiation to reduce the risk of recurrence of breast cancer and subsequent secondary primary breast cancers [7]. ET consists of 1 pill taken at any time of the day, 1 time per day, with or without food. Individuals who take exemestane are required to avoid grapefruit and Seville, that is "bitter" oranges, as they reduce its efficacy [8]. Additionally, individuals taking ET should avoid or limit alcohol, as it reduces the efficacy of the medications [8].

Breast cancer survivors who take tamoxifen for 5 years posttreatment have a one-third reduced risk of breast cancer mortality through 15 years after diagnosis and an average of 39% reduced recurrence rate through 10 years after diagnosis compared to those who are not using tamoxifen [9]. Van Liew et al [10] conducted a systematic review of adherence to oral ET and found that breast cancer survivors who adhered to ET for more than 80% of the time had the lowest risk of the recurrence of breast cancer. Rates of long-term adherence to ET are low among certain breast cancer survivors despite its critical role in preventing the recurrence of breast cancer. Estimated rates of adherence to ET range from an average of 79% during the first year after treatment to an average of 56% during the fourth or fifth year [11]. A recent systematic review examining modifiable factors associated with adherence to ET identified the presence of side effects as being associated with ET adherence (in 53% of univariate studies and 100% of multivariate studies) [12]. Furthermore, positive decisional balance, positive emotions, negative emotions, depressive symptoms, perceived risk of breast cancer recurrence, intention to take ET, self-efficacy, participation in the decision to take ET, and quality of the relationship with a health care provider were shown to be associated with ET adherence across multiple studies. A recent systematic review and meta-analysis found that, overall, interventions that aimed to improve adherence to ET were effective, but there was considerable variation across the studies, with 53.3% of randomized controlled trials and 57.9% of nonrandomized controlled trials reporting an intervention effect to improve adherence. Interventions that reduced the cost of medications through legislation and the use of generic medications were consistently effective. Psychosocial interventions that focused on individual participants had limited efficacy when they were provided education about breast cancer, ET, and side effects. Those psychosocial interventions that targeted forgetting and involved medical personnel in managing side effects were sometimes effective [13]. Some psychosocial interventions to improve ET adherence may have limited effectiveness because there is a poor understanding of barriers to adherence and circumstances in which medication taking occurs or does not occur.

Further examination of adherence to ET may be crucial in reducing the burden of breast cancer and improving the overall quality of life and long-term survival among breast cancer survivors. The collection of ecologically valid data, such as sensors (ie, wearable sensors, medication event monitoring system [MEMS] pill cap sensors, and smartphone sensors) and ecological momentary assessment (EMA) data, may facilitate a better understanding of medication-taking behaviors. Baglione et al [14] described the framework of a multiscale modeling intervention (MMI) that presents a new approach to intervention development by capturing real-time data using mobile and wearable sensing systems and working toward incorporating theory-driven approaches to intervention development for adherence to ET. Through the lens of social cognitive theory (SCT) [15], identifying personal (ie, physiological, cognitive, and affective states), environmental (ie, social and physical environment), and behavioral (ie, medication-taking in the context of other behaviors) influences that impact adherence to ET may help provide better insights into the context of adherence and nonadherence. The combination of a theory-driven approach along with data collection using smartphone and wearable sensors, surveys, and EMA will allow for a much richer understanding of the real-world context of medication-taking behaviors, which will ultimately aid in the development of personalized interventions.

As part of the current project, in 2020, the study team conducted an informal assessment of commercially available wearable and MEMS sensors that could be used to collect ecologically valid data from breast cancer survivors to better understand their risk of ET nonadherence and aid in the development of an MMI system. Wearable sensors and MEMS devices were identified via online searches and through the knowledge of study team members. Sensors were reviewed and tested based on a wide variety of factors that would be important for feasible data collection and intervention as part of the MMI system (eg, cost, battery life, platform, etc; see Tables S1 and S2 in Multimedia Appendix 1). All data obtained from the review and testing of the MEMS devices were entered into an Excel spreadsheet. Multiple members of the project team reviewed these data to make final sensor selections. Based on this review, the team

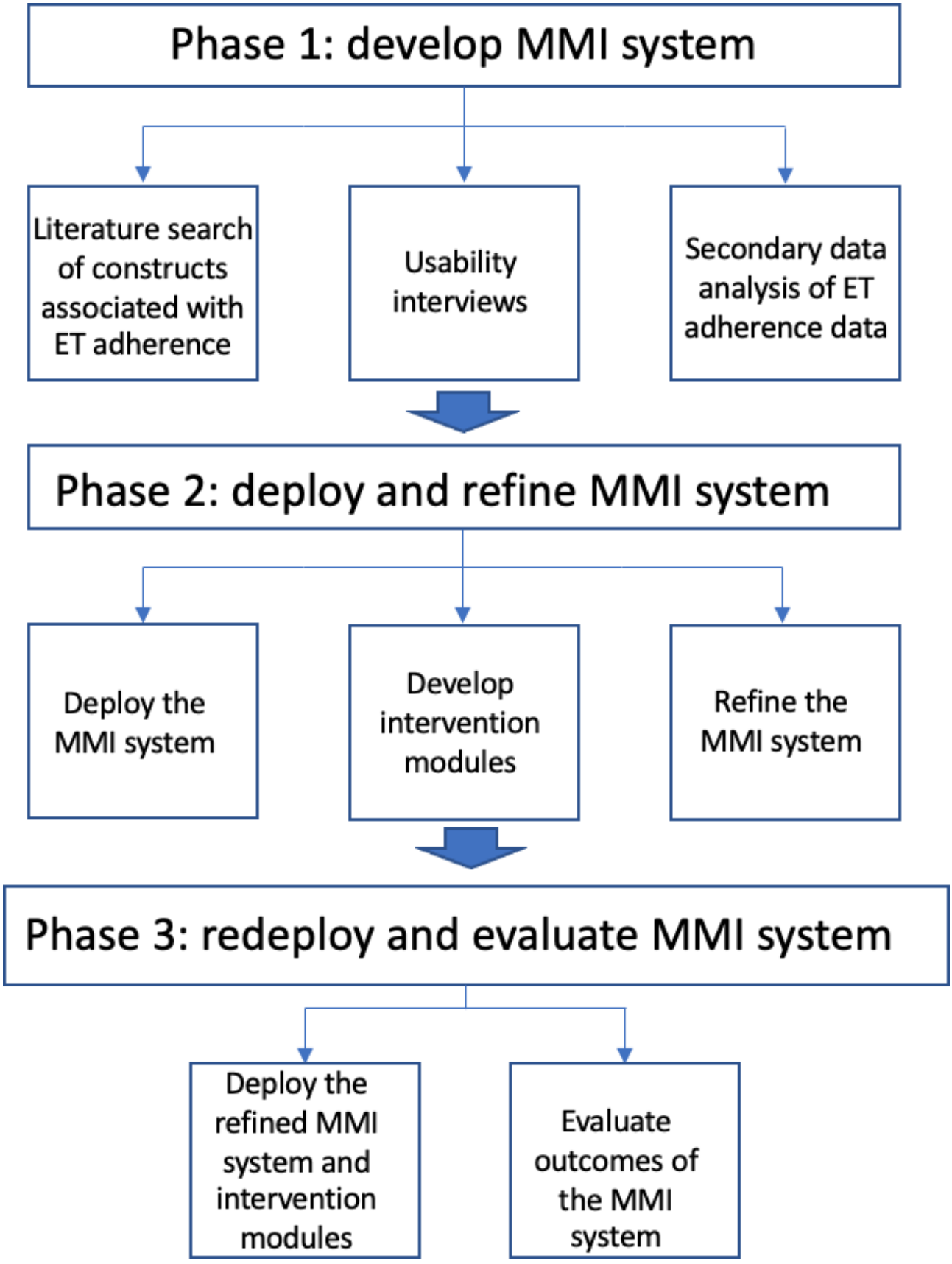
XSL•FO
RenderX

selected Fitbit Sense as the final wrist-worn sensor and RxCap as the final MEMS device for the MMI system.

In this paper, we describe the development of an MMI system to facilitate adherence to daily oral ET for breast cancer survivors who were diagnosed with stages 0 to 3 breast cancer in the past 5 years, have completed all surgeries, radiation, and chemotherapy treatments for breast cancer, and were prescribed ET. This represents the first phase of a larger project (Figure 1) that aims to use machine learning to predict when breast cancer survivors are most likely to miss their ET medications to deploy personalized interventions. The purpose of this initial phase was (1) to determine the acceptability of the proposed MMI system, (2) to identify modifiable predictors of ET medication adherence among breast cancer survivors, and (3) to select surveys/items measuring constructs associated with ET adherence among breast cancer survivors for inclusion in the MMI system. To develop the MMI system, we conducted 2 studies. Study 1 consisted of usability interviews with 25 breast cancer survivors prescribed ET. Study 2 consisted of (1) a secondary data analysis of ET adherence data from 32 breast cancer survivors and (2) a review of research literature of constructs and measurement of constructs associated with ET adherence among breast cancer survivors. The information obtained from these 2 studies allowed us to (1) develop long-form and EMA surveys and determine the frequency of administration of EMAs and (2) determine survey items and sensor data critical to include in the deployment of the prototype MMI system.

XSL•FO
RenderX

Figure 1. Flow diagram of multiscale modeling intervention (MMI) for medication adherence project. ET: endocrine therapy.



XSL•FO
RenderX

Methods

Overview

This phase of the project describes the development of the MMI system for a future 6-month deployment of the system among 20 breast cancer survivors.

Study 1: Usability Interviews With Breast Cancer Survivors

Participants

Using purposive sampling [16], participants were breast cancer survivors (n=25 women) prescribed ET who were recruited via social media posts (eg, Facebook) targeted toward patients diagnosed with breast cancer. These posts were made in breast cancer–related groups and channels. Those who were interested were asked to respond to the post with their contact information. A team member subsequently communicated with individuals who expressed interest in participating via email. Purposive sampling, unlike random sampling, is used to select cases who will provide appropriate and useful information [16]. Participants who expressed interest in participating in the study were sent an eligibility screener on Qualtrics. Eligibility criteria included participants who (1) had the ability to speak and read English, (2) were between the ages of 21 and 70 years, (3) were diagnosed with stages 0 to 3 breast cancer in the past 5 years, (4) had completed all surgeries, radiation, and chemotherapy for breast cancer, and (5) were prescribed ET. If a participant was deemed eligible based on the Qualtrics screener, they were subsequently contacted and scheduled for a screening via Zoom videoconference in order to confirm eligibility. If eligibility was confirmed, participants were scheduled for a 60- to 90-minute semistructured interview via Zoom videoconference. Interviews took place between September 2020 and July 2022.

Procedure

Using semistructured interview guides, 2 doctoral-level researchers (MG and CG-A) conducted 5 rounds of interviews assessing breast cancer survivors' perceptions of and willingness to use an EMA smartphone app, smartwatch, smart pill bottle, smart pill box, and the entire MMI system. All interview guides were developed by the study team and pilot-tested with undergraduate research assistants and individuals outside of the study team (see Multimedia Appendix 2 for all interview guides). One interviewer (MG), a male doctoral student, held a bachelor's degree in psychology, significant interest in systematic approaches to intervention development, and prior experience in clinical health psychology research. The second interviewer (CG-A), a male doctoral student, held a master's degree in psychology, significant interest in research methodology, and prior experience in qualitative research. Each round consisted of 5 participants who were not known to the researcher conducting interviews, and only 1 interviewer was present during each semistructured interview. Five interviews were conducted in each round based on the experience with prior human-centered technology design research [17]. Participants were interviewed alone. There were no repeat interviews. The purpose of round 1 interviews (n=5 participants) was to present a short survey on an EMA smartphone app and ask about participants' willingness to use and their thoughts on the usability of the EMA application. The purpose of round 2 (n=5 participants) was to present a smartwatch and EMA smartwatch application and ask about participants' willingness to use it and their thoughts on the usability of the smartwatch and associated smartphone app. The purpose of rounds 3 (n=5 participants) and 4 (n=5 participants) was to ask about participants' current medication-taking behaviors, demonstration of the use of a smart pill bottle and box, willingness to use a smart pill bottle or smart pill box, and willingness to use a smartphone app associated with the smart pill bottle. The purpose of round 5 (n=5 participants) was to ask about participants' willingness to use the entire MMI system, including completion of the long-form and EMA surveys, demonstration of the use of a pill bottle sensor, using a smartwatch, using apps associated with the 2 sensors, and their comfort with the EMA application collecting various streams of data (eg, HealthKit, activity, and Bluetooth encounters). All interviews were audio-recorded using a digital audio recorder. In addition, participants completed a short survey using Qualtrics, which assessed the following sociodemographic and clinical variables: age, Latinx ethnicity, race, education, marital status, sex assigned at birth, gender identity, employment status, occupation, country of birth, years residing in the United States, primary language, number of people in the household, annual household income, year of breast cancer diagnosis, stage of breast cancer diagnosis, types of treatment for breast cancer, and whether they experienced a breast cancer recurrence.

Data Analysis

All recorded audio files were transcribed verbatim by 2 undergraduate research assistants. Transcripts were not returned to participants for comment. A member of the research team (MG) reviewed all transcripts and developed a summary document for each of the 5 rounds of usability interviews (n=5 for each round). Data were summarized based on the participants' answers to specific questions in the interview, such as *"In general, how do you feel about the experience of answering the surveys on the smartwatch?"* For some items, responses were summed (eg, how many participants stated it was easy to use a feature). Descriptive statistics (ie, frequencies, means, and SDs) of participant characteristics were calculated via SPSS (version 29; IBM).

Ethical Considerations

This cross-sectional qualitative study was approved by the University of Virginia Institutional Review Board (IRB) (IRB-HSR number 21406) to which the San Diego State University and University of Alabama IRBs agreed to rely on the regulation of the study. This paper follows COREQ (Consolidated Criteria for Reporting Qualitative Research) reporting guidelines for qualitative research (a complete COREQ checklist can be found in Checklist 1) [18]. All participants provided written consent prior to starting the usability interviews, which took place between September 2020 and July 2022. Data were stored in accordance with the University of Virginia's information security policies, and deidentified data were analyzed. Participants were provided with a US $75 gift card for completion of the 60- to 90-minute semistructured

XSL•FO
RenderX

interview and self-report survey. This study adhered to local, national, regional, and international law and regulations regarding protection of personal information, privacy, and human rights.

Study 2

Secondary Data Analysis

As previously reported by study team members Kaur et al [19], to select long-form surveys and EMA items/surveys, the research team analyzed data from 32 early-stage breast cancer survivors [19,20]. In the primary study from which these data were obtained, eligible participants provided informed consent after participating in the screening process. After completing baseline surveys during the primary study, breast cancer survivors' daily medication adherence was assessed using a MEMS pill bottle (Aardex) over the 8-month time period. As part of the primary study, participants were compensated with gift cards worth up to US $90: US $30 for the baseline survey, US $30 for the 4-month survey, and US $30 for the 8-month survey. The data from the primary study were stored in accordance with San Diego State University and Sharp HealthCare Information Security policies, and deidentified data were analyzed by Kaur et al [19]. The initial IRB approval by the San Diego State University IRB (vIRB number 1239088) and the Sharp HealthCare IRB (IRB number 1308810) permitted the secondary use of data.

Kaur et al [19] conducted a randomized neural network (RNN) analysis to determine important features (eg, attitudes toward medications, stress, and mood) of ET adherence 4 weeks following the completion of the 346-item baseline survey. Kaur et al [19] extracted the weights on different input variables of the trained RNN models, which are equivalent to the contribution of the respective variables in the final outcome prediction. Additionally, Kaur et al [19] were guided by principles of SCT [15], which evaluates how environmental factors, personal factors, and a person's behavior interact through a process of reciprocal determinism. In other words, the SCT indicates that each of these aspects (eg, environmental factors, personal factors, and person's behavior) influences each other dynamically, which is important for the MMI system. Therefore, Kaur et al [19] used the SCT [15] when identifying personal, behavioral, and environmental constructs associated with daily medication adherence in the RNN analysis. Two authors (KJW and LEB) were largely responsible for categorizing constructs as personal, behavioral, or environmental factors associated with daily medication adherence in the RNN analysis. These findings were presented to other members of the research team for their feedback, which were incorporated by a third author (NK). Details regarding the RNN analyses are reported in the prior publication [19]. Kaur et al [19] chose the survey item importance starting with the fourth week of ET MEMS data collection, as the Hawthorne effect associated with the new use of the MEMS cap and enrollment in the study would have attenuated by week 4. In contrast, during weeks 2 and 3, participants were probably still affected by the study enrollment, completing the baseline surveys, and acclimating to the MEMS cap. Thus, only the 4-month and 8-month data were used to prevent noise from baseline data from corrupting the models. Kaur et al [19] used descriptive statistics to summarize demographic and clinical data. Details of the final survey items are provided in the *Results* section.

Review of Research Literature

Because the secondary analysis of data did not include all potential modifiable factors associated with ET medication adherence, we also examined the systematic review of Toivonen et al [12] (2020) to further develop the long-form survey and EMA components for the MMI system. The synthesis of Toivonen et al [12] comprised 68 studies that examined modifiable factors associated with adherence to adjuvant ET in breast cancer survivors. Possible modifiable factors of adjuvant ET in the literature were classified into the following 6 categories by Toivonen et al (2020) [12]: side effects (n=44 studies), attitudes toward adjuvant ET (n=29 studies), psychological factors (n=30 studies), health care provider–related factors (n=26 studies), sociocultural factors (n=25 studies), and general/quality-of-life factors (n=24 studies). Across individual studies within the synthesis of Toivonen et al [12], 1 study team member and author (CG-A) abstracted which modifiable factors were significantly associated with adherence to adjuvant ET as well as the different surveys/items used to measure these factors. One author (CG-A) also retrieved articles describing the development and psychometric evaluation of such surveys or items. This author (CG-A) then collaborated with another author (KJW), who has expertise in survey development, to determine the quality of these surveys or items. Surveys or items were considered for initial inclusion in the long-form survey and EMA components for the MMI system if they had adequate psychometric properties based on a minimum of a Cronbach $\alpha \geq 0.70$ in at least 1 study sample (when psychometric information was available).

The final inclusion of items and surveys in the long-form and EMA surveys was determined by the monetary cost, length of the surveys and items, and the previously described secondary data analysis findings. To be included in the MMI system, the surveys/items had to be free or a low cost, short enough to ensure the completion of the full battery of long-form surveys within 45 minutes or completion of EMA surveys within 5 minutes, and measure significant modifiable factors of adherence to adjuvant ET according to the secondary data analysis and the systematic literature review of Toivonen et al [12]. All authors participated in the decision-making process to select items and surveys during weekly team meetings in which the RNN analysis and the literature review were discussed. For the long-form surveys, a draft of the baseline survey was created by one author (CG-A), revised by a second author (KJW), and then presented to all group members for their suggestions and revisions. Once the baseline survey was finalized through consensus of the study team, the 2 follow-up surveys were created by one author (CG-A) and edited by a second author (KJW). The follow-up surveys were then presented to the study team for their suggestions and revisions. The EMA surveys were drafted by 2 study authors (ANB and KJW) and reviewed by all other study team members over several rounds of revisions. The long-form and EMA surveys were finalized once the study team agreed that no further edits were required.

XSL•FO
RenderX

Results

Study 1

Sample Characteristics

Out of 25 participants, 84% were White (n=21), 4% were African American/Black (n=1), and 12% were identified as Hispanic/Latinx (n=3; Table 1). The mean age of the sample was 53.6 (SD 10.1) years. Twenty-two out of 25 participants were married (88%), 13 were employed for wages (52%), 22 of them were born in the United States (88%), 24 had a primary language of English (96%), 10 were diagnosed with stage 1 breast cancer (43%), and 24 of them did not have a breast cancer recurrence (96%). Twenty-nine respondents agreed to participate in the interviews and 25 were interviewed (86% completion rate). Three participants who were scheduled to participate in an interview did not attend the appointment (75%), and 1 indicated she was no longer interested in participating prior to the interview (25%).

XSL•FO
RenderX

Table . Sample demographics and clinical characteristics of study 1 participants.

Characteristic and variable	Value
Race/ethnicity (n=25), n (%)	
White	21 (84)
African American/Black	1 (4)
Hispanic/Latinx	3 (12)
Sex (n=24), n (%)	
Female	24 (100)
Gender (n=25), n (%)	
Female	25 (100)
Marital status (n=25), n (%)	
Legally married	22 (88)
In a committed relationship	1 (4)
Divorced/separated	1 (4)
Widowed	1 (4)
Employment status (n=25), n (%)	
Employed for wages	13 (52)
Self-employed	2 (8)
Out of work	1 (4)
Homemaker	3 (12)
Student	1 (4)
Retired	5 (20)
Country of birth (n=25), n (%)	
United States	22 (88)
Canada	1 (4)
Denmark	1 (4)
Ukraine	1 (4)
Annual household income in US dollars (n=25), n (%)	
10,000-19,999	1 (4)
20,000-29,999	0
30,000-39,999	0
40,000-49,999	0
50,000-74,999	0
75,000-99,999	8 (32)
100,000-124,999	5 (20)
125,000-149,999	2 (8)
150,000 or more	9 (36)
Primary language (n=25), n (%)	
English	24 (96)
Danish	1 (4)
Breast cancer stage (n=23), n (%)	
Stage 0	2 (9)
Stage 1	10 (43)
Stage 2	5 (22)

XSL•FO
RenderX

Characteristic and variable	Value
Stage 3	6 (26)
Type of treatment (n=25), n (%)	
Single mastectomy	3 (12)
Double mastectomy	11 (44)
Lumpectomy	13 (52)
Another type of surgery	1 (0.04)
Radiation	17 (68)
Chemotherapy	16 (64)
Other	4 (16)
Breast cancer recurrence (n=25), n (%)	
Yes	1 (4)
No	24 (96)
Age in years (n=25), mean (SD)	53.6 (10.1)

Round 1 Interviews

All participants stated that filling out the survey on the smartphone app was easy and straightforward, and they would find it easy to use the smartphone app on their own phones and are willing to use the app for the future 6-month study duration. Additionally, several participants reported preferences for completing a survey up to 2 times per day with up to 10 questions at a time. Participants also had preferences for completing surveys early in the day or later in the evening. Furthermore, a few participants reported concerns about "being a slave to their phone," receiving multiple notifications throughout the day while busy, and being busier on the weekends. However, participants reported that they would prefer an option to receive a notification to complete a survey and be able to attend to it at a later time. Finally, participants reported recommendations for lighter colors in the mobile app, the ability to customize the colors of the app, and the inclusion of larger text.

When asked about how easy or difficult it was to answer the questions using the EMA smartphone app, one participant said:

> *That was easy...That was easy. I think that the answers were just, it was laid out very simple. There wasn't a lot of distraction; so the answer was very easy to just select, to read the question and select the appropriate answer.* [P3]

Round 2 Interviews

Participants described the use of the smartwatch as easy, user-friendly, and clear and liked it as it was convenient, easy, and fast. Most participants reported that they would find it easy to remember to wear the watch every day for the duration of the 6-month study period, would accept answering questions on the smartwatch application for the 6-month study period, and that the use of the app on the smartphone would be easy. Several participants also reported that they would not prefer to wear the watch while showering, sleeping, or dressing up in a nice outfit. Several participants preferred larger text on the app and stated that the smartwatch demonstrated was too big. Only 1 participant reported that they might have difficulty using a smartwatch for the 6-month study duration due to not having a preference for wearing jewelry or watches.

When asked about their experience of answering surveys on a smartwatch emulator, one participant said:

> *I think it's very easy, very user friendly. So, I think it's great...I think, so if that just pops up on your watch without you having to do anything, it's very user friendly; it has the options there. Anybody could do it—like a 5-year-old could do it.* [P23]

Rounds 3 and 4 Interviews

The mean number of reported medicines taken was 3.1 (SD 1.4), and the mean number of vitamins or supplements taken was about 3.6 (SD 2.1). Almost all participants reported taking their medicines or supplements in the morning and/or in the evening. A majority of participants stored their medicines and supplements in the kitchen, on their nightstand, or a combination of both, and 5 participants reported keeping their medicines in a pill box, while 4 reported keeping them in their original container. Of note, only 2 participants reported having someone who helps them manage their medications. Generally, participants were willing to use the smart pill bottle or smart pill box and the associated smartphone app for the duration of the 6-month study period.

When asked about how easy or difficult it would be to fit using the smart pill bottle into their everyday life, one participant said:

> *I don't think it would be difficult at all...I mean I think it would be something that, let's say that pharmacy gave to me and said here do you want to try this out, I would, I would absolutely try it.* [P42]

When asked about how easy or difficult it would be to fit using the smart pill box into their everyday life, one participant said:

> *I think it'd be fairly easy to use...It's, it's got every day's medication. I see it goes by the day. It's going*

XSL•FO
RenderX

to tell me if I took it. And it's not that big or bulky...I would say it's of decent size. [P71]

Round 5 Interviews

In general, participants were willing to use the entire MMI system for the duration of the 6-month study period. However, participants reported that the survey (see below) they completed was too long; several participants requested that the survey be shortened. In general, participants were comfortable with the EMA app collecting various streams of data and preferred the option to select which streams of data were collected or not collected. However, 2 participants described the collection of data streams as "intrusive."

When asked about their willingness to use the entire MMI system for 6 months of data collection, one participant said:

Yeah. I would...Because I think it's important to find ways to help other people. I think that's important, but yeah; and I like the fact that I can opt out of certain things. [P83]

Study 2

Survey and EMA Evaluation

As described by Kaur et al [19], data from 32 breast cancer survivors who completed surveys at all 3 time periods (baseline, 4 mo, and 8 mo; Table 2) were used in the RNN. Of the 32 breast cancer survivors who participated in the study, 29 (90.6%) had MEMS data for the entire 8-month period. As indicated by Kaur et al [19], the analysis included bottle-opening events before or after the survey periods to evaluate the algorithms' prediction performance. A total of 2604 data samples were adopted, and each one included subjective values from survey data and the ground truth of medication-taking behavior captured by the MEMS device.

Table . Sample demographics and clinical characteristics of study 2 participants[a].

Characteristic and variable	Value
Race (n=32), n (%)	
White	14 (44)
African American/Black	4 (13)
Asian	2 (6)
Native Hawaiian/Other Pacific Islander	1 (3)
Native American/Alaskan Native	1 (3)
Unknown	10 (31)
Ethnicity (n=32), n (%)	
Latin (o/a/e/x)	25 (78)
Not Latin (o/a/e/x)	7 (22)
Gender (n=32), n (%)	
Female	32 (100)
Preferred language (n=32), n (%)	
English	7 (22)
Spanish	24 (75)
Other	1 (3)
Breast cancer stage (n=32), n (%)	
DCIS[b]	5 (16)
Early breast cancer	25 (78)
Locally advanced breast cancer	2 (6)
Intervention arm (n=32), n (%)	
Patient navigation intervention	15 (47)
Usual care	17 (53)
Age in years (n=25), mean (SD)	51.6 (8.9)

[a]Adapted from Kaur et al [19].

[b]DCIS: ductal carcinoma in situ.

As described by Kaur et al [19], the developed RNN models outperformed traditional neural network models in terms of adherence prediction accuracy under 2 types of randomness: subjective values and decision rules randomness. After training

XSL•FO
RenderX

the adherence prediction models, we preferred to gain insights into the importance of different input variables toward medication adherence prediction. We extracted the weights on different input variables of the trained RNN models, which are equivalent to the contribution of the respective variables in the final outcome prediction. We had 346 individual items as input variables, and using a threshold of the 70th percentile for feature importance (absolute input variable weights), we ended up with 104 survey items with significant contributions to 4-week medication adherence. Within the context of SCT [15], RNN analysis findings identified the following constructs associated with daily medication adherence: (1) environment: patient-provider interactions; instrumental, emotional, and informational social support; barriers to health care; and patient satisfaction with care; (2) personal: perceptions of control/agency/self-efficacy, numerous medication-related side effects, spirituality, perceived stress, decision regret, and fatigue; and (3) behavioral: having a routine for medication taking.

With regard to systematic review findings [12], 42 surveys were reviewed as predictors of ET medication adherence. In addition, there were additional SCT [15] constructs found to be associated with ET medication adherence among breast cancer survivors, as indicated by a recent systematic review [12]: (1) environment: communication with a health care provider and health care provider relationship quality, and (2) personal: mood/depression, perceived cognitive function, and perceived susceptibility of recurrence.

When combined, the RNN analysis [19] and the review of existing literature resulted in the identification of 34 surveys that could be included in the MMI system during the planned deployment, measuring constructs identified in both approaches as important predictors of ET adherence. Given participant feedback on reducing the length of the survey to be administered (see Round 5 usability interview findings), 2 measures of cancer health-related quality of life were removed due to the lengthiness of the surveys and because it was believed that health-related quality of life was also assessed by other measures being considered for inclusion in the MMI system. The final survey for the MMI system consisted of 32 measures and demographic items (Table 3).

XSL•FO
RenderX

Table . Measures included in the multiscale modeling and intervention (MMI) system.

Construct	Measure	Description	Psychometric properties
Overall mental and physical health	PROMIS[a] Global Health	A 10-item measure of a person's overall evaluation of their health	The scale has evidence for construct validity and adequate internal consistency (Global Physical Health Cronbach α=0.81; Global Mental Health Cronbach α=0.86) [21].
Medical comorbidities	Self-Administered Comorbidity Measure	An instrument that assesses for medical comorbidities and functions	The instrument has evidence for construct validity and adequate test-retest reliability (intraclass correlation coefficient=0.94, Pearson *r*=0.81) [22].
Side effects and symptoms associated with endocrine therapy	Breast Cancer Prevention Trial Checklist	A 42-item measure of vasomotor symptoms, urinary incontinence, cognitive and mood symptoms, vaginal symptoms, and weight gain/appearance concerns	The subscales have evidence for construct validity and adequate internal consistency (Cronbach α>0.70) [23,24].
Sleep-related impairment	Sleep-Related Impairment Short Form 8a	An 8-item measure of perceptions of sleep alertness, sleepiness, and tiredness during usual waking hours, and the perceived functional impairments during wakefulness	The scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.91) [25].
Sleep disturbance	Sleep Disturbance—Short Form 8b	An 8-item measure of perceptions of sleep quality, sleep depth, and restoration associated with sleep	The scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.84) [25].
Fatigue	Fatigue Symptom Inventory	A 13-item measure of multiple aspects of fatigue, including severity, frequency, and interference with daily functioning	The scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.94) [26,27].
Region of pain	Pain Region Scale	A measure of degree of bother for pain in the stomach, head, neck or back, widespread, and in the arms, legs, or joints	Psychometric evidence is not available because the Pain Region scale was developed by the research team.
Pain intensity	PROMIS Pain Intensity 3 a Short Form	A 3-item measure of degree to which a person hurts	The scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.86) [28].
Catastrophic pain−related thoughts	Pain Catastrophizing Scale	A 13-item measure of catastrophic pain-related thoughts, including rumination, magnification, and helplessness	The scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.75-0.93) [29].
Hot flash interference	Hot flash−Related Daily Interference Scale	A 3-item measure of perceived hot flash interference, referring to the degree to which hot flashes interfere with 9 aspects of daily life and overall quality of life.	The scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.90) [30].
Cognitive function	PROMIS Cognitive Function-Short Form 8a	An 8-item measure of perceived cognitive deficits	The scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.89-0.97) [31].
Personality traits	Mini IPIP[b]	A 20-item measure of the following 5 personality dimensions: extraversion, agreeableness, conscientiousness, neuroticism, and intellect/imagination	The subscales have evidence for construct validity and adequate internal consistency (Cronbach α>0.70) [32].
Perceived control	Pearlin Mastery Scale	A 7-item measure of perceptions of control or the extent to which one regards one's life chances as being under one's own control in contrast to being fatalistically ruled	The scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.70) [33].

XSL•FO
RenderX

Construct	Measure	Description	Psychometric properties
Spirituality	The Functional Assessment of Chronic Illness Therapy-Spiritual Well-Being	A 12-item measure of spirituality of patients with chronic and/or life-threatening illness	The scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.80) [34].
Perceived stress	The Perceived Stress Scale	A 10-item measure of the degree to which individuals feel that situations in their lives are unpredictable, uncontrollable, and overloading.	The scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.80) [35].
Generalized anxiety	Generalized Anxiety Disorder Scale	A 7-item measure of anxiety symptoms within the previous 2 weeks	The scale has evidence for construct validity and adequate internal consistency (Cronbach α>.80) [36].
Depression	Patient Health Questionnaire Depression Scale	An 8-item measure of depressive symptoms within the previous 2 weeks	The scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.86) [37,38].
Fear of cancer recurrence	Fear of Cancer Recurrence Inventory-Short Form	A 9-item measure of fear of cancer recurrence, referring to the fear or worry that the cancer will return or progress in the same organ or in a different part of the body	The full scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.71) [39].
Adherence to endocrine therapy	The Medication Adherence Report Scale	A 5-item measure of adherence with endocrine therapy	The scale has some evidence for construct validity and adequate internal consistency (Cronbach α range=0.67-0.89) [40].
Self-efficacy in adherence to medications	Medication Adherence Self-Efficacy Scale	A 26-item measure of self-efficacy in adherence to prescribed medications	The original scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.95) [41].
Beliefs about medicines	The Beliefs about Medicines Questionnaire	An 18-item measure of commonly held beliefs about medicines	The scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.70) [42,43].
Perceived necessity of endocrine therapy	Perceptions of endocrine therapy necessity items	Two items are used to measure perceived necessity of endocrine therapy	Psychometric evidence is not available on the items, but they have been used in previous cancer-related studies.
Decision regret	The Decision Regret Scale	A 5-item measure of distress or remorse after choosing to take or not take endocrine therapy	The scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.70) [44].
Satisfaction with cancer health care	Patient Satisfaction with Cancer Care	An 18-item measure of perceived satisfaction with health care received since diagnosis of cancer	The measure has evidence for construct validity and high internal consistency (Cronbach α range=0.95 - 0.96) [45-47].
Emotional support	PROMIS Emotional Support	An 8-item measure of perceived feelings of being cared for and valued as a person, having confident relationships	The scale has adequate construct validity and internal consistency (Cronbach α=0.97) [48].
Instrumental support	PROMIS Instrumental Support	An 8-item measure of perceived availability of assistance with material, cognitive, or task performance	The scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.95) [48].
Informational support	PROMIS Informational Support	A 10-item measure of perceived availability of helpful information or advice	The scale has evidence for construct validity and adequate internal consistency (Cronbach α=0.96) [48].
Barriers to receiving care	The Barriers to Care Scale	A 28-item measure of barriers to health care related to geography/distance, medical and psychological issues, community stigma, and personal resources	The original scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.70) [49].

XSL•FO
RenderX

Construct	Measure	Description	Psychometric properties
Cancer self-efficacy	The Communication and Attitudinal Self-Efficacy	A 12-item measure of self-efficacy in dealing with cancer and related health services	The measure has evidence for construct validity and adequate internal consistency (Cronbach α>0.80) [50].
Participation in health care decisions and quality of relationship with health care provider	Participation in Decision and Relationship Quality Survey	A 9-item measure of support received from doctors and other health professionals, the patient's role in decision-making, information shared with the patient about potential side effects, and their experiences with feeling listened to and respected by their provider	Psychometric evidence is not available on the items, but they have been used in previous cancer-related studies.
Social-psychological aspects of patient-physician interactions	Interpersonal Processes of Care Survey: Short Form	An 18-item measure of the social-psychological aspects of the patient-physician interaction	The scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.70) [51].
Self-efficacy when interacting with physicians	Patient-Perceived Self-Efficacy	A 10-item measure of subjective sense of patients' confidence when interacting with their physician	The scale has evidence for construct validity and adequate internal consistency (Cronbach α>0.90) [52].

[a]PROMIS: patient-reported outcomes measurement information system.

[b]Mini IPIP: Mini International Personality Item Pool.

Summary of the MMI System

The sensors selected for the final MMI system include an RxCap smart pill bottle and its associated smartphone app, the Fitbit Sense smartwatch and its associated smartphone app, and the Sensus Mobile smartphone app, which includes smartphone sensors and can administer EMA surveys. The timing of the EMA surveys includes morning and evening surveys, randomized surveys, surveys administered every 2 weeks, and surveys administered every 4 weeks. Additionally, the study team identified 32 constructs that will be included in the long-form baseline, 3-month, and 6-month surveys for the MMI system.

Discussion

Principal Findings

The purpose of this study was to describe the development of the MMI system, a novel approach to improving adherence to ET among early-stage breast cancer survivors that aims to understand medication adherence barriers and medication-taking behaviors. Overall usability testing interview findings suggest that participants were willing to use each component of the entire MMI system for the future 6-month-long study duration. A secondary analysis of data identified 104 survey items with significant contributions to 4-week medication adherence. Forty-two additional surveys were reviewed as predictors of ET medication adherence, and as per the SCT and other health and communication theories, included personal factors, environmental factors, and the behavior itself. When combined, the RNN analysis [19] and the review of existing literature resulted in the identification of 34 surveys that could be included in the MMI system during the planned deployment, measuring constructs identified in both approaches as important predictors of ET adherence.

Comparison to Prior Work

While interventions that aimed to improve ET adherence have been found to be effective overall, there is significant heterogeneity in the effects of psychosocial interventions, with reminder interventions being the most consistently helpful [13]. The lack of effectiveness of some psychosocial ET adherence intervention approaches suggests a lack of understanding of the barriers to ET adherence and the context of medication-taking behavior. The MMI system is designed to overcome challenges in improving adherence to daily ET among breast cancer survivors in stages 0 to 3 who have completed surgery, chemotherapy, and radiation. Personalized approaches to improving medication-taking behaviors may address the varied efficacy of prior psychosocial interventions to increase ET adherence among breast cancer survivors [13] by understanding personal, environmental, and behavioral factors as described in the SCT, along with understanding reciprocal determinism from the SCT, that is, the fact that all of the factors (including the behavior and medication adherence) influence each other in a dynamic way.

Strengths and Limitations

The strengths of the current work are that we used a novel and comprehensive approach to creating an MMI system through RNN analysis findings, usability interviews, and a review of research literature of constructs and the measurement of constructs associated with ET adherence among breast cancer survivors. Additionally, our research highlights the use of theory, data-driven models, and participant feedback to inform intervention development. This is not without limitations. First, the breast cancer survivors recruited during the usability interviews, which occurred during the COVID-19 pandemic lockdowns, were primarily White. Thus, the usability interview findings may not generalize to other ethnic/racial groups. However, the secondary analysis participants were diverse, and the study team will focus on recruiting more diverse participants for phase 2 of the study (see Figure 1) to mitigate this limitation.

XSL•FO
RenderX

Second, participants were primarily recruited via social media posts, so they may be inherently different, such as potentially being more familiar with and/or eager to adopt technologies included in the MMI system, compared to the general population of breast cancer survivors. To mitigate this limitation in the phase 2 deployment, the study team has increased recruitment strategies to include breast cancer survivors from various sources (ie, ResearchMatch and in-person recruitment). Third, the secondary data analysis was conducted on diverse female breast cancer survivors as part of a different intervention study; therefore, the original survey development was not intended to be used for the analytic purposes used in this study. Therefore, we were unable to include newly identified modifiable factors in the secondary data analysis. However, the purpose of the RNN was to develop and test the analytical framework that provides critical insight and refinement into the desired models to be run in phases 2 and 3 (Figure 1) of the project. Fourth, the analysis of the interview transcripts was completed by a single reviewer. Future research may benefit from incorporating additional reviewers into the analysis of the transcripts to better confirm the findings.

Future Directions

The next phase of the project (phase 2) will include the deployment of the entire MMI system with 20 breast cancer survivors over 6 months, collecting data on constructs from multiple health-related theories. Using the data collected during the planned phase 2 deployment, the RNN analysis will be used to identify the most important factors that contribute to ET nonadherence at various time frames among a new sample of breast cancer survivors, including more dynamic features, such as medication-taking patterns, data collected from wrist-worn sensors, smartphone sensors, and EMA. In subsequent deployment (phases 2 and 3, see Figure 1), we will continue to collect information on perceptions of devices, systems, and surveys in the MMI system to better refine it. The study team will then work toward reducing the amount of data collected to essential surveys, EMA, and sensors. Furthermore, the study team will begin to develop intervention components based on the group analysis and individualized model analyses in preparation for phase 3, which is the deployment of the refined MMI system and the implementation of the individualized intervention components that deliver intervention content based on a person's particular risks for ET nonadherence across different time frames (eg, daily and weekly). These intervention strategies may draw from multiple health and communication theories and can be matched to app features (eg, information reminders, reinforcement, and praise), which will likely improve medication adherence across different timescales [53]. Future large-scale testing of the MMI system and intervention approaches will be required to determine its effectiveness in increasing ET adherence among breast cancer survivors.

Acknowledgments

The authors would like to acknowledge Maddie Becker for her help with qualitative transcriptions of participant usability testing interviews.

Funding

The research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award numbers R01CA239246 and R21CA161077. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The author CG-A was funded by the UC San Diego Strategic Enhancement of Excellence through Diversity (SEED) Fellowship and the San Diego State University (SDSU) University Graduate Fellowship. The author MG was funded by a diversity supplement to the National Cancer Institute under award number R01CA239246 and by the SDSU University Graduate Fellowship.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: LEB, KJW
Data curation: MG, CG-A, NK, ANB
Formal analysis: MG, NK
Funding acquisition: LEB, KJW
Investigation: MG, CG-A
Supervision: LEB, KJW
Writing – original draft: MG, CG-A, NK, LEB, KJW
Writing – review & editing: MG, CG-A, NK, ANB, SL, LEB, KJW

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Characteristics of wearable sensor and medication event monitoring system devices reviewed.
[PDF File, 108 KB - cancer_v12i1e68255_app1.pdf]

Multimedia Appendix 2
Interview guides.
[DOCX File, 28 KB - cancer_v12i1e68255_app2.docx]

Checklist 1
COREQ checklist.
[PDF File, 452 KB - cancer_v12i1e68255_app3.pdf]

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin 2025;75(1):10-45. [doi: 10.3322/caac.21871] [Medline: 39817679]
3. Giaquinto AN, Sung H, Miller KD, et al. Breast cancer statistics, 2022. CA Cancer J Clin 2022 Nov;72(6):524-541. [doi: 10.3322/caac.21754] [Medline: 36190501]
4. Engelhardt EG, Garvelink MM, de Haes J, et al. Predicting and communicating the risk of recurrence and death in women with early-stage breast cancer: a systematic review of risk prediction models. J Clin Oncol 2014 Jan 20;32(3):238-250. [doi: 10.1200/JCO.2013.50.3417] [Medline: 24344212]
5. PDQ® Adult Treatment Editorial Board. Breast cancer treatment. National Cancer Institute (NCI). 2024. URL: https://www.cancer.gov/types/breast/patient/breast-treatment-pdq [accessed 2025-04-17]
6. Trayes KP, Cokenakes SEH. Breast cancer treatment. Am Fam Physician 2021 Aug 1;104(2):171-178. [Medline: 34383430]
7. Runowicz CD, Leach CR, Henry NL, et al. American Cancer Society/American Society of Clinical Oncology breast cancer survivorship care guideline. CA Cancer J Clin 2016;66(1):43-73. [doi: 10.3322/caac.21319] [Medline: 26641959]
8. Heery M. Precautions for patients taking tamoxifen or aromatase inhibitors. J Adv Pract Oncol 2023 Sep;14(6):478-482. [doi: 10.6004/jadpro.2023.14.6.2] [Medline: 37808073]
9. Davies C, Godwin J, Early Breast Cancer Trialists' Collaborative Group (EBCTCG), et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011 Aug 27;378(9793):771-784. [doi: 10.1016/S0140-6736(11)60993-8] [Medline: 21802721]
10. Van Liew JR, Christensen AJ, de Moor JS. Psychosocial factors in adjuvant hormone therapy for breast cancer: an emerging context for adherence research. J Cancer Surviv 2014 Sep;8(3):521-531. [doi: 10.1007/s11764-014-0374-2] [Medline: 24986227]
11. Moon Z, Moss-Morris R, Hunter MS, Carlisle S, Hughes LD. Barriers and facilitators of adjuvant hormone therapy adherence and persistence in women with breast cancer: a systematic review. Patient Prefer Adherence 2017;11:305-322. [doi: 10.2147/PPA.S126651] [Medline: 28260867]
12. Toivonen KI, Williamson TM, Carlson LE, Walker LM, Campbell TS. Potentially modifiable factors associated with adherence to adjuvant endocrine therapy among breast cancer survivors: a systematic review. Cancers (Basel) 2020 Dec 31;13(1):107. [doi: 10.3390/cancers13010107] [Medline: 33561076]
13. Bright EE, Finkelstein LB, Nealis MS, et al. A systematic review and meta-analysis of interventions to promote adjuvant endocrine therapy adherence among breast cancer survivors. J Clin Oncol 2023 Oct 1;41(28):4548-4561. [doi: 10.1200/JCO.23.00697] [Medline: 37531593]
14. Baglione AN, Gong J, Boukhechba M, Wells KJ, Barnes LE. Leveraging mobile sensing to understand and develop intervention strategies to improve medication adherence. IEEE Pervasive Comput 2020;19(3):24-36. [doi: 10.1109/MPRV.2020.2993993] [Medline: 33510585]
15. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory: Prentice-Hall; 1986.
16. Campbell S, Greenwood M, Prior S, et al. Purposive sampling: complex or simple? Research case examples. J Res Nurs 2020 Dec;25(8):652-661. [doi: 10.1177/1744987120927206] [Medline: 34394687]
17. Tandon A, Cobb B, Centra J, et al. Human factors, human-centered design, and usability of sensor-based digital health technologies: scoping review. J Med Internet Res 2024 Nov 15;26:e57628. [doi: 10.2196/57628] [Medline: 39546781]
18. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
19. Kaur N, Gonzales M, Garcia Alcaraz C, Barnes LE, Wells KJ, Gong J. Theory-guided randomized neural networks for decoding medication-taking behavior. IEEE EMBS Int Conf Biomed Health Inform 2021 Jul;2021. [doi: 10.1109/bhi50953.2021.9508614] [Medline: 34505062]

XSL•FO
RenderX

20. Boukhechba M, Baee S, Nobles AL, Gong J, Wells K, Barnes LE. A social cognitive theory-based framework for monitoring medication adherence applied to endocrine therapy in breast cancer survivors. IEEE EMBS Int Conf Biomed Health Inform 2018 Mar;2018:275-278. [doi: 10.1109/BHI.2018.8333422] [Medline: 29862383]
21. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Qual Life Res 2009 Sep;18(7):873-880. [doi: 10.1007/s11136-009-9496-9] [Medline: 19543809]
22. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Rheum 2003 Apr 15;49(2):156-163. [doi: 10.1002/art.10993] [Medline: 12687505]
23. Alfano CM, McGregor BA, Kuniyuki A, et al. Psychometric properties of a tool for measuring hormone-related symptoms in breast cancer survivors. Psychooncology 2006 Nov;15(11):985-1000. [doi: 10.1002/pon.1033] [Medline: 16470891]
24. Cella D, Land SR, Chang CH, et al. Symptom measurement in the Breast Cancer Prevention Trial (BCPT) (P-1): psychometric properties of a new measure of symptoms for midlife women. Breast Cancer Res Treat 2008 Jun;109(3):515-526. [doi: 10.1007/s10549-007-9682-9] [Medline: 17851765]
25. Lei DK, Yousaf M, Janmohamed SR, et al. Validation of Patient-Reported Outcomes Information System sleep disturbance and sleep-related impairment in adults with atopic dermatitis. Br J Dermatol 2020 Nov;183(5):875-882. [doi: 10.1111/bjd.18920] [Medline: 32017011]
26. Hann DM, Denniston MM, Baker F. Measurement of fatigue in cancer patients: further validation of the Fatigue Symptom Inventory. Qual Life Res 2000;9(7):847-854. [doi: 10.1023/a:1008900413113] [Medline: 11297027]
27. Hann DM, Jacobsen PB, Azzarello LM, et al. Measurement of fatigue in cancer patients: development and validation of the Fatigue Symptom Inventory. Qual Life Res 1998 May;7(4):301-310. [doi: 10.1023/a:1024929829627] [Medline: 9610214]
28. Stephan A, Stadelmann VA, Leunig M, Impellizzeri FM. Measurement properties of PROMIS short forms for pain and function in total hip arthroplasty patients. J Patient Rep Outcomes 2021 May 30;5(1):41. [doi: 10.1186/s41687-021-00313-1] [Medline: 34056667]
29. Osman A, Barrios FX, Kopper BA, Hauptmann W, Jones J, O'Neill E. Factor structure, reliability, and validity of the Pain Catastrophizing Scale. J Behav Med 1997 Dec;20(6):589-605. [doi: 10.1023/a:1025570508954] [Medline: 9429990]
30. Carpenter JS. The Hot Flash Related Daily Interference Scale: a tool for assessing the impact of hot flashes on quality of life following breast cancer. J Pain Symptom Manage 2001 Dec;22(6):979-989. [doi: 10.1016/s0885-3924(01)00353-0] [Medline: 11738160]
31. Henneghan AM, Van Dyk K, Zhou X, et al. Validating the PROMIS cognitive function short form in cancer survivors. Breast Cancer Res Treat 2023 Aug;201(1):139-145. [doi: 10.1007/s10549-023-06968-2] [Medline: 37330430]
32. Donnellan MB, Oswald FL, Baird BM, Lucas RE. The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. Psychol Assess 2006 Jun;18(2):192-203. [doi: 10.1037/1040-3590.18.2.192] [Medline: 16768595]
33. Pearlin LI, Schooler C. The structure of coping. J Health Soc Behav 1978 Mar;19(1):2-21. [Medline: 649936]
34. Peterman AH, Fitchett G, Brady MJ, Hernandez L, Cella D. Measuring spiritual well-being in people with cancer: the functional assessment of chronic illness therapy--Spiritual Well-being Scale (FACIT-Sp). Ann Behav Med 2002;24(1):49-58. [doi: 10.1207/S15324796ABM2401_06] [Medline: 12008794]
35. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
36. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
37. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
38. Razykov I, Ziegelstein RC, Whooley MA, Thombs BD. The PHQ-9 versus the PHQ-8--is item 9 useful for assessing suicide risk in coronary artery disease patients? Data from the Heart and Soul Study. J Psychosom Res 2012 Sep;73(3):163-168. [doi: 10.1016/j.jpsychores.2012.06.001] [Medline: 22850254]
39. Lebel S, Simard S, Harris C, et al. Empirical validation of the English version of the Fear of Cancer Recurrence Inventory. Qual Life Res 2016 Feb;25(2):311-321. [doi: 10.1007/s11136-015-1088-2] [Medline: 26341969]
40. Chan AHY, Horne R, Hankins M, Chisari C. The Medication Adherence Report Scale: a measurement tool for eliciting patients' reports of nonadherence. Br J Clin Pharmacol 2020 Jul;86(7):1281-1288. [doi: 10.1111/bcp.14193] [Medline: 31823381]
41. Ogedegbe G, Mancuso CA, Allegrante JP, Charlson ME. Development and evaluation of a medication adherence self-efficacy scale in hypertensive African-American patients. J Clin Epidemiol 2003 Jun;56(6):520-529. [doi: 10.1016/s0895-4356(03)00053-2] [Medline: 12873646]
42. Horne R, Weinman J. Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. J Psychosom Res 1999 Dec;47(6):555-567. [doi: 10.1016/s0022-3999(99)00057-4] [Medline: 10661603]

XSL•FO
RenderX

43. Horne R, Weinman J. Self-regulation and self-management in asthma: exploring the role of illness perceptions and treatment beliefs in explaining non-adherence to preventer medication. Psychol Health 2002 Jan;17(1):17-32. [doi: 10.1080/08870440290001502]
44. Brehaut JC, O'Connor AM, Wood TJ, et al. Validation of a decision regret scale. Med Decis Making 2003;23(4):281-292. [doi: 10.1177/0272989X03256005] [Medline: 12926578]
45. Fiscella K, Whitley E, Hendren S, et al. Patient navigation for breast and colorectal cancer treatment: a randomized trial. Cancer Epidemiol Biomarkers Prev 2012 Oct;21(10):1673-1681. [doi: 10.1158/1055-9965.EPI-12-0506] [Medline: 23045542]
46. Jean-Pierre P, Fiscella K, Freund KM, et al. Structural and reliability analysis of a patient satisfaction with cancer-related care measure: a multisite patient navigation research program study. Cancer 2011 Feb 15;117(4):854-861. [doi: 10.1002/cncr.25501] [Medline: 20922802]
47. Wells KJ, Winters PC, Jean-Pierre P, et al. Effect of patient navigation on satisfaction with cancer-related care. Support Care Cancer 2016 Apr;24(4):1729-1753. [doi: 10.1007/s00520-015-2946-8] [Medline: 26438146]
48. Rung Wu J, Chen X, Iwanaga K, et al. Psychometric validation of the PROMIS Social Support Scale in a sample of individuals with chronic health conditions and disabilities: a factor analytic study. Rehabil Couns Bull 2024 Oct;68(1):37-46. [doi: 10.1177/00343552231199243]
49. Heckman TG, Somlai AM, Peters J, et al. Barriers to care among persons living with HIV/AIDS in urban and rural areas. AIDS Care 1998 Jun;10(3):365-375. [doi: 10.1080/713612410] [Medline: 9828979]
50. Wolf MS, Chang CH, Davis T, Makoul G. Development and validation of the Communication and Attitudinal Self-Efficacy scale for cancer (CASE-cancer). Patient Educ Couns 2005 Jun;57(3):333-341. [doi: 10.1016/j.pec.2004.09.005] [Medline: 15893217]
51. Stewart AL, Nápoles-Springer AM, Gregorich SE, Santoyo-Olsson J. Interpersonal processes of care survey: patient-reported measures for diverse groups. Health Serv Res 2007 Jun;42(3 Pt 1):1235-1256. [doi: 10.1111/j.1475-6773.2006.00637.x] [Medline: 17489912]
52. Maly RC, Frank JC, Marshall GN, DiMatteo MR, Reuben DB. Perceived efficacy in patient-physician interactions (PEPPI): validation of an instrument in older persons. J Am Geriatr Soc 1998 Jul;46(7):889-894. [doi: 10.1111/j.1532-5415.1998.tb02725.x] [Medline: 9670878]
53. Geuens J, Geurts L, Gerling K, Croon RD, Abeele VV. A dyad of lenses for the motivational design of mHealth: bridging the gap between health theory and app design. Presented at: 2019 IEEE International Conference on Healthcare Informatics (ICHI); Jun 10-13, 2019. [doi: 10.1109/ICHI.2019.8904839]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
EMA: ecological momentary assessment
ET: endocrine therapy
IRB: institutional review board
MEMS: medication event monitoring system
MMI: multiscale modeling intervention
RCT: randomized controlled trial
RNN: randomized neural network
SCT: social cognitive theory

Edited by M Balcarras; submitted 31.Oct.2024; peer-reviewed by C Baxter, E Evbuomwan; revised version received 12.Mar.2026; accepted 12.Mar.2026; published 30.Apr.2026.

Please cite as:
Gonzales IV M, Garcia-Alcaraz C, Kaur N, Baglione AN, Livermon S, Barnes LE, Wells KJ
Improving Long-Term Adherence to Endocrine Therapy Among Breast Cancer Survivors: Development of a Multiscale Modeling and Intervention System
JMIR Cancer 2026;12:e68255
URL: https://cancer.jmir.org/2026/1/e68255
doi:10.2196/68255

© Manuel Gonzales IV, Cristian Garcia-Alcaraz, Navreet Kaur, Anna N Baglione, Sarah Livermon, Laura E Barnes, Kristen J Wells. Originally published in JMIR Cancer (https://cancer.jmir.org), 30.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is

XSL•FO
RenderX

properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Feasibility and Implementation of an eHealth Dashboard for the Remote Monitoring of Dutch Patients With Chronic Myeloid Leukemia: Multimethods Approach

Lynn Verweij[1*], PhD; Julie E M Swillens[2*], PhD; Sanne J J P M Metsemakers[1], MSc; Yolba Smit[1], MD, PhD; Esri Wener[1], MSc; Geneviève I C G Ector[3], MD, PhD; Rosella P M G Hermens[2*], PhD; Nicole M A Blijlevens[1*], MD, PhD

[1]Department of Hematology, Radboud Institute for Health Sciences, Radboud University Medical Center, Geert Grooteplein Zuid 8, Nijmegen, The Netherlands

[2]IQ Health Science Department, Radboud University Medical Center, Nijmegen, The Netherlands

[3]Department of Internal Medicine, Rijnstate Hospital, Arnhem, The Netherlands

[*]these authors contributed equally

Corresponding Author:
Lynn Verweij, PhD
Department of Hematology, Radboud Institute for Health Sciences, Radboud University Medical Center, Geert Grooteplein Zuid 8, Nijmegen, The Netherlands

Abstract

Background: Chronic myeloid leukemia (CML) has evolved into a chronic condition as a consequence of effective tyrosine kinase inhibitor (TKI) therapy, leading to an expanding demographic of patients necessitating lifelong monitoring. The use of eHealth solutions has the potential to facilitate sustainable and patient-centered care by enabling remote monitoring and enhancing guideline adherence. The Dutch CMyLife digital care platform incorporates a CML Dashboard intended for health care professionals (HCPs). This dashboard is designed to provide insight into real-world CML care and enable remote monitoring.

Objective: This study aimed to evaluate the feasibility of the CML Dashboard for remote monitoring of CML care, to assess the usefulness of dashboard-derived data, and to identify barriers and facilitators for its implementation in routine clinical practice.

Methods: We conducted a multimethods early-stage evaluation, determining the feasibility of the CML Dashboard and its barriers and facilitators. Quantitative data were generated through the CMyLife app and displayed in the CML Dashboard. These data were collected over a 2.5-year period and analyzed descriptively. Concurrently, semistructured interviews were conducted with HCPs treating patients with CML to explore barriers and facilitators for implementation. The analysis of interview transcripts was conducted using framework analysis, using established implementation models.

Results: Of the 199 patients who were registered, 177 provided data to the dashboard. The dashboard provided insight into patient characteristics, TKI treatment, BCR::ABL1 values, and monitoring intervals. However, the completeness and reliability of the data were limited, as the majority of the data were manually entered by the patients. The absence of complete data and the presence of inconsistencies in monitoring time points reduced the immediate usability of the dashboard for clinical decision-making and quality monitoring. A total of 8 HCPs participated in the interviews. The facilitators included an enhanced overview of patient data, support for consultation preparation, and the potential for remote monitoring and quality improvement. Key barriers identified were the lack of integration with electronic medical records, the necessity for additional log-in procedures, the limited adoption of the app by patients, and the perception of its added value as being inferior to existing systems.

Conclusions: The prospect of implementing a remote monitoring system for CML care holds promise. Such a system, using a digital dashboard, has the potential to enhance oversight and contribute to the enhancement of quality assurance. Nonetheless, the implementation of automated data exchange with electronic medical records, enhanced data completeness, and the augmentation of integration into clinical workflows are imperative prior to the large-scale implementation and the transition toward home-based CML care.

(JMIR Cancer 2026;12:e76096) doi:10.2196/76096

KEYWORDS

digital care platform; remote monitoring; patient empowerment; barriers and facilitators; CMyLife

XSL•FO
RenderX

Introduction

The use of health care resources and the associated costs are increasing due to an aging population and a growing number of patients with chronic diseases, including malignant ones [1]. Consequently, the development of more efficient care delivery methods is imperative to ensure the improvement and sustainability of future care. Remote monitoring is a method that enables patients to manage their own disease trajectory. A number of studies have demonstrated the efficacy of remote monitoring in the treatment of chronic cardiovascular diseases and diabetes mellitus [2,3].

Another chronic disease for which remote monitoring shows great promise is chronic myeloid leukemia (CML). The advent of tyrosine kinase inhibitors (TKIs) in 2001 transformed the management of CML, culminating in a 5-year survival rate of 89% [4] and a reduction in the loss of life years reduced to less than 3 years [5]. Therefore, the global prevalence of patients with CML is increasing significantly, and it is projected to become one of the most prevalent hematological malignancies by 2050 [6]. The feasibility of remote monitoring is contingent upon the regular monitoring of the BCR::ABL1 biomarker, which serves as an indicator of disease activity [7].

Literature showed that the current in-hospital monitoring of patients with CML, as outlined in evidence-based CML guidelines, is deemed to be suboptimal [8-10]. The distribution of expertise in the treatment of patients with CML is diffused, resulting in suboptimal biomarker monitoring and variability in guideline adherence [9,11,12]. This finding suggests that reduced adherence to TKIs may contribute to disease progression and increased health care expenditures [13]. Moreover, the timely recognition of adverse effects by health care professionals (HCPs) is crucial for enhancing treatment adherence, as studies have demonstrated that the adverse effects of TKI treatment were the primary factor contributing to intentional noncompliance [14].

The extant literature suggests that the active involvement of patients in their own care, for instance, through eHealth interventions, can positively affect these outcomes. Research has demonstrated that patient engagement has the potential to enhance medication adherence, facilitate early reporting and management of adverse effects, and support more consistent monitoring according to guidelines [15-17].

eHealth tools have the potential to enhance cooperation between patients and HCPs, thereby empowering patients to assume a more active role in their own care. This shift can lead to a reduction in the use of in-person care services and an increase in remote health care services, without compromising the quality of care received [18]. The CMyLife platform was developed with and for patients diagnosed with CML and is centered on the patient [19]. The platform aims to inform and empower patients and to provide them with the tools necessary to manage their disease. The system incorporates a variety of features, including the cmyCML website and the CMyLife app for patients, as well as the CML Dashboard for HCPs. These features enable patients and their HCPs to respond more appropriately in the event that their BCR::ABL1 results are suboptimal or side effects are too severe, according to CML guideline milestones. Previous research showed that the cmyCML website and mobile app enhanced information provision, as well as patient empowerment and adherence to guidelines and medications [17]. The CML Dashboard for HCPs represents a pioneering development in the field of CML care, as it is the first system to facilitate remote monitoring in real-world settings. To date, there is no literature on the use of such an eHealth solution for remote CML monitoring.

While eHealth solutions for remote monitoring offer opportunities to support chronic CML care, their successful implementation may be challenged by usability, data quality, and integration into routine clinical workflows. Therefore, the aim of this study is to explore the feasibility of the use of the Dutch CML Dashboard for remote patient monitoring, including the relevance of the CML Dashboard–derived data, and the identification of barriers and facilitators to implementation.

Methods

Study Design

We conducted an early-stage evaluation, using a multimethod approach to explore the feasibility of the CML Dashboard, the usefulness of the data, and the barriers and facilitators for implementation. A quantitative approach was used to explore user data. This study did not evaluate clinical effectiveness or quality improvement compared to standard care. A qualitative approach was used to explore the barriers and facilitators of the implementation of the CML Dashboard among HCPs. To determine these influencing factors, we organized semistructured interviews with HCPs who treat patients with CML.

Ethical Considerations

Upon initial access to the CMyLife and subsequent log-in, patients were required to provide explicit consent for the use of their data. The confidentiality and integrity of all data were ensured in accordance with the provisions of the General Data Protection Regulation and the Dutch security guideline (NEN7510). The Institutional Medical Ethical Committee “CMO Regio Arnhem-Nijmegen” has confirmed that this study is exempt from the provisions of the Medical Research Involving Human Subjects Act (WMO) because no traceable data were used and there was minimal impact on study participants (dossier number: 516006001). We used the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist to report the qualitative part of this research [20].

Setting

In the Netherlands, the delivery of care for patients with CML is predominantly characterized by a protocol-based approach, albeit with a dispersed distribution across health care facilities. Patients with CML are treated in 8 university hospitals and 68 general hospitals. Previous research has shown that monitoring of the BCR::ABL1 gene and the survival of patients with CML is better in hospitals with more experience in treating them [11]. Treatment of CML consists of oral TKIs, requires daily intake, and is subject to the Dutch “expensive medicine regulations.” For this reason, only medical specialists can prescribe TKIs. In the Netherlands, health insurance is mandatory and universal,

XSL•FO
RenderX

which ensures that everybody has access to CML treatment. In accordance with the evidence-based CML guidelines, the BCR::ABL1 levels of patients with CML need to be monitored throughout their lives. The latest Dutch guideline recommends to measure BCR::ABL1 values at least once every 3 months in the first year after diagnosis and 4 to 6 months each successive year [21]. Optimal response to treatment, as measured by the molecular response level, is achieved after 3 months, 6 months, and ≥12 months when BCR::ABL1 values are ≤10 (% international scale [IS]), ≤1 (% IS), and ≤0.1 (% IS), respectively (treatment milestones). In usual care, most patients do not get direct insight into their personal BCR::ABL1 values, achievement of milestones, or correct timing of BCR::ABL1 measurements. Outside studies, there is no validated tool for measuring side effects in order to monitor them in time, which could result in suboptimal treatment and delayed detection of intolerance.

Interventions

CMyLife Digital Care Platform

CMyLife [22] is a digital care platform codeveloped with and for patients with CML in 2016. The platform has been described in detail before [17,19]. However, since CMyLife is constantly being adjusted to the wishes and needs of its users, improvements have been made over time. CMyLife provides patients and HCPs with different features to arrange CML care according to (inter)national guidelines: the cmyCML website [23], the CMyLife app, and the CML Dashboard. The cmyCML website contains general information about CML, its treatment, and its impact on daily life. It also provides a forum where patients can connect with peers and the option to ask HCPs questions [23].

The CMyLife App

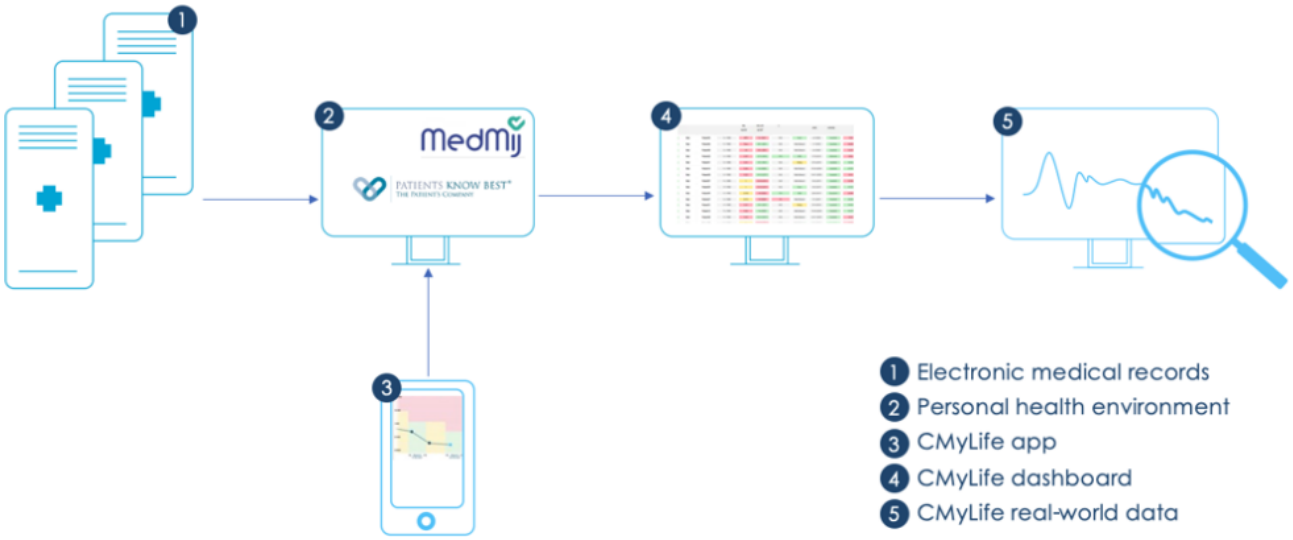
The CMyLife app provides personalized feedback based on the evidence-based Dutch guideline for each patient individually. This feedback is based on BCR::ABL1 results, guideline milestones, and monitoring schedules. Patients with atypical CML presentations or treatment trajectories that substantially deviated from standard guideline-based care were able to use the CMyLife app; however, they were excluded from receiving guideline-based visualizations and messages within the app. BCR::ABL1 values are registered by patients by hand in the app. Up to now, only BCR::ABL1 values from patients treated in one specific hospital were automatically sent from the laboratory to the personal health environment (PHE). For other hospitals, this connection is not live yet. In addition, the CMyLife app enables patients to set a reminder for medication intake in order to improve TKI adherence, and it is connected to patients' PHE to store personal health data. Data exchange was enabled through RESTful (Representational State Transfer) APIs and HL7-based standards (Health Level Seven International), in combination with a MedMij-certified PHE. A transition toward HL7 Fast Healthcare Interoperability Resources is being made in future versions to improve interoperability and scalability. The PHE is a secure and central digital cloud platform that allows patients to collect, manage, and share their medical data from various HCPs in one place. Supported by the government [24], this initiative is based on the MedMij standard [25], setting the standard for the application of a national Dutch PHE. The PHE that is used is Patients Know Best. This PHE respects the privacy and security of data, in accordance with applicable laws and regulations, including the General Data Protection [26]. Using the app, patients are informed about treatment goals and are better prepared for consultation with their HCPs [27].

The CML Dashboard

This study focused on the CML Dashboard for HCPs, which is based on data generated in the CMyLife app for patients. Figure 1 shows how data flows to the dashboard. The CML Dashboard was developed together with HCPs in the Netherlands in 2022 using the design-thinking method, in which HCPs actively participate in the development. In 2023, HCPs were introduced to this dashboard by the CMyLife team (consisting of CMyLife experts easily approachable for questions and help) and received information on how to access and use it. HCPs were able to use the dashboard according to their own needs. The dashboard provided HCPs an overview of their patients with CML using the CMyLife app. It provided them with an overview of CML treatment, insight into the disease course (BCR::ABL1 results) of their patients, guideline-driven monitoring appointments, and side effects of CML treatment. At a glance, HCPs could use real-world patient data in the dashboard to identify which patients need more attention (indicated by color-coded treatment milestones derived from evidence-based clinical guidelines), which patients are on a favorable disease course, and adjust the level of care accordingly.

XSL•FO
RenderX

Figure 1. Dataflow from patients and health care professionals to the dashboard. Dataflow from 1 to 2 is only possible in one hospital.

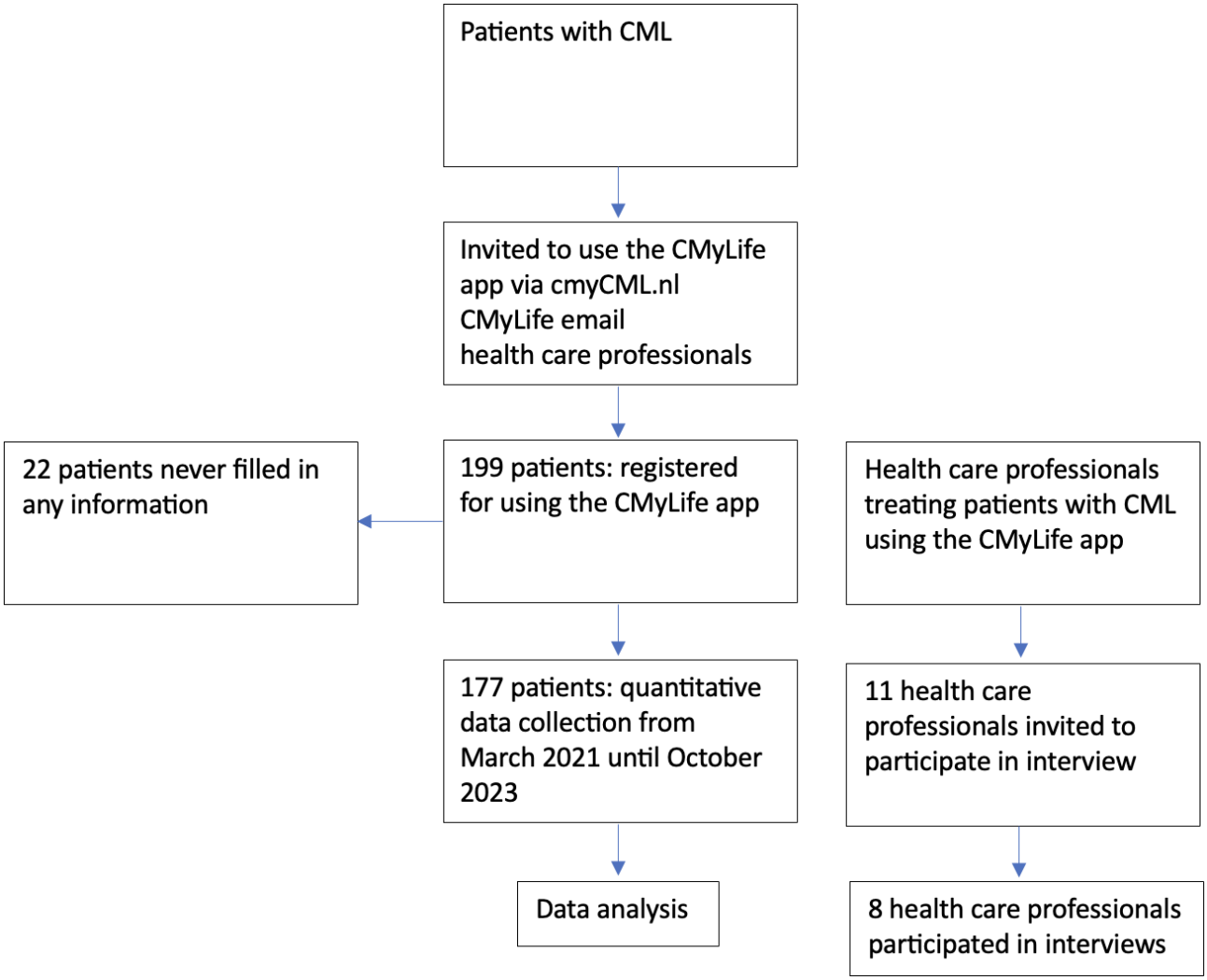


Study Population

Quantitative data were retrieved from the dashboard between March 2021 and October 2023. Patients became aware of the CMyLife app via their HCPs or through communication channels of the CMyLife platform. Patients from all hospitals treating CML could register for the CMyLife app via the cmyCML website, CMyLife email, or their HCPs. Once registered, the CMyLife team contacted the patient, introduced the app, and helped them install it, providing instructions on how to use it. To use the app, a PHE needs to be connected, and the team member helped with that as well. For the qualitative part, we recruited HCPs who provide care to patients with CML, such as hematologists and nurse specialists within hematology. Inclusion criteria were that these professionals treated patients with CML and had a private introductory meeting of the CML Dashboard by one of the members of the CMyLife team. Based on their prior introduction to the CML Dashboard, potential participants were recruited via convenience sampling by the CMyLife team to make an appointment for an interview. To ensure variability, we included HCPs from both university and general hospitals, as well as professionals with 5 to 30 years of experience in clinical practice. They received an invitation to an interview via email, including the information letter. If they were interested, they were then contacted by phone to arrange the interview. The recruitment of HCPs stopped when we had a diversified sample and no new opportunities, barriers, facilitators, or improvements were mentioned, so data saturation was achieved. Figure 2 shows the participant flowchart.

XSL•FO
RenderX

Figure 2. Participant flowchart. CML: chronic myeloid leukemia.



Data Collection

App data were sent to the dashboard via the PHE. In October 2023, all available descriptive user data from April 2018 until October 2023 were retrieved from the dashboard. Data consisted of basic characteristics such as age, gender, and current TKI used and other CML-related information, such as BCR::ABL1 values, TKI-related information (current and previous TKIs, dosage, duration of TKI therapy, and intake), BCR::ABL1 monitoring time points, side effects of CML treatment, messages sent to patients, and messages read by patients. The date of diagnosis could not be registered or traced due to all personal health data in the dashboard being secured in compliance with the General Data Protection Regulation (GDPR) and the Dutch guideline for health care–related information security (NEN7510). Only the date of starting TKI treatment was recorded. Access to patient data was restricted to authorized HCPs, and only nontraceable data were used for research purposes.

As for the qualitative part, we used a qualitative descriptive design within the pragmatic paradigm [28,29]. We composed an interview guide based on the theoretical frameworks of Flottorp et al [30] and Grol and Wensing [31]. These frameworks define barriers to and facilitators for change in health care practice at 6 different levels: (1) the innovation itself, (2) the patient, (3) the individual professional, (4) the organizational context, (5) the social context, and (6) the economic and political context. This guide included questions on actual use and potential use of the CML Dashboard in clinical practice, experiences with the CML Dashboard including barriers and facilitators for implementation, structured by the domains of the theoretical frameworks, and improvements of the CML Dashboard. The interview guide can be found in Multimedia Appendix 1. Although this research is based on the assumption that multiple subjective realities exist, the study used systematic data collection and reflexive practices to approach the material as consistently and objectively as possible. This balance allowed for openness to participants' meanings while maintaining methodological rigor. A pilot interview was conducted with a member of the CMyLife team who used the dashboard to test the interview guide. After testing, the interview guide was adjusted accordingly. There was no prior relationship between the interviewer and participants.

Before the start of the interview, informed consent was obtained. In addition, characteristics, such as gender, years of experience, type of hospital, and medical profession, were collected. With the use of this interview guide, semistructured online interviews via Microsoft Teams were conducted by a senior researcher

XSL•FO
RenderX

(JEMS) with experience in conducting interviews with HCPs about digital tools in health care. The interviews were also audiotaped using Microsoft Teams and lasted about 20 to 30 minutes.

Data Analysis

Descriptive statistics were used to explore the data generated from the CML Dashboard. Patients who did not use the CMyLife app were excluded from the analysis. Other incomplete data were not excluded because our goal was to explore what data and outcomes could be generated by the CML Dashboard. SPSS (IBM SPSS Statistics for Windows, version 29) was used for the analyses. Molecular response levels were calculated based on the BCR::ABL1 values. The first 5 monitoring time points were analyzed because they represent the first year of monitoring.

For qualitative analysis, automatically generated transcripts (verbatim by Microsoft Teams) were cleaned by a student to enable framework analysis. Transcripts were coded using ATLAS.ti GmbH version 23.1.1.0. Two researchers (JEMS and EW) independently coded the interview transcripts, using the framework of Flottorp et al [30] and Grol and Wensing [31] for coding the barriers and facilitators. The codes were discussed to reach a consensus. When agreement could not be achieved, a third researcher was consulted (RPMGH).

Results

Quantitative Results

Patient Characteristics

Of the 199 included patients, 22 patients did not fill in any information via the CMyLife app and therefore left no data in the dashboard. They were excluded, leaving 177 patients for analysis. Basic characteristics of these patients are summarized in Table 1. Most patients were male (105/177, 59.3%), and the mean age was 59.1 (SD 12.7) years.

Table . Basic characteristics of participating patients with chronic myeloid leukemia.

Characteristics	Patients with CML[a] (n=177)
Sex, n (%)	
Male	105 (59.3)
Female	70 (39.6)
Missing	2 (1.1)
Age (years), n (%)	
18 - 64	79 (44.6)
65 or older	46 (26.0)
Missing	52 (29.4)
Age (years), mean (SD)	59.1 (12.7)

[a]CML: chronic myeloid leukemia.

Treatment Data and Its Usefulness

Treatment characteristics are summarized in Table 2. Of the 177 patients, the first used TKI was registered by 146 (82.5%) patients. Currently used TKI registration was missing in 35 (19.8%) patients, and the first registered BCR::ABL1 values were missing in 11 (6.2%) patients. The 146 patients were first treated for their CML 5.3 years ago on average (minimum 0 years and maximum of 22 years). Patients who switched TKI or increased or decreased their dosage did so 173.6 weeks after starting the first TKI on average, and 79 weeks after the first switch, increase, or decrease. Almost 60% (87/146) of patients who registered their TKI never adjusted the TKI or TKI dosage. Nearly 27 (18.5%) of the 146 patients increased or decreased the dosage of their first used TKI. Of the 27 patients, the dosage of 11 (40.7%) patients was decreased, and the dosage of 16 (59.3%) was increased. TKIs were switched once, twice, three times, or four times in 26% (38/146), 8.2% (12/146), 4.8% (7/146), and 1.4% (2/146) of patients, respectively. Overall, 35 (19.8%) of the 177 patients did not fill in their currently used TKI. Lastly, dosages of medication were registered in two different variables. One variable indicated whether patients took their medication once or twice a day, while the other captured dosage in milligrams. It was unclear whether the reported dosage in milligrams referred to the total daily dose or to the amount taken twice a day. Patients did not fill in side effects regularly, accurately, and completely. Of the 177 patients, 120 (67.8%) registered side effects (mean 6.53, SD 16.3), 47 (26.6%) registered side effects only once, and 1 (0.6%) registered side effects 129 times.

XSL•FO
RenderX

Table . Treatment characteristics at registration of tyrosine kinase inhibitors.

Characteristics	Patients with CML[a] (n=177)
First TKI[b] used, n (%)	
Imatinib	101 (57.1)
Dasatinib	20 (11.3)
Nilotinib	20 (11.3)
Bosutinib	4 (2.3)
Ponatinib	1 (0.6)
Missing	31 (17.5)
Currently used TKI, n (%)	
Imatinib	59 (33.3)
Dasatinib	39 (22.0)
Nilotinib	27 (15.3)
Bosutinib	8 (4.5)
Ponatinib	4 (2.3)
Asciminib	4 (2.3)
Discontinued TKI therapy	1 (0.6)
Missing	35 (19.8)
Years since first treatment for CML	
Mean (SD)	5.3 (4.9)
Missing, n (%)	15 (8.5)
Duration TKI in weeks	
Duration first TKI, n (%)	93 (52.5)
Duration first TKI, mean (SD)	173.6 (221.2)
Duration second TKI, n (%)	56 (31.6)
Duration second TKI, mean (SD)	79.0 (112.3)
Duration third TKI, n (%)	30 (16.9)
Duration third TKI, mean (SD)	77.9 (101.2)
Duration fourth TKI, n (%)	19 (10.7)
Duration fourth TKI, mean (SD)	65 (148.5)
Duration fifth TKI, n (%)	11 (6.2)
Duration fifth TKI, mean (SD)	40.9 (35.2)
Switch TKI (n=146), n (%)	
Never switched	87 (59.6)
Switched once	38 (26.0)
Switched twice	12 (8.2)
Switched three times	7 (4.8)
Switched four times	2 (1.4)

[a]CML: chronic myeloid leukemia.

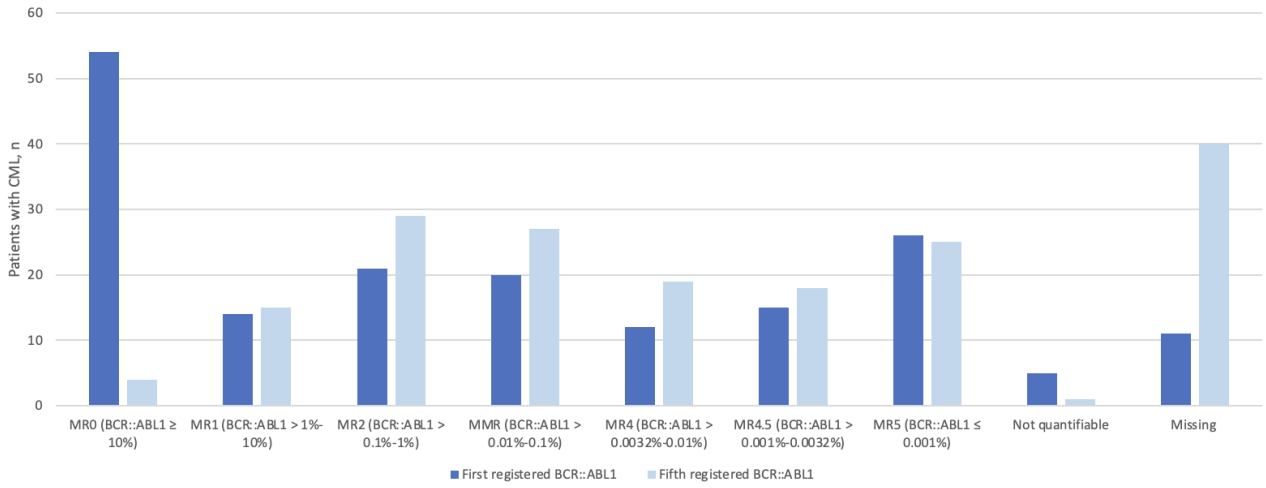
[b]TKI: tyrosine kinase inhibitor.

XSL•FO
RenderX

BCR::ABL1 Values, Monitoring Time Points, and Their Usefulness

In total, 168 patients registered 2297 BCR::ABL1 values (Table 3). The first and fifth registered BCR::ABL1 values are shown in Figure 3. The first registered BCR::ABL1 values ranged from 0% to 250% (mean 19.66%). Besides, 46.4% (78/168) of patients were in major molecular response (MMR) or lower (MR4, MR4.5, or MR5). The fifth registered BCR::ABL1 values ranged from 0% to 20.3% (mean 0.93%). The fifth registered BCR::ABL1 values showed half of the patients to be in MMR, or lower (MR4, MR4.5, or MR5).

Figure 3. Molecular response of the first and fifth registered BCR::ABL1 values (n=168). Not quantifiable indicates that the BCR::ABL1 value was at a level too low or too uncertain to be precisely quantified. CML: chronic myeloid leukemia; MMR: major molecular response.



XSL•FO
RenderX

Table . BCR::ABL1 values and molecular responses.

Characteristics	Patients with CML[a] (n=177)
First registered BCR::ABL1 values	
Min	0
Max	250
Mean (SD)	19.66 (39.32)
Missing, n (%)	11 (6.2)
Molecular response of the first registered BCR::ABL1, n (%)	
MR0 (BCR::ABL1 ≥ 10%)	54 (30.3)
MR1 (BCR::ABL1 > 1% - 10%)	14 (7.9)
MR2 (BCR::ABL1 > 0.1% - 1%)	21 (11.8)
MMR[b] (BCR::ABL1 > 0.01% - 0.1%)	20 (11.2)
MR4 (BCR::ABL1 > 0.0032% - 0.01%)	12 (6.7)
MR4.5 (BCR::ABL1 > 0.001% - 0.0032%)	15 (8.4)
MR5 (BCR::ABL1 ≤ 0.001%)	26 (14.6)
Not quantifiable	5 (2.8)
Missing	11 (6.2)
Molecular response of the fifth registered BCR::ABL1, n (%)	
MR0 (BCR::ABL1 ≥ 10%)	4 (2.2)
MR1 (BCR::ABL1 > 1% - 10%)	15 (8.4)
MR2 (BCR::ABL1 > 0.1% - 1%)	29 (16.3)
MMR (BCR::ABL1 > 0.01% - 0.1%)	27 (15.2)
MR4 (BCR::ABL1 > 0.0032% - 0.01%)	19 (10.7)
MR4.5 (BCR::ABL1 > 0.001% - 0.0032%)	18 (10.1)
MR5 (BCR::ABL1 ≤ 0.001%)	25 (14.0)
Not quantifiable	1 (0.6)
Missing	40 (22.5)

[a]CML: chronic myeloid leukemia.

[b]MMR: major molecular response.

In total, patients registered 2354 monitoring time points. Patients filled in a minimal 0 dates, and most registered monitoring time points were 61. The mean time between the first two monitoring time points was 159.9 days, with a minimum of 1 day and a maximum of 7380 days. The mean time between the second and third monitoring time points was 92.9 days, with a minimum of 0 days and a maximum of 900 days. The mean time between the third and fourth monitoring time points was 81.3 days, with a minimum of −282 days and a maximum of 364 days. The mean time between the fourth and fifth monitoring time points was 97.9 days, with a minimum of −267 days and a maximum of 735 days. A negative value means that the registered monitoring time point was a date before the previous registered time point.

Qualitative Results

HCP Characteristics

A total of 11 HCPs were invited to participate in an interview, of which 3 were not able to participate or did not respond. We conducted 8 interviews with HCPs, resulting in data saturation. Characteristics of study participants are shown in Table 4.

XSL•FO
RenderX

Table . Overview of study participants' characteristics (n=8).

Characteristics	Values
Sex, n (%)	
Male	4 (50)
Female	4 (50)
Years of experience, median (range)	16.5 (5-30)
Type of hospital, n (%)	
University hospital	3 (37)
General hospital	5 (63)
Medical profession, n (%)	
Hematologist	5 (63)
Physician assistant/nurse practitioner	3 (37)

Barriers and Facilitators

In Table 5, an overview of the barriers and facilitators of the CML Dashboard implementation is shown, categorized by the domains of Flottorp et al [30] and Grol and Wensing [31]. This overview also includes potential barriers and facilitators for its future implementation, for example, in the case of remote monitoring.

XSL•FO
RenderX

Table . Current and potential barriers and facilitators of the CML[a] Dashboard.

Facilitators	Barriers
Innovation factors	

Strengths of the CML Dashboard
• Overview at the health care professional level
• Overview at the patient level
  • Overview of patients' symptoms
  • Overview of patients' treatment/medication (including switches)
• Facilitates consultation preparation
• Facilitates consultation with patients
• Increases awareness and knowledge of the CML guideline
• Increases workflow efficiency
• Is a safety net for no-show patients
	Limited additional value/information of the CML Dashboard
	• …in comparison to the EMR[b]
	• …in comparison to personal contact with patients
	• …in comparison to other sources of CML clinical guideline information
	• …in comparison to the existing communication channels of the hospital
	• …for health care professionals with sufficient knowledge of the CML clinical guideline
	• …if treating a small number of patients with CML
	• …for straightforward clinical pathways

Potential strengths of the CML Dashboard
• Remote monitoring
• Audit and feedback on health care professionals, hospitals, regional and national levels for quality improvement
  • Monitoring medication safety
  • Monitoring patient visits/contacts
  • Monitoring patient symptoms
  • Research (database)	

Feasibility of the CML Dashboard • Presentation of patients' symptoms (colors and scale for severeness and detailed description) • Presentation of BCR::ABL values (colors in relation to CML treatment milestones) • Presentation of medication switch • Compliance with legislation	Limited feasibility of the CML Dashboard • Patient information not informative or of additional value • Additional screen needed (pop-up screen) • Inefficient layout • Messages (function) unclear • Slow system

Potential feasibility of the CML Dashboard
• Customizable layout/information view
• Alert/pop-up not reaching milestone BCR::ABL value
• Links to important websites/information included
  • Guideline information
  • Information contact other departments (eg, cardiology)
• Monitoring patient visits/contacts
• Scale up to other hematological diseases
• Information on study/trial participation
• Allocation of diverse access roles
• Demand-based patient communication
• Access data in the CML Dashboard via PHE[c]
• Personalized BCR::ABL1 cutoff values
• Self-ordering of lab tests
• Self-planning consult hospital	

Compatibility of the CML Dashboard • Insight into patients' symptoms	Limited compatibility of the CML Dashboard • No connection with/or integration in EMR • CML guideline not applicable to all patients • Lack of patients' filling in information in the CMyLife app • Information "overkill" for health care professionals due to diverse information sources • Lack of information on the use of the CMyLife app
Potential compatibility of the CML Dashboard • Integrated version of the CML Dashboard in the EMR • Reminder to patients to use the CMyLife app	
The CML Dashboard is deemed accessible	Limited accessibility of the CML Dashboard • Additional authenticator step • Additional log-in • Lack of instructions on how to log in

XSL•FO
RenderX

Facilitators	Barriers
Individual professional factors	
Agreement with the CML Dashboard • Positive attitude	*Not discussed*
Awareness and familiarity with the CML Dashboard • Experience with the dashboard in the development phase • Experience with dashboard patient consult (preparation) • Introduction to the system	Limited awareness and familiarity with the CML Dashboard • Little to no clinical experience with the CML Dashboard
Intention and motivation to use the CML Dashboard • Increased motivation when patients use the CMyLife app	No intention and motivation to use the CML Dashboard • Decreased motivation when patients do not use the app
Not discussed	Use of the CML Dashboard is not a routine practice
Expected outcome of using the CML Dashboard for remote monitoring[d] • Fewer consults/hospital visits[d] • More self-directed patients[d]	*Not discussed*
Patient factors	
Need to have more ownership in the care process in case of remote monitoring[d]	No need to have more ownership in the care process in case of remote monitoring[d]
Social context factors	
Encouraging others to start using the CML Dashboard	*Not discussed*
Referral process improved when the CML Dashboard was integrated into EMR[e]	*Not discussed*
Organizational context factors	
Available team to keep it up-to-date and to provide support service	*Not discussed*
Mandated organization remote monitoring[d]	*Not discussed*
Economic and political factors	
Reimbursement for remote monitoring[d]	*Not discussed*
Not discussed	Legal obligation to document consult[d]

[a]CML: chronic myeloid leukemia.

[b]EMR: electronic medical record.

[c]PHE: personal health environment.

[d]Barriers and facilitators for remote monitoring.

[e]Barrier for improvement: integrated version of the CML Dashboard in EMR.

Innovation Factors

The most influential factors related to the innovation itself concerned the CML Dashboard. These factors included the strength, feasibility, compatibility, and accessibility of the CML Dashboard. Multiple strengths of the CML Dashboard for current use were mentioned by participants, such as having an overview of patients with CML-related information on multiple levels. This information was used for both the preparation of a consultation and the consultation itself, and helped to not miss patients. The information provided increased the knowledge of the CML guidelines and the efficiency of the workflow. However, according to some HCPs, the information was of limited additional value to existing information sources and their own knowledge, especially when considering the low number of patients with CML and straightforward clinical pathways. The CML Dashboard was deemed feasible and compatible. Compared to other current available resources, the CML Dashboard was not deemed efficient in use because of some feasibility and compatibility issues. This was mainly related to the log-in to a stand-alone dashboard, including the authenticator step. Some HCPs experienced logging in to the dashboard as rather easy, while others thought it was not easy at all. Having multiple information channels across different screens (pop-ups) led some users to experience an overload of additional and double information when using the dashboard alongside the information already available. In addition, a lack of information on patient participation in the CMyLife app and patients' treatment plans that deviate from complying with guideline recommendations were experienced as barriers.

Individual Professional Factors

Several factors are related to individual HCPs. Most interviewees were satisfied with the general content of the

XSL•FO
RenderX

dashboard. However, the intention and motivation among participants was quite low to use the CML Dashboard in current clinical practice. Motivation was influenced by the limited use of the CMyLife app among patients. Most HCPs were aware of the dashboard. However, only some of them had used the dashboard in clinical practice during (preparation for) consultations. Last, participants found it hard to integrate the use of the dashboard into daily work routines.

Social Setting and Organizational Factors

Users of the CML Dashboard encouraged others to also start using the dashboard. Participants experienced a lack of guidance on how to access the dashboard. Therefore, a facilitator was an easily available team, responsible for services such as updates and customer support.

Recommendations to Current Use of the CML Dashboard: Potential Innovation Factors

In Table 5, potential options to enhance the strength, feasibility, and compatibility of the current CML Dashboard, mentioned by the interviewees, are shown. Regarding its strength, the dashboard could play a central role in the remote monitoring of patients with CML in the future. Facilitators of this new app mentioned were the expected outcomes of having fewer in-person consultations and patients becoming more self-directed, although opinions differed on whether patients want to be more self-directed. A barrier was seen for complying with remote monitoring to existing regulations, like taking notes of the patient, and to existing funding policies. In addition, HCPs wondered who the coordinating party should be when shifting care to the home environment. Next, the dashboard could be used as a quality control instrument on several levels: individually, locally, regionally, or even nationally. More specific applications mentioned were monitoring of clinical visits, monitoring of medication safety, and identifying general trends in patient symptoms, medication use, and their responses (national level). Lastly, using this database for research purposes was positively mentioned by an interviewee. With the potential options and purposes in mind, remote monitoring motivated HCP to start using the CML Dashboard. Other factors increasing their motivation were an increasing workload due to a foreseen growth in patients with CML, due to increased survival, and more use of the CMyLife app by patients with CML.

Next, to improve feasibility and compatibility, and to overcome the lack of stand-alone functionality, it was suggested to incorporate the dashboard within existing structures (electronic health records and PHE). This would also enable easier access to information of and referral to other medical specialists. With regard to the compliance with the guideline, additional features could be an alert of a changing/increasing BCR::ABL value, including the allowance of customized cutoff values, adding a severeness scale of patient symptoms, and links to important websites with additional information on CML care and its clinical guidelines. Regarding the overview function, the overview of patients could also be expanded to include patients with other hematological conditions. Furthermore, interviewees mentioned that it would be relevant to have information on study/trial participation, cardiovascular care/history, and on patients' use of the CMyLife app. For the visualization of this information, a customizable layout was suggested. For access to the information, allocation of diverse roles/access was mentioned as well as access of other HCPs to the dashboard (eg, cardiologists). The contact with patients can be expanded with demand-based communication and ordering functions, like planning a clinical visit or a lab test. Concerning this function, supporting staff should have access to the dashboard to be able to manage the appointments (laboratory and hospital) of patients with CML.

Discussion

General Findings

This study demonstrated an early-stage evaluation of a promising eHealth tool for future remote monitoring of patients with CML by the use of the CMyLife app and the CML Dashboard. By assessing the usefulness of data from the CML Dashboard at this stage, the study laid the groundwork for potential optimizations of the CML Dashboard. Quantitative data showed useful information about the patients using the CMyLife app, their therapy, BCR::ABL1 values, and BCR::ABL1 monitoring time points. However, the data had some limitations, and adjustments should be made regarding data provision (automatically from electronic medical records and improving instructions for patients) before it can be used accordingly. Although many strengths of the CML Dashboard were mentioned by the end users, the CML Dashboard was not yet widely adopted in clinical practice. The main reasons were the lack of compatibility with other information systems used in clinical practice such as electronic health records, feasibility issues such as difficulties with logging in into the CML Dashboard, and lack of additional information relative to other information resources, also hindered by the lack of use of the CMyLife app. Regarding potential applications, the CML Dashboard could have a central role in remote monitoring and quality assurance. However, organizational steps, such as improved data input and reimbursement, need to be taken first before CML care can be moved completely toward the home environment.

At present, imatinib, dasatinib, bosutinib, and nilotinib are available as first-line TKIs for the treatment of CML in the Netherlands, but drug safety profile, comorbidities, and impact on quality of life of patients should be considered while selecting any first-line TKI treatment [32-34]. To date, little real-world evidence is available about what TKI is most appropriate in what cases and the way HCPs make treatment decisions [35]. Moreover, the timing of switching TKIs remains a controversial topic [34]. Our quantitative results showed that the CML Dashboard can collect and therefore provide valuable insights into these data, and the use of the dashboard for quality monitoring was mentioned as an important application in the qualitative part. However, almost all data were registered by patients themselves, except for BCR::ABL1 values from patients treated in one specific hospital. This influences the usefulness of the data. BCR::ABL1 values registered by patients are less reliable as it is not certain if they recorded the values correctly and if they registered all values. After improvement in data input, increased use of both the CML Dashboard and the

XSL•FO
RenderX

CMyLife app may result in insight into relations between patient characteristics, TKIs, switches, and side effects in the real world, which could lead to better selection of treatment tailored to specific patient characteristics, eventually leading to improvement of CML guidelines and, therefore, quality of CML care.

Patients registered at the start of their first TKI treatment, but reliability could not be checked, which makes it hard to analyze response to treatment and reaching treatment milestones in accordance with the guidelines, as it is not clear what BCR::ABL1 values are after 3 months, 6 months, and ≥12 months after the start of treatment. Therefore, it would be of added value to add a question to the app that registers this information. Moreover, monitoring time points registered by patients have shown to be not reliable. The period between two time points was negative, which means that a registered monitoring time point was a date before the previous registered time point. Nevertheless, registered BCR::ABL1 values showed a decrease over time and no worsening of the disease.

In the future, according to the value-based health care principles of Porter [36], a widely supported compromised set of quality indicators based on the CML guideline should be agreed upon by involved stakeholders to monitor, or proactively redirect them according to guideline adherence in the remote setting. This compromised set of indicators will also be important for the other future application for the dashboard, quality monitoring on diverse levels (hospital, region, and national). When care of CML is organized more centrally, these expert hospitals can measure this set of indicators and together, by using audit and feedback, improve the quality of care for patients with CML. In addition, the data mentioned above could potentially help in proving relations between patient characteristics, TKIs, switches, and side effects in the real world. An important element of this quality indicator set will be a patient-reported outcome measure specifically aimed at patients with CML. Using this patient-reported outcome measure, it could enable the collection of structured data from a patient perspective. To date, no validated questionnaire for patient outcome data was used within CML care [37].

To overcome the lack of stand-alone functionality, HCPs suggested in this study to incorporate the dashboard within existing structures (electronic medical records and PHE). Currently, aligning information systems in health care is a tremendous challenge, and therefore, national strategies have been developed in recent years to improve data infrastructure and interoperability [38]. Because access to the dashboard was a barrier as well, it was preferred that no additional log-in be needed, but GDPR legislation obligates the use of a 2-factor authentication. In addition, the World Health Organization mentions safety and security as essential aspects, and similar levels of security and authority are seen in related eHealth systems in other countries [39-41]. Several national initiatives are developed and tested, which should simplify access of HCPs to eHealth platforms, including patient information such as the CML Dashboard [42-44].

There is a potential central role of the CML Dashboard in the remote monitoring of patients with CML. In this study, the use of the CML Dashboard was low. This could be due to the slow transformation of providing remote care in general, and especially for patients with CML. In addition, experience with the treatment of CML and expertise are scattered [11,12]. This threatens optimal guideline adherence and adequate monitoring of disease activity [11,12]. eHealth can address health care challenges and support patients in managing their health and support HCPs to manage their patients more efficiently. As CML care is largely protocolled, more remote care could be enabled with the right tools. It could also enable the possibility to monitor patients with CML from home when disease activity is low and stable and only visit the hospital when needed, in accordance with the guidelines. Not only do all HCPs get insight into the course of their patients' disease in one glance, as the CML Dashboard highlights those who are in need of attention. When remote care is introduced in CML care, a well-functioning dashboard to monitor patients with CML BCR::ABL1 values, BCR::ABL1 monitoring time points, and side effects of treatment is therefore a must. Impeding rules and regulations and a lack of proper funding were already mentioned as barriers in this study. One strategy to overcome these barriers is agenda setting. In the Europe's Beating Cancer Plan, the use of digital health and telemonitoring to decrease inequality is encouraged [45]. Furthermore, in 2022, a shared view on the future of health care in the Netherlands was published [46]. In this document, it is emphasized multiple times that appropriate care takes place at home when possible, supported by digital tools, which is also called hybrid care delivery. These digital tools should aim to provide people-centered and sustainable care. When a properly functioning dashboard is used to remotely monitor patients with CML, or in a next step to proactively guide them by using alerts, hospital care can be reduced and patients become more self-directed. However, this asks for a couple of requirements, such as organizing care in networks and gaining financial resources for the transformation [46].

Next to new applications, patients' use of the app remains an essential requisite in the implementation of the CML Dashboard. Therefore, barriers and facilitators of the implementation of the CMyLife app experienced by patients should be taken into account. This should result in improved active participation on the patient side, using the CMyLife app accurately. This will equip hematology HCPs with the needed information to balance the type of care needed at a certain moment in time, according to clinical guidelines.

Strengths and Limitations

The strengths of this study are its methodological diversity, theoretical grounding, focus on early-stage evaluation, and inclusion of a diverse participant sample. These strengths collectively contribute to the study's robustness and potential to generate valuable insights into the potential of the CML Dashboard, thereby informing future research and clinical practice. The incorporation of both qualitative and quantitative research methodologies enhances the depth and breadth of data collection and analysis. Qualitative methods allowed for in-depth exploration of perceptions, experiences, and contextual factors, while quantitative approaches facilitated exploration of the available data, providing insight into the usefulness and opportunities of the CML Dashboard in its current form.

XSL•FO
RenderX

Additionally, the use of established theoretical frameworks provided a solid conceptual basis for the qualitative part of the study and interpretation of results. This ensured the study's alignment with existing theoretical perspectives and enhanced its generalizability. Moreover, the inclusion of a diverse set of patients and HCPs enriches the study's findings and increases its external validity. By capturing perspectives and experiences from a range of demographic, clinical, and professional backgrounds, the study can better reflect the complexities and nuances inherent in real-world health care contexts, enhancing the relevance and applicability of its conclusions. However, user experiences from patients should also be determined to improve participation from the patient side.

As for the quantitative research, there were a few limitations to the data. First, all data were registered by patients themselves, and it was not possible to check the reliability. This could have resulted in missing values, inconsistencies, and potential errors in time points and treatment details. For example, a period of 1 day or a period of 7380 days between monitoring time points is very unlikely. Second, patients registered what TKI they used and how often they switched. However, the reason for switching TKIs (lack of response or intolerability) was not registered, which is of importance because milestones of the different treatment lines differ. Another limitation of the quantitative data is the relatively high proportion of missing data for certain variables, such as age and current TKI use, which is inherent to patient-reported, manually entered data. As a result, findings may be less generalizable to the broader CML population, particularly patients who are less digitally engaged. At last, a limitation for the qualitative research was that the interviewed HCPs had little experience with the CML Dashboard in real-world clinical practice. However, their limited experiences were of great value in this early-stage exploration study with the aim to improve the CML Dashboard.

Conclusions

Adjustments should be made to the CML Dashboard to make the data useful for real-world data-driven guideline adjustments and remote monitoring. Organizational steps need to be taken first before CML care can be moved completely toward the home environment. HCPs specified several barriers and facilitators, and potential applications of the CML Dashboard. The CML Dashboard could have a central role in remote monitoring and quality assurance as it could enable patients to act as active participants in their own care, functioning not only as care recipients but also as providers of care and producers of health-related data. This aligns with the concept of patient empowerment and self-management, where patients are equipped with the knowledge, tools, and autonomy to manage their health conditions within specified guidelines. Our dashboard could be the first application that gives HCPs a timely oversight of the achievement/failure of treatment milestones of their patients and enables the use of real-world data for policy adjustments in daily clinical practice.

Acknowledgments

The authors would like to thank all researchers, nurses, and physicians who participated and recruited participants, all patients who participated in this study, the A.T. Twintey project team, and the patient association Hematon. Generative AI was used to make textual adjustments to improve readability. JEMS and LV were co-first authors of this article, and RPMGH and NMAB were co-senior authors.

Funding

This research was partly supported by The Netherlands Organization for Health Research and Development (ZonMw; project numbers: 516006001, 05160062510003). This organization may not gain or lose financially through this publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 138 KB - cancer_v12i1e76096_app1.docx]

References

1. Yabroff KR, Lund J, Kepka D, Mariotto A. Economic burden of cancer in the United States: estimates, projections, and future research. Cancer Epidemiol Biomarkers Prev 2011 Oct 1;20(10):2006-2014. [doi: 10.1158/1055-9965.EPI-11-0650]
2. Katalenich B, Shi L, Liu S, et al. Evaluation of a remote monitoring system for diabetes control. Clin Ther 2015 Jun 1;37(6):1216-1225. [doi: 10.1016/j.clinthera.2015.03.022] [Medline: 25869625]
3. Pelter MN, Quer G, Pandit J. Remote monitoring in cardiovascular diseases. Curr Cardiovasc Risk Rep 2023 Nov;17(11):177-184. [doi: 10.1007/s12170-023-00726-1]
4. Druker BJ, Guilhot F, O'Brien SG, et al. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. N Engl J Med 2006 Dec 7;355(23):2408-2417. [doi: 10.1056/NEJMoa062867] [Medline: 17151364]

XSL•FO
RenderX

5. Bower H, Björkholm M, Dickman PW, Höglund M, Lambert PC, Andersson TL. Life expectancy of patients with chronic myeloid leukemia approaches the life expectancy of the general population. J Clin Oncol 2016 Aug 20;34(24):2851-2857. [doi: 10.1200/JCO.2015.66.2866] [Medline: 27325849]
6. Maas CCHM, van Klaveren D, Ector GICG, et al. The evolution of the loss of life expectancy in patients with chronic myeloid leukaemia: a population-based study in the Netherlands, 1989-2018. Br J Haematol 2022 Mar;196(5):1219-1224. [doi: 10.1111/bjh.17989] [Medline: 34865221]
7. Ossenkoppele G, Janssen JJWM, Posthuma E, et al. Aanbevelingen voor de behande-ling van chronische myeloïde leukemie. Ned Tijdschr Hematol 2011;8(6):237-247 [FREE Full text]
8. Ector G, Geelen IGP, Dinmohamed AG, et al. Adherence to quality indicators in chronic myeloid leukemia care: results from a population-based study in the Netherlands. Leuk Lymphoma 2023 Jan 28;64(2):424-432. [doi: 10.1080/10428194.2022.2142055]
9. Hillis CM, Busque L, Stakiw J, Forrest DL. Adherence to chronic myeloid leukemia monitoring and treatment guidelines in Canadian registries. Blood 2016 Dec 2;128(22):3079-3079. [doi: 10.1182/blood.V128.22.3079.3079] [Medline: 27247135]
10. Goldberg SL, Akard LP, Dugan MJ, Faderl S, Pecora AL. Barriers to physician adherence to evidence-based monitoring guidelines in chronic myelogenous leukemia. J Oncol Pract 2015 May;11(3):e398-e404. [doi: 10.1200/JOP.2014.001099] [Medline: 25758446]
11. Geelen IGP, Thielen N, Janssen J, et al. Impact of hospital experience on the quality of tyrosine kinase inhibitor response monitoring and consequence for chronic myeloid leukemia patient survival. Haematologica 2017 Dec;102(12):e486-e489. [doi: 10.3324/haematol.2017.175265] [Medline: 28860340]
12. Geelen IGP, Thielen N, Janssen J, et al. Treatment outcome in a population-based, “real-world” cohort of patients with chronic myeloid leukemia. Haematologica 2017 Nov;102(11):1842-1849. [doi: 10.3324/haematol.2017.174953] [Medline: 28860339]
13. Jabbour E, Saglio G, Radich J, Kantarjian H. Adherence to BCR-ABL inhibitors: issues for CML therapy. Clin Lymphoma Myeloma Leuk 2012 Aug;12(4):223-229. [doi: 10.1016/j.clml.2012.04.002] [Medline: 22633166]
14. Hall AE, Paul C, Bryant J, et al. To adhere or not to adhere: rates and reasons of medication adherence in hematological cancer patients. Crit Rev Oncol Hematol 2016 Jan;97:247-262. [doi: 10.1016/j.critrevonc.2015.08.025] [Medline: 26412718]
15. Pouls BPH, Vriezekolk JE, Bekker CL, et al. Effect of interactive eHealth interventions on improving medication adherence in adults with long-term medication: systematic review. J Med Internet Res 2021 Jan 8;23(1):e18901. [doi: 10.2196/18901] [Medline: 33416501]
16. Wong ZS, Siy B, Da Silva Lopes K, Georgiou A. Improving patients’ medication adherence and outcomes in nonhospital settings through eHealth: systematic review of randomized controlled trials. J Med Internet Res 2020 Aug 20;22(8):e17015. [doi: 10.2196/17015] [Medline: 32663145]
17. Verweij L, Ector G, Smit Y, et al. Effectiveness of digital care platform CMyLife for patients with chronic myeloid leukemia: results of a patient-preference trial. BMC Health Serv Res 2023 Mar 8;23(1):228. [doi: 10.1186/s12913-023-09153-9] [Medline: 36890512]
18. Innovative care for chronic conditions: building blocks for actions. : World Health Organization; 2002 Oct 31 URL: https://www.who.int/publications/i/item/innovative-care-for-chronic-conditions-building-blocks-for-actions [accessed 2026-04-16]
19. Ector GI, Westerweel PE, Hermens RP, et al. The development of a web-based, patient-centered intervention for patients with chronic myeloid leukemia (CMyLife): design thinking development approach. J Med Internet Res 2020 May 15;22(5):e15895. [doi: 10.2196/15895] [Medline: 32412424]
20. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
21. Chronic myeloid leukemia. Nederlandse Vereniging voor Hematologie. 2023 Dec 6. URL: https://publicatie.hematologienederland.nl/richtlijnen/chronische-myeloide-leukemie-2 [accessed 2026-04-16]
22. CMyLife. URL: www.cmylife.nl [accessed 2026-04-16]
23. cmyCML. URL: https://www.cmycml.nl [accessed 2026-04-16]
24. PGO.nl. URL: www.pgo.nl/partners [accessed 2026-04-16]
25. MedMij. URL: www.medmij.nl [accessed 2026-04-16]
26. Patients Know Best. URL: https://www.patientsknowbest.nl [accessed 2026-04-16]
27. Verweij L, Metsemakers S, Ector G, et al. Improvement, implementation, and evaluation of the CMyLife digital care platform: participatory action research approach. J Med Internet Res 2023 Sep 15;25:e45259. [doi: 10.2196/45259] [Medline: 37713242]
28. Colorafi KJ, Evans B. Qualitative descriptive methods in health science research. HERD Health Environ Res Des J 2016 Jul;9(4):16-25. [doi: 10.1177/1937586715614171] [Medline: 26791375]
29. Doyle L, McCabe C, Keogh B, Brady A, McCann M. An overview of the qualitative descriptive design within nursing research. J Res Nurs 2020 Aug;25(5):443-455. [doi: 10.1177/1744987119880234] [Medline: 34394658]

XSL•FO
RenderX

30. Flottorp SA, Oxman AD, Krause J, et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. Implement Sci 2013 Mar 23;8(1):35. [doi: 10.1186/1748-5908-8-35] [Medline: 23522377]
31. Grol R, Wensing M. What drives change? Barriers to and incentives for achieving evidence-based practice. Med J Aust 2004 Mar 15;180(S6):S57-S60. [doi: 10.5694/j.1326-5377.2004.tb05948.x] [Medline: 15012583]
32. Ono T. Which tyrosine kinase inhibitors should be selected as the first-line treatment for chronic myelogenous leukemia in chronic phase? Cancers (Basel) 2021 Oct 12;13(20):5116. [doi: 10.3390/cancers13205116] [Medline: 34680265]
33. Saglio F. Speaker presentations. Leukemia Res 2017;61(Suppl 1):S1-S14. [doi: 10.1016/S0145-2126(17)30549-0]
34. García-Gutiérrez V, Breccia M, Jabbour E, Mauro M, Cortes JE. A clinician perspective on the treatment of chronic myeloid leukemia in the chronic phase. J Hematol Oncol 2022 Jul 11;15(1):90. [doi: 10.1186/s13045-022-01309-0] [Medline: 35818053]
35. Baccarani M, Bonifazi F, Soverini S, et al. Questions concerning tyrosine kinase-inhibitor therapy and transplants in chronic phase chronic myeloid leukaemia. Leukemia 2022 May;36(5):1227-1236. [doi: 10.1038/s41375-022-01522-3] [Medline: 35338251]
36. Porter ME. What is value in health care? N Engl J Med 2010 Dec 23;363(26):2477-2481. [doi: 10.1056/NEJMp1011024] [Medline: 21142528]
37. Smit Y, Metsemakers S, Janssen J, et al. Measuring chronic myeloid leukaemia TKI-related toxic effects in the real world: a systematic review and critical assessment of content validity of patient-reported outcome measures. Lancet Haematol 2023 Oct;10(10):e849-e859. [doi: 10.1016/S2352-3026(23)00173-4] [Medline: 37604177]
38. Nationale strategie voor het gezondheidsinformatiestelsel. : Ministry of Health, Welfare and Sport, The Netherlands; 2024 Oct 31 URL: https://www.rijksoverheid.nl/documenten/publicaties/2024/10/31/nationale-strategie [accessed 2026-04-16]
39. Mariano B. Towards a global strategy on digital health. Bull World Health Organ 2020 Apr 1;98(4):231-231A. [doi: 10.2471/BLT.20.253955] [Medline: 32284641]
40. Kapoor M. Top real world evidence (RWE) platforms. Mahalo Digital Ventures. 2025 Mar 20. URL: https://www.mahalo.health/insights/top-real-world-evidence-rwe-platforms [accessed 2026-04-16]
41. Liao T, Qiu L, Zhu J, Li J, Zhang Y, Yang L. A mHealth-based nursing model for assessing the health outcomes of the discharged patients with nasopharyngeal carcinoma: a pilot RCT. BMC Nurs 2022 Aug 1;21(1):210. [doi: 10.1186/s12912-022-00993-0] [Medline: 35915490]
42. UZI-register. Ministry of Health, Welfare and Sport, The Netherlands. URL: https://www.uziregister.nl [accessed 2026-04-16]
43. Digital transformation. NHS England. URL: https://www.england.nhs.uk/digitaltechnology [accessed 2026-04-06]
44. Balla D, Belani H, Berdun J, et al. Building sustainable digital health services in Europe: lessons learned from the COVID-19 pandemic. : CroRIS; 2020 URL: https://www.croris.hr/crosbi/publikacija/rad-ostalo/789805 [accessed 2026-04-30]
45. A cancer plan for Europe. European Commission. 2025 May 27. URL: https://commission.europa.eu/topics/public-health/european-health-union/cancer-plan-europe_en [accessed 2026-04-16]
46. Integraal Zorgakkoord: ‘Samen werken aan gezonde zorg’ [in Dutch]. : Ministry of Health, Welfare and Sport, The Netherlands; 2022 Sep 16 URL: https://www.rijksoverheid.nl/documenten/rapporten/2022/09/16/integraal-zorgakkoord-samen-werken-aan-gezonde-zorg [accessed 2026-04-16]

Abbreviations

CML: chronic myeloid leukemia
COREQ: Consolidated Criteria for Reporting Qualitative Research
GDPR: General Data Protection Regulation
HCP: health care professional
IS: international scale
PHE: personal health environment
RESTful: Representational State Transfer
TKI: tyrosine kinase inhibitor

Edited by M Balcarras; submitted 29.Apr.2025; peer-reviewed by O Nguyen, Y Chu; revised version received 17.Mar.2026; accepted 18.Mar.2026; published 07.May.2026.

Please cite as:
Verweij L, Swillens JEM, Metsemakers SJJPM, Smit Y, Wener E, Ector GICG, Hermens RPMG, Blijlevens NMA
Feasibility and Implementation of an eHealth Dashboard for the Remote Monitoring of Dutch Patients With Chronic Myeloid Leukemia: Multimethods Approach
JMIR Cancer 2026;12:e76096
URL: https://cancer.jmir.org/2026/1/e76096
doi:10.2196/76096

XSL•FO
RenderX

© Lynn Verweij, Julie E M Swillens, Sanne J J P M Metsemakers, Yolba Smit, Esri Wener, Geneviève I C G Ector, Rosella P M G Hermens, Nicole M A Blijlevens. Originally published in JMIR Cancer (https://cancer.jmir.org), 7.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cultural Adaptation of a Web-Based Ostomy Care Intervention for Hispanic Patients With Cancer and Caregivers: Mixed Methods Study

Xiaomeng Wang[1], MS; Lixin Song[2], PhD; Amy Vondenberger[1], BS; Fei Yu[3], PhD; Jia Liu[1], MSN, PhD; Roxana Delgado[1], PhD; Robert Svatek[4,5], MD; Natalia Villegas Rodriguez[6], MSN, PhD; Lurheinna Rosado Rivera[1], BSN; Mark Bonnen[5,7], MD

[1]School of Nursing, The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

[2]School of Nursing, Mays Cancer Center, The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Dr, San Antonio, TX, United States

[3]School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[4]Department of Urology, Long School of Medicine, The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

[5]Mays Cancer Center, The University of Texas Health Science Center, San Antonio, San Antonio, TX, United States

[6]School of Nursing, University of North Carolina at Chapel Hill, Chapel Hill, United States

[7]Department of Radiation Oncology, Long School of Medicine, The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

Corresponding Author:
Lixin Song, PhD
School of Nursing, Mays Cancer Center, The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Dr, San Antonio, TX, United States

Abstract

Background: Ostomy creation for cancer treatment negatively impacts the quality of life of both patients and caregivers. Hispanic patients with cancer and caregivers often face additional challenges, including limited access to supportive care programs.

Objective: This study aimed to examine the experiences and preferences of Hispanic patients with cancer living with ostomies and their caregivers to inform the cultural adaptation of an existing intervention program and the design of Ostomy Self-Care Program (Programa de AutoCuidado de Estoma [PACE]).

Methods: In this 2-stage study, conducted between March and August 2023 in San Antonio, Texas, we used a qualitatively driven mixed methods design, starting with an initial survey followed by qualitative interviews to explore the experiences, needs, and intervention preferences of Hispanic patients and caregivers managing ostomy care. We used Braun and Clarke's 6-phase thematic analysis approach to analyze the qualitative data and performed descriptive analysis for the quantitative data. Subsequently, we applied affinity diagramming and persuasive systems design principles to guide the design of PACE.

Results: In total, 14 Hispanic participants managing an ostomy (9 patients with cancer and 5 caregivers) completed a survey and participated in interviews, continuing until data saturation was reached. Participants had a mean age of 58.9 (SD 13.01, range 37‐79) years, and most (n=12) reported a high school diploma or General Educational Development as their highest education level. Around 5 (36%) participants scored below 26 on the eHealth Literacy Scale (eHEALS), indicating low digital health literacy, and the average Charlson Comorbidity Index (CCI) was 3.21 (SD 1.86, range 0‐6). Overall, 3 major themes emerged from the qualitative data analysis, namely perceptions of living with an ostomy, seeking support, and postsurgery challenges. Additionally, two primary themes emerged from participant interviews: (1) importance of preferred language and multimedia delivery and (2) patients and caregivers desire early introduction, multimodal delivery of materials, and inclusion of peer and family support. These themes informed the design and development of a culturally appropriate, web-based, bilingual PACE intervention that integrates content visualization, cultural adaptations, and persuasive technologies—strategies designed to encourage user engagement.

Conclusions: Our findings emphasize the importance of understanding the ostomy care experiences, supportive care needs, and intervention preferences of Hispanic patients and caregivers. Informed by stakeholders' insights, we culturally adapted the original intervention program using persuasive systems design principles to design and develop the PACE intervention, aiming to enhance engagement among Hispanic patients with cancer and caregivers, support effective self-management of ostomy care, and improve health outcomes.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e70354) doi:10.2196/70354

KEYWORDS

ostomy; Hispanic; eHealth; digital platform; caregiver; patient with cancer; cultural adaptation; persuasive technology.

Introduction

Ostomy creation can be performed as part of cancer treatment for bladder, colorectal, cervical, ovarian, and uterine cancers. The surgery involves creating an opening on the anterior abdominal wall to allow defecation or urination [1]. This procedure can complicate the patient's ability to resume normal daily activities and further impact quality of life (QOL) [2]. While supportive care and management can sustain QOL after ostomy creation, a gap remains in available care resources that are tailored to minorities, including Hispanic patients with cancer and caregivers.

Compared to their non-Hispanic counterparts, Hispanic patients with cancer and their caregivers in the United States face additional challenges throughout cancer survivorship. These barriers include socioeconomic difficulties [3], limited access to health care [4], language barriers, and financial strain [5]. Additionally, this population experiences a higher rate of being uninsured [6]. Approximately 51.3% of Hispanics in the United States need translation services during medical consultations [7]. These language barriers can hinder effective communication and potentially lead to misunderstanding home care instructions [8]. Patients undergoing an ostomy creation surgery typically use more health care resources [9], which poses a further detriment to the Hispanic population, as they often face significant challenges in obtaining health care resources[10]. In addition, cultural beliefs and practices play a crucial role in illness management. In Hispanic culture, families traditionally take on the responsibility of caring for their sick members [11], and they tend to be more involved in ostomy care than any other medical or nursing tasks [12]. A study shows that involving family members in group training sessions can significantly improve outcomes among Hispanic patients and families[13]. Recognizing the needs of both Hispanic patients and their families is essential for providing adequate support and intervention in postostomy care.

Web-based interventions offer a promising way to support self-management for patients with chronic conditions and their families [14]. Research suggests that, compared to non–web-based interventions, these web-based interventions are effective in promoting knowledge acquisition and behavioral change [15]. As a cost-effective solution, they can particularly benefit underserved groups, such as Hispanic communities [16], through overcoming barriers to health care access. Offering content in Spanish further addresses language barriers and increases access [17]. Additionally, it provides a convenient way for patients to share information within their community, fostering peer support among people who have similar experiences [18].

As part of our preparatory work to address the care needs of Hispanic patients with cancer and caregivers, our multidisciplinary team previously translated the PRISMS (Patient-Reported Outcomes-Informed Symptom Management System) [19] into Spanish. PRISMS is an eHealth intervention that integrates a program with wearable devices to support patients with cancer and their caregivers during the postostomy care transition from hospital to at-home self-management. PRISMS aims to reduce preventable emergency room use and readmissions [19]. To ensure the cultural appropriateness of the translated materials, we engaged 10 bilingual translators, including clinical nurses, nursing students, and faculty, from Chile, Mexico, Puerto Rico, Argentina, and Colombia. These bilingual, native Spanish speakers translated the English version of PRISMS into Spanish to accommodate various dialects while sensitively addressing cultural nuances.

Building on previous work, our current study explored the experiences of Hispanic patients with cancer and caregivers regarding ostomy care, as well as their preferences for ostomy care interventions, to inform cultural adaptation of PRISMS-Spanish using a mixed methods approach. We presented the PRISMS-Spanish materials to participants to facilitate discussion and gather their feedback to help inform our design of a web-based ostomy care intervention program—Programa de AutoCuidado de Estoma (PACE)—tailored for this specific underserved population. While some references cited in this paper use terms, such as "Latino/Latina," we use "Hispanic" consistently throughout this study to align with the terminology used in participant recruitment.

Methods

Study Design Overview

In stage 1, we used a qualitatively driven mixed methods design with sequential data collection and parallel analysis. First, we made an initial phone call to each participant and collected quantitative data. Then we scheduled a Zoom (Zoom Communications) interview with each of them to collect qualitative data, which was typically within 1 month after the initial phone contact. The survey data provided a clear understanding of the characteristics, health literacy, and long-term mortality risk of Hispanic patients with cancer and caregivers. In contrast, the interview data and thematic analysis offered in-depth insights into their ostomy care experiences and preferences for intervention features and functions. We used Braun and Clarke's 6-phase thematic analysis approach [20] to guide our data analysis.

In stage 2, we used affinity diagramming and the Persuasive Systems Design (PSD) model developed by Oinas-Kukkonen and Harjumaa [21] to guide PACE program design for effective user engagement [22,23].

XSL•FO
RenderX

Stage 1: Exploration of Ostomy Care Experiences and Intervention Preference

We aimed to understand the experiences and information needs of Hispanic patients and caregivers. Additionally, we gathered intervention preferences based on participant feedback on PRISMS-Spanish materials.

Participants

Participants were recruited from the Mays Cancer Center in San Antonio, Texas, the seventh largest city in the United States, with a 65% Hispanic majority population [24]. Patients were eligible if they (1) were Hispanic adults with one of three types of ostomies (colostomy, ileostomy, or urostomy) for the treatment of bladder, colorectal, cervical, ovarian, or uterine cancer; (2) could read and speak Spanish; and (3) passed a cognitive assessment using the Mini-Mental State Examination [25]. Caregivers were included if they were adults serving as the patient's primary caregiver and could read and speak Spanish. Eligible participants were contacted via phone to confirm their willingness to participate and to address any questions.

Data Collection

Quantitative and qualitative data were collected through a survey administered during an initial phone call, followed by a Zoom interview that began with a presentation of the PRISMS-Spanish materials. Each participant encounter was conducted in either English or Spanish, based on their language preference. The survey questionnaires, interview guides, and intervention materials were translated into Spanish by a certified translator and reviewed by bilingual research staff and nursing students.

The trained bilingual research staff and research assistant (AV and LRR) conducted the initial phone survey before the interview in a secure, private office space at the University of Texas Health Science Center at San Antonio (UTHSA). They read each survey question aloud and recorded participants' responses into REDCap (Research Electronic Data Capture; Vanderbilt University) hosted by the UTHSA. The survey items included demographic questions, type of ostomy, cancer diagnosis, Charlson Comorbidity Index (CCI) [26], and the Electronic Health Literacy Scale (eHEALS) [27]. Participants were asked to complete the CCI and the eHEALS questionnaires to measure comorbidities and digital health literacy, which were critical for chronic illness management and for informing the PACE design. The CCI score is the summation of the number of comorbidities that participants reported. The eHEALS score, ranging from 8 to 40, assessed participants' digital literacy, including knowledge, comfort level, and perceived abilities in using electronic health information, with scores below 26 indicating low health literacy.

Next, participants were scheduled for a Zoom interview session at their convenience. During the session, translated PRISMS-Spanish materials were first presented, either via screen share or by email if needed, followed by open-ended questions. The semistructured interview guide was collaboratively developed by the research team, informed by insights from the PRISMS study [19] and a review of existing interventions related to ostomy care, digital health literacy, and culturally tailored programs for Hispanic populations. Interview questions focused on four main areas: (1) experience of ostomy care, (2) feedback on PRISMS-Spanish materials, (3) preferences for intervention and delivery, and (4) preferred implementation strategies. Interview sessions lasted approximately 45 - 60 minutes and were conducted between March and August 2023 by the research staff. All interviews were audio-recorded and transcribed verbatim via Zoom and reviewed by the project coordinator and trained research staff who had prior experience working with Spanish-speaking patients. We determined that data saturation was achieved when no new codes or themes emerged across interviews. This approach aligns with prior research indicating that qualitative data saturation often occurs within the first 12 interviews, with key themes emerging as early as 6 [28].

Data Analysis

Quantitative data from the initial phone surveys were exported from REDCap into Microsoft Excel and RStudio (Posit; version 4.4.1) for analysis. We performed descriptive statistics, including means, ranges, frequencies, SDs, and percentages, to summarize participant demographics, CCI scores, and eHEALS (Table 1). These analyses provided an overview of the sample's characteristics and informed the interpretation of qualitative findings.

XSL•FO
RenderX

Table . Participant characteristics.

Variable	Patients (n=9), n	Caregivers (n=5), n	Both (n=14), n
Sex			
Male	4	2	6
Female	5	3	8
Race			
White	6	4	10
Other	3	1	4
Education			
<=7th-8th Grade	2	0	2
12th Grade	6	4	10
Bachelor's degree	1	1	2
Employment status			
Work	3	2	5
Retired, age/choice	2	3	5
Disability/illness	3	0	3
Unable to work/perform normal activities	1	0	1
Marital status			
Single/never married	1	0	1
Married/domestic partner	8	4	12
Partner			
Long-term partner	0	1	1
Ostomy type			
Urostomy	2	3	5
Ileostomy	3	0	3
Colostomy	4	2	6
Age (years)			
30 - 39	2	0	2
40 - 49	0	1	1
50 - 59	3	1	4
60 - 69	2	2	4
>70	2	1	3
CCI[a]			
0	1	0	1
1-2	3	1	4
3-4	3	3	6
≥5	2	1	3
eHEALS score[b]			
<26	4	1	5
>26	5	4	9

[a]CCI: Charlson Comorbidity Index.

[b]eHEALS: eHealth Literacy Scale.

XSL•FO
RenderX

The deidentified qualitative interview data were analyzed using Braun and Clarke's 6-phase approach to thematic analysis [20]. This approach was applied to ensure a credible analysis procedure and continued until no new themes were identified. The approach is detailed as follows:

(1) Familiarization: Two researchers (AV and LRR) independently read the interview transcripts and noted initial ideas and patterns. This process allowed researchers to gain an in-depth understanding of the participants' overall experiences and needs.

(2) Generating initial codes: AV and LRR generated initial codes systematically across the entire dataset, using NVivo (Lumivero) qualitative data analysis software (version 10.2.1) and a preliminary codebook based on interview topics.

(3) Searching for themes: the research team collated the initial codes into broader themes by grouping related codes to generate potential themes to capture participants' experiences.

(4) Reviewing themes: Team meetings were held to review themes and resolve discrepancies in theme identification, aiming for a minimum coding consistency agreement of 80% [29].

(5) Defining and naming themes: 3 researchers (AV, LRR, and XW) collaboratively defined and named each theme, ensuring that each theme was distinct and related to the aim of this study. Detailed descriptions were developed for each theme, supported by illustrative quotes from participants.

(6) Producing the report: Finally, the team synthesized the findings into a cohesive report, which was used to inform the cultural adaptation of PRISMS-Spanish and the design of PACE.

Although quantitative and qualitative data were collected sequentially, analysis was conducted in parallel. Quantitative descriptive statistics (eg, demographics, CCI, and eHEALS) were reviewed with qualitative codes and themes side by side. The integration occurred during the interpretation phase, where quantitative results were used to contextualize and support qualitative findings. The integrated results were synthesized into joint displays (Tables 2 and 3) to inform the design of the PACE intervention.

XSL•FO
RenderX

Table . Hispanic patients' and caregivers' experiences with ostomy care.

Themes and subthemes	Participant quotes
Ostomy perception	
Emotion	• "Because it is kind of a shocking thing to see your internal organs hanging out of your body even though it's a little bit, it's still a hole in your belly and you got your guts or hanging out." [Female, ileostomy, patient, eHEALS[a] score=35, CCI[b]=2] • "All the ugly stuff that nobody wants to show." [Female, ileostomy, patient, eHEALS score=32, CCI=3]
Necessity	• "And as well as many other people, it definitely has changed my perception to something that is scary to something that is a necessity at times if needed." [Female, colostomy, patient, eHEALS score=38, CCI=4] • "Knowing that I had it and I needed it to save my life." [Female, colostomy, patient, eHEALS score=38, CCI=4]
Support	
Health care provider	• "But I don't feel that knowledge is there for a lot of health care workers." [Female, colostomy, patient, eHEALS score=38, CCI=4]
Family and caregiver	• "So I really depended on him (husband) to help me put a new one back on." [Female, colostomy, patient, eHEALS score=38, CCI=4] • "… but what really helped me a lot was her (wife) and my family that was a hundred percent behind me and you know they, gave me the faith because I didn't have the faith at first." [Male, colostomy, patient, eHEALS score=23, CCI=6] • "I have an alternative. I do my husband's changes for his urostomy bag. But my son also knows how to provide the care because his wife has an ostomy bag. So in a pinch if I'm not available, he can substitute for me." [Female, urostomy, caregiver, eHEALS score=32, CCI=4] • "So they supported me through the cancer, but they weren't really sure how to help me or how to feel about it because it deals with stool and . I know they were really like kind of just supporting but not really familiar with it." [Female, colostomy, patient, eHEALS score=38, CCI=4] • "I think so, that they also involve families in these talks for them to also see how family members and the sick should be treated" [Male, colostomy, caregiver, eHEALS score=29, CCI=3]
Peer support	• "I had a lot more difficulties with my ostomy then a lot of people but you know, I ended up going to peers that worked with an ostomy as opposed to the ostomy care nurses that they were giving me because they too still weren't familiar with how to take care of my problem." [Female, colostomy, patient, eHEALS score=38, CCI=4]
Self-care	• "… it depends on what the circumstances are, but I think that the patient should do as much as possible to help themselves." [Female, ileostomy, patient, eHEALS score=35, CCI=2]
Challenges	
Radiation side effects	• "I couldn't even lift up my arms after radiation ... if they have the ostomy bag for some reason, if they're also undergoing radiation. Yeah, that's bad." [Female ileostomy, patient, eHEALS score=35, CCI=2]
Chemotherapy side effects	• "here's added things so sometimes chemos depending on which one some can cause a lot of diarrheas. Some can cause constipation." [Female, colostomy, patient, eHEALS score=38, CCI=4]

XSL•FO
RenderX

Themes and subthemes	Participant quotes
Skin irritation	• "You know, a lot of times people see, okay, it's just, it's just a little rash or whatever But actually, it's maybe an infection or something else and they see it as it just being a little rash when it's really not, you know." [Male, ileostomy, patient, eHEALS score=34, CCI=0] • "The skin irritation just lasted the entire time." [Female, ileostomy, patient, eHEALS score=35, CCI=2] • "Once in a while, he'll have a little bit of irritation, but I'll just clean it." [Female, urostomy, caregiver, eHEALS score=32, CCI=4]
Leakage	• "One reason was because, you know, I was having like a leak. And we're like, well maybe, what are we doing wrong?" [Male, ileostomy, patient, eHEALS score=34, CCI=0] • "Because I'm a heavier-set girl and obesity is rampant in the Latin community. My ostomy was retracted in so it was not sticking out like normal people.it had a lot more problems with leaking because it couldn't get a good seal around the stoma." [Female, colostomy, patient, eHEALS score=38, CCI=4]
Self-isolation	• "I feel in the house we had everything organized and everything there for us to do the job, but taking it out of the houses where it was a little harder for me." [Female, colostomy, patient, eHEALS score=38, CCI=4]
Caregiver stress and emotion	• "I'm gonna say like a depression myself, but I kept it to myself." [Female, colostomy, caregiver, eHEALS score=19, CCI=1] • "We didn't know what else to do. but you know at that point my dad had lost a lot of blood." [Male, urostomy, eHEALS score=34, CCI=4]

[a]CCI: Charlson Comorbidity Index.

[b]eHEALS: eHealth Literacy Scale.

XSL•FO
RenderX

Table . Preferences for the ostomy care intervention among Hispanic patients and caregivers.

Themes and subthemes	Participant quotes
Importance of preferred language and multimedia delivery	
Content and delivery format	• "Okay, the information is really, really helpful." [Female urostomy, caregiver, eHEALS[a] score=32, CCI[b]=4] • "The YouTube (videos) really does help out a lot." [Male, ileostomy, patient, eHEALS score=34, CCI=0] • "You have some really good tips like though there was one that we were told about and it really had helped me now that I when I apply the bag on my husband." [Female, urostomy, caregiver, eHEALS score=32, CCI=4]
Language	• "… for one who does not speak English. There is nothing, there is nothing clearer than Spanish, so the explanation is excellent, it is very good." [Female, colostomy, patient, eHEALS score=25, CCI=1]
Need for early, multimodal materials and peer and family support	
Delivery platform	• "I do on a phone and do a desktop." [Female, urostomy, patient, eHEALS score=40, CCI=6] • "For me it was more telephone because that's what I used more as far as online and researching stuff." [Female, ileostomy, patient, eHEALS score=32, CCI=3] • "I happen to like computers, … but I think not a lot of people, especially if they're older ... So he's not computer savvy either. So I think maybe like a handout or in a notebook kind of form would be better. With pictures." [Female, urostomy, patient, eHEALS=40, CCI=6]
Delivery timing	• "This should be shown to the patients before they leave the hospital because they might not have the support once they get home." [Female, urostomy, caregiver, eHEALS score=32, CCI=4] • "Like before you get going home." [Female, urostomy, caregiver, eHEALS score=32, CCI=4] • "I think it'd be better if you were to be able to give them all the information ahead of time." [Male, ileostomy, patient, eHEALS score=34, CCI=0]
Family and peer support	• "We definitely needed help. So I definitely think helping caregivers or showing caregivers how to help is really important." [Female, colostomy, Patient, eHEALS score=38, CCI=4] • "I think it's very important that anybody in the family that's willing to be involved that they would get involved in case the person that normally does it for you or helps you with it is not available for whatever reason. I think other family members should be involved also." [Male, urostomy, patient, eHEALS score=33, CCI=3] • "I think it would be wise and I think it really helps if you ask somebody, somebody personal, somebody that's gone through it." [Male, ileostomy, patient, eHEALS score=34, CCI=0]

[a]CCI: Charlson Comorbidity Index.

[b]eHEALS: eHealth Literacy Scale.

Stage 2: PACE Web-Based Intervention Program Design

Overview

Findings from stage 1 informed the cultural adaptation of PRISMS-Spanish and the design of the PACE prototype. We used affinity diagramming to organize the intervention contents and materials, and the PSD model to guide the design of the PACE program. Our multidisciplinary team, including digital health researchers, nursing faculty, user interface (UI) designers, web developers, informatics specialists, and clinicians from UTHSA and the University of North Carolina at Chapel Hill, held weekly hybrid meetings (in-person and via Zoom) over 6 months (from September 2023 to March 2024). Design decisions were made collaboratively during structured sessions focused on usability, cultural relevance, and visual accessibility, during which wireframes and prototypes were reviewed. Ideas and comments were synthesized and iteratively incorporated into the prototypes using affinity diagramming on a physical whiteboard, guided by the PSD model. A project coordinator documented meeting notes and key design decisions throughout the process.

XSL•FO
RenderX

Affinity Diagramming

We used affinity diagramming [30] to organize ostomy care materials across various web pages based on the content interrelationships to delineate the tasks and features of the new PACE program. Affinity diagramming is a commonly used user experience design approach that helps organize design ideas systematically [30]. Based on PRISMS-Spanish, the results from stage 1, and the affinity diagram, we generated 5 different types of prototypes for PACE by collaborating with a UI design class. Our multidisciplinary team analyzed and evaluated each prototype. Key features, such as website layout, visual and graphical representations, and navigation options were finalized to ensure the program would meet the needs of patients and caregivers.

Persuasive Systems Design

To ensure the appropriate application of these persuasive principles, we incorporated expert evaluation during the design process. We used PSD to guide the design of the culturally adapted PACE to optimize user engagement in the web-based intervention [31]. The PSD model includes 28 design principles across 4 categories (Textbox 1). Experts in health care, UI design, and health informatics provided feedback on the integration between the PSD model and PACE.

Textbox 1. Persuasive system design model.

Categories of persuasive techniques and system design principles

- Primary task support: reduction, tunneling, tailoring, personalization, self-monitoring, simulation, and rehearsal.
- Dialogue support: praise, rewards, reminders, suggestion, similarity, liking, and social role.
- Credibility support: trustworthiness, expertise, surface credibility, real-world feel, authority, third-party endorsements, and verifiability.
- Social support: social learning, social comparison, normative influence, social facilitation, cooperation, competition, and recognition.

Ethical Considerations

This study was approved by the University of Texas Health San Antonio's Institutional Review Board (IRB; number 20220693EX). Referred patients were contacted via phone to confirm their eligibility and willingness to participate and to answer their questions. All participants received study information and signed a written informed consent form in their preferred language (English or Spanish) after their questions were answered and before any research activities began. The consent form included the following information: (1) the primary data collection (surveys and interviews) would be used to inform intervention design; (2) interview recordings and transcripts would be deidentified and stored separately from study documents containing participants' identifiable information on a password-protected, encrypted cloud drive to ensure data security and confidentiality; and (3) each participant would receive a US $80 gift card upon completing both the survey and the interview session.

Results

Stage 1: Exploration of Ostomy Care Experiences and Intervention Preference

Participant Characteristics

Data collection and analysis continued until no new themes related to care needs and ostomy experiences emerged during the interviews. The research team successfully enrolled 9 patients and 5 caregivers, reaching data saturation with a total of 14 participants. These participants had a mean age of 58.9 (SD 13.01, range: 37 - 79) years. The mean CCI score was 3.21 (SD 1.86, range 0 - 6), indicating a moderate comorbidity burden. The mean eHEALS score was 29.71 (SD 6.62, range 18 - 40). Most participants (n=10) reported a high school diploma or General Educational Development equivalent as their highest level of education, while 2 participants had completed up to the seventh grade, and 2 participants held a bachelor's degree. The patients and caregivers reported an annual income between US $10,000 and US $90,000. Most patients had their ostomy surgery 3 - 4 years before the interview.

Thematic Analysis Results

Hispanic Patients and Caregivers' Experiences With Ostomy Care

As outlined in Table 2, the thematic analysis identified three major themes and associated subthemes among Hispanic patients with cancer and caregivers: (1) perceptions of ostomy and its impact on their lives, (2) experiences in seeking support, and (3) challenges faced following discharge from the hospital.

Perceptions of Ostomy and Its Impacts

Patients and caregivers frequently expressed fear and negative perceptions regarding ostomies and used terms like "ugly" and "shocking" to describe them (refer to "Emotion" in Table 2). On the other hand, the subtheme of necessity reflected a gradual shift in ostomy perception as patients and caregivers began to acknowledge the life-saving function of an ostomy and were able to accept and live with it (refer to "Necessity" in Table 2).

Support Seeking

Support emerged as a significant theme from the interviews, with 4 sources identified, such as health care provider support, family and caregiver support, peer support, and self-care. Many participants reported limited access to resources with sufficient knowledge for changing ostomy bags and teaching others to perform the task (refer to "Health care provider" in Table 2). As shown in the "Family and caregiver" section in Table 2, patients with CCI scores ranging from 4 to 6 emphasized the crucial role of family support. Some indicated that family members often act as backups for one another when providing care. While some patients do not formally recognize their family members as caregivers, they still rely on them for ostomy bag

XSL•FO
RenderX

changes. Additionally, patients expressed a desire to get help from people who have had similar experiences (ie, "Peer support" in Table 2. Conversely, patients with relatively low CCI scores highlighted the importance of self-reliance during postsurgical home management (refer to "Self-care" in Table 2).

Challenges Post Surgery

Patients faced significant physical and emotional difficulties after ostomy surgery. Many needed to continue cancer treatment, such as radiation and chemotherapy, which further complicated their daily ostomy care (refer to "Radiation side effects" and "Chemotherapy side effects" in Table 2). Side effects of these treatments, such as fatigue and diarrhea, made it harder for patients to manage bag changes, thereby increasing the risk of leaks. Additionally, treatments made patients' skin more sensitive and prone to irritation (refer to "Skin irritation" in Table 2). As shown in the "Leakage" section in Table 2, leakage challenges were associated with obesity, which is a common issue in the Hispanic community. Leakage not only caused feelings of helplessness but also led to self-isolation. Many patients reported feeling more secure and comfortable staying at home (refer to "Self-isolation" in Table 2). Notably, caregivers reported high levels of stress and depression while caring for patients with ostomy cancer (refer to "Caregiver stress and emotion" in Table 2).

Preference for the PACE Ostomy Care Intervention Program

"insert subheading"

Overall, participants felt excited about the existence of a program like PRISMS and expressed satisfaction with the translated PRISMS-Spanish materials provided during the interviews. They also shared their preferences regarding intervention delivery timing, devices commonly used in information searching (mobile phones, desktops, etc), and supporting resources.

Importance of Preferred Language and Multimedia Delivery

Participants reported a preference for multimedia, including text documents, audio, and videos, when presented with intervention materials. They found practical tips and tricks extremely useful, especially regarding the optimal timing for changing ostomy appliances, barrier rings, and support belts. Although some tips had been mentioned by their health care team before, participants appreciated the tips as valuable reminders (refer to "Content and delivery format" in Table 3). Participants with low eHEALS scores (<26) were more likely to be impacted by language barriers when accessing health care resources. This is supported by the "Language" (refer to "Language" in Table 3) subtheme, where participants expressed a preference for Spanish-translated materials.

Need for Early Introduction, Multimodal Materials Delivery and Peer and Family Support

Participants expressed their preference for the intervention delivery platform, including the timing and support they would like to receive. Many participants were proficient with computers and mobile devices, but some suggested that providing printed handouts with pictures would also be helpful (refer to "Delivery platform" in Table 3). Participants also stated that information about ostomy care should be provided at the preoperative appointment or, at the latest, during hospitalization, although most patients were sent home with discharge paperwork after having a nurse demonstrate a bag change (refer to "Delivery timing" in Table 3). Feedback also emphasized the importance of training multiple support people in changing an ostomy bag, depending on family dynamics (refer to "Family and peer support" in Table 3). These comments further illustrated that the program should be accessible to all relevant family members.

Stage 2 Design of the Web-Based PACE Intervention

Based on the feedback in stage 1, we developed the culturally adapted PRISMS-Spanish intervention, renamed as PACE. In this section, we focus on introducing the design of the PACE model, guided by affinity diagramming and user engagement strategies informed by the PSD model.

Content and Cultural Adaptation

Based on the interview findings and thematic analysis, we primarily focused on culturally adapting and customizing the existing PRISMS/PRISMS-Spanish program to develop the PACE intervention for Hispanic patients with cancer and their family members. PACE retained core contents from the PRISMS, including text and demonstrative videos on ostomy care, providing information on managing common complications (eg, dehydration and skin issues), performing safe physical activity, and preventing falls. Additionally, we introduced a new preoperation checklist to help patients and their families prepare thoroughly. In response to participants' preferences, we maintained the PRISMS-Spanish for information delivery using multimedia formats, including audio, video, and text. For instance, a step-by-step video demonstration of the crusting technique and a text listing situations in which patients should contact a doctor post surgery were included in PACE. To further enhance user experience, audio, and video were adapted to accommodate preferences for speaking speeds, tone, and captioning.

New culturally adaptive elements were added to engage Hispanic users by incorporating familiar visuals and culturally resonant content. For example, images of traditional Hispanic dishes, such as empanadas and ceviche, were included to promote healthy eating. Culturally significant colors like orange and blue were used to evoke a sense of energy and warmth. These cultural adaptations aimed to develop an inclusive platform for PACE that fosters connection and engagement within the Hispanic community, optimizing support for patients with cancer and caregivers.

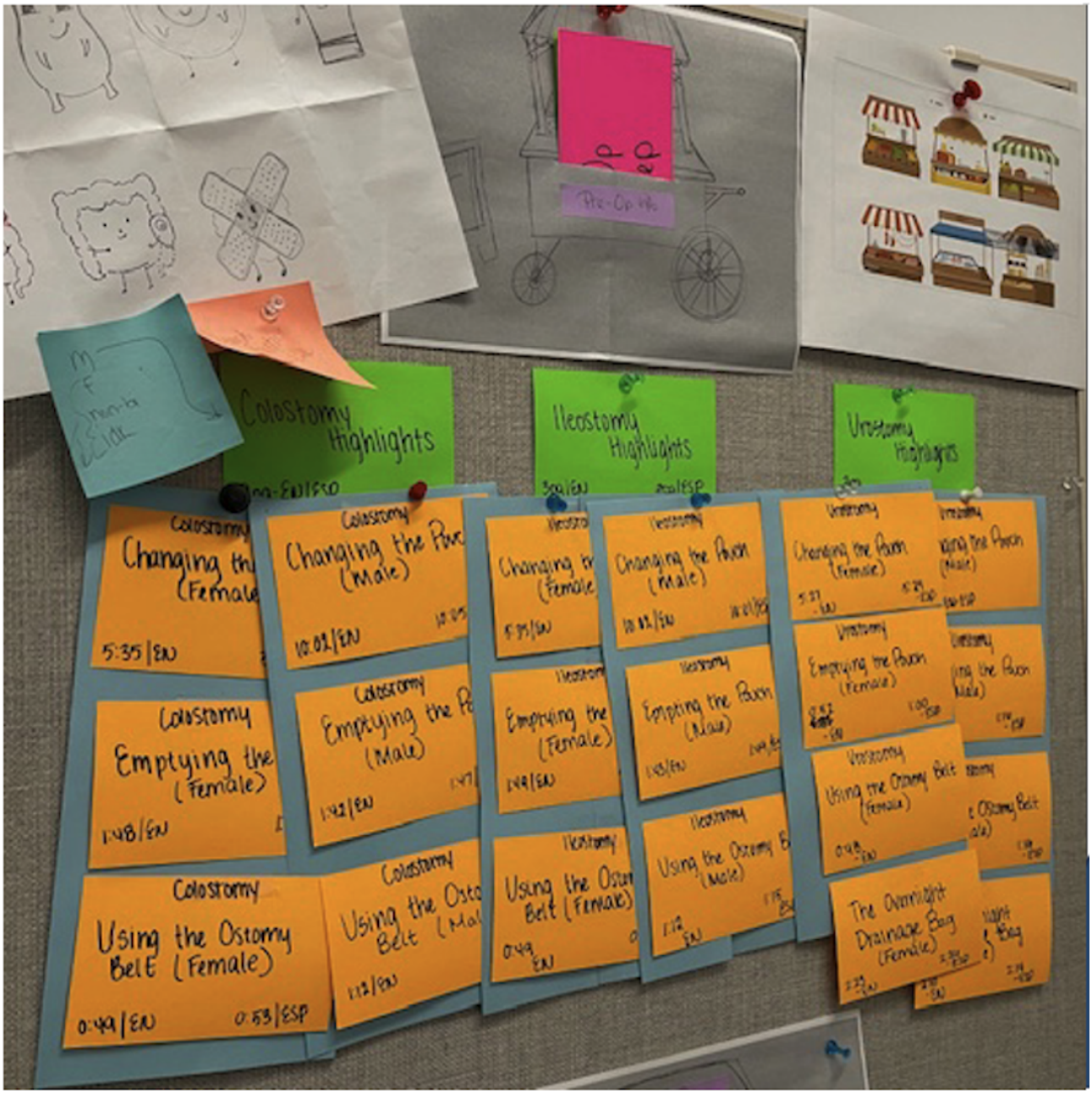
Affinity Diagram of the PACE Website

An affinity diagram was created to organize the content in a logical and user-friendly layout (Figure 1). Each piece of material, whether text, videos, or audio files, was represented as a color-coded data point and written on sticky notes, which were mapped into meaningful sections informed by interview feedback on delivery timing. This process aided in designing a

XSL•FO
RenderX

user experience that guides users through the entire journey of ostomy care, from preoperative preparation to postsurgery self-management at home, with specific adaptations for the Hispanic population.

Figure 1. Affinity diagram of the PACE (Programa de AutoCuidado de Estoma) website.



Similar to our PRISMS intervention, our team identified 6 main sections for the PACE website: “Pre-Op,” “Doctor Office Visits,” “Learning Your Ostomy Device,” “Ostomy Care & Recovery,” “General Physical Health,” and “Good Mental Health” (Figure 2). “Pre-Op” offers comprehensive preoperative preparation, including a printable medication list template. “Doctor Office Visits” provides checklists for health care appointments, listing symptoms and test results to track. “Learning Your Ostomy Device” delivers essential knowledge on managing an ostomy, including changing, emptying, and understanding different types of bags and related equipment. “Ostomy Care & Recovery” addresses skin irritations, chemotherapy side effects, and common complications. “General Physical Health” offers recommendations for physical activities and tips on nutrition and dressing for fitness. Finally, “Good Mental Health” provides strategies for maintaining mental well-being for both patients and caregivers.

XSL•FO
RenderX

Figure 2. Home page of the PACE (Programa de AutoCuidado de Estoma) website.



Persuasive Elements

Overview

After analyzing the interview data and designing the website layout, the research team (LS, AV, JL, LRR, and XW) conducted an expert evaluation leading to the integration of 7 persuasive principles of the PSD Model into the website design, as detailed in Table 4.

Table . Persuasive elements in PACE (Programa de AutoCuidado de Estoma).

PSD[a] categories and selected principles for PACE[b] design	Application of PSD principles
Primary task support	
Tailoring	Users receive tailored task recommendations based on their group (Patient/Caregiver and family member).
Self-monitoring	Users can track their progress on each task.
Dialogue support	
Rewards	Users earn avatars as rewards for progress in learning sections.
Reminders	Users receive email reminders to log in to the PACE website regularly.
Credibility support	
Trustworthiness	The PACE website offers truthful, fair, and unbiased information through various media, including videos, audio, text, and images.
Real-world feel	Team members from the research and development teams are introduced on the login page.
Social support	
Social learning	An online forum is available in PACE, allowing users to write comments, ask questions, or share experiences anonymously.

[a]PSD: Persuasive Systems Design.

[b]PACE: Programa de AutoCuidado de Estoma.

XSL•FO
RenderX

Primary Task Support

We incorporated tailoring and self-monitoring principles. Thematic analysis revealed that many participants desired multiple family members to be trained in ostomy care. To address this, we applied the tailoring principle by enabling role selection and rendering customized content. During registration, users could choose roles (patient, caregiver, or other support) and have the flexibility to change roles later in the user profile settings. The content and layout of the website were rearranged according to user roles, enhancing the user experience. For UI design, our team decided to use a card layout (clickable containers with images and a brief content summary) with a progress bar at the bottom (Figure 2B) by displaying completed tasks and allowing users to resume where they left off to support self-monitoring. Users could track their progress in completing recommended videos and audios on their profile page.

Dialogue Support

To reduce fear and stress and encourage engagement, we established a reward mechanism where users could unlock new cartoon ostomy avatars and decorations after engaging with content. Additionally, email reminders were also set to prompt participants to revisit the website.

Credibility Support

To ensure the PACE program's credibility, we applied the principles of trustworthiness and real-world feel. We presented truthful, fair, and unbiased information through videos, text, and images based on user preferences. The information was developed based on scientific evidence and clinical guidelines with guidance from health care professionals. Videos were recorded by a certified Wound and Ostomy Care Nurse from the University of North Carolina at Chapel Hill. Additionally, a detailed introduction of the research team was also displayed on the login page for reference.

Social Support

The social learning principle was facilitated through an online forum, "Peer Support" (Figure 2). After logging in to the PACE website, patients and their family members could join support groups and seek advice. They could post, comment, and bookmark discussions, creating a space for the Hispanic community to not only exchange ostomy care knowledge but also provide emotional support to those community members in need, fostering a sense of belonging.

Discussion

Principal Findings

The main findings of this study revealed that Hispanic patients with cancer and caregivers experienced significant emotional distress and faced multiple challenges related to ostomy care management post surgery. Participants emphasized the need for culturally tailored interventions and expressed preferences for bilingual multimedia content, early delivery of the intervention, and the inclusion of family and peer support in ostomy care management. Drawing on these insights and feedback, our multidisciplinary team culturally adapted the PRISMS-Spanish and designed and developed the bilingual PACE intervention.

First, participants in our study highlighted a strong preference for content in Spanish alongside English, aligning with existing literature that emphasizes the importance of bilingual support in efficiently overcoming linguistic barriers [32]. Participants who completed the eHEALS survey in Spanish and scored below 26 reported difficulties reading English-language materials and expressed a strong preference for Spanish-translated content. The limited availability of Spanish-language digital resources likely limited their opportunity to access online health information, contributing to lower confidence in using digital health tools and lower self-reported digital health literacy. These findings illustrate the importance of providing bilingual digital health intervention. Thus, we developed the bilingual PACE program by integrating the original PRISMS and the PRISMS-Spanish materials, ensuring cultural and contextual relevance for Hispanic patients and caregivers. Additionally, the bilingual PACE intervention was designed to facilitate family involvement in ostomy care, particularly among Hispanic children and younger generations who are fluent in English and more acculturated to American culture [33,34]. To accommodate participants' preferences and address linguistic barriers, the bilingual platform includes Hispanic dubbing for audio and videos, in addition to text and graphics, which were recorded at appropriate speaking speeds in both English and Spanish, accommodating both patients and their family members who may not speak Spanish or English but still act as caregivers. Recent findings from the Pew Research Center indicate that approximately 65% of third- or later-generation Latinos report limited proficiency in Spanish [35]. By incorporating a bilingual focus, PACE fosters engagement with the entire family, regardless of language proficiency. Our study participants expressed satisfaction with the translation and bilingual design during interviews, highlighting its effectiveness in meeting their needs.

Second, the perceptions of our study participants also played a key role in shaping the delivery format of the PACE intervention. Hispanic patients and caregivers expressed a strong preference for content visualization through a multimedia platform, aligning with the findings of Joshi et al [36] in which Hispanic participants find multimedia programs easy to navigate and effective for obtaining relevant health information. In response, PACE delivers content in various formats, including video, audio, images, and text, to cater to diverse learning preferences. To further enhance engagement, we incorporated culturally relevant elements, such as images of Hispanic foods and vibrant colors like blue and orange, which resonate with the community. Additionally, based on participants' preferences and feedback, similar to the original PRISMS, PACE was designed to be accessible across different devices, including personal computers, mobile phones, and tablets.

Third, participants' feedback affirmed the comprehensive content of the original PRISMS and guided the design of PACE to improve QOL for patients and caregivers. Participants shared their experiences living with an ostomy, often expressing feelings of embarrassment and body image issues following surgery. By increasing knowledge and familiarity, PACE helps

XSL•FO
RenderX

Hispanic patients and caregivers integrate ostomy care into their daily lives with greater ease and positivity.

Additionally, our study revealed that family members of Hispanic cancer survivors are often expected to provide emotional support, assist with daily living, and manage overall health care [37]. This cultural expectation includes providing comprehensive information about the illness [38], treatment details, anticipated side effects, hands-on care skills [39,40], and strategies for stress reduction [41,42]. However, these informational needs are often unmet [41-44]. To fill this gap, PACE delivers tailored information for both patients and their caregivers and family members, enhancing their knowledge and skills in ostomy care. This approach not only enables them to provide better support to patients but also helps maintain their own QOL.

Finally, and most importantly, we used PSD principles to enhance user engagement in the web-based PACE interventions. A systematic review [45] demonstrates that PSD can significantly enhance adherence to web-based interventions, particularly those targeting chronic conditions and lifestyle changes [46,47]. This is especially relevant for patients with cancer living with ostomies, who face substantial changes in their health, lifestyle, behaviors, and daily routines [48]. To promote effective self-management, patients and their families must consistently engage with web-based interventions [49]. However, a notable challenge is that Hispanics are reportedly less likely to use health-related applications compared to other racial groups [50]. Moreover, the absence of user input during program development can further limit the success of these interventions [51]. In this study, we applied PSD principles in the design of the PACE intervention to enhance user engagement and ostomy care self-management. Building on the features of the original PRISMS intervention and incorporating participant feedback, we implemented key PSD principles to enhance engagement. Findings from this study are consistent with prior research [52-54], demonstrating that integrating persuasive principles into intervention design can significantly improve user engagement and support the long-term effectiveness of web-based platforms. The following are examples of how we integrate PSD principles into the design of PACE.

1. Self-monitoring: Participants were encouraged to track their progress in learning ostomy care, increasing awareness, and reinforcing skill development.
2. Rewards and reminders: These features were incorporated to foster continuous use of the program, promoting the acquisition of strong ostomy care skills [55].
3. Trustworthiness and readability: The PACE intervention ensured information was truthful, fair, and unbiased, building trust with users through a user-friendly interface.
4. Social learning: An online forum was integrated, enabling peer support and communication within the PACE intervention.

Strengths and Limitations

To our knowledge, this is the first study to apply affinity diagramming and PSD principles in the design of an ostomy care intervention. Given the limited access to high-quality online health resources among the Hispanic population [56] and building on evidence that Spanish-language, tailored health information in Spanish improves reach and effectiveness in this underserved group [57], our multidisciplinary team intentionally incorporated bilingual and culturally sensitive features to enhance patient engagement and family involvement in ostomy care. During cancer survivorship, ostomy creation as part of cancer treatment is a life-changing event. Traditional education, typically delivered through brief in-hospital demonstrations, can be overwhelming for patients recovering from surgery and coping with changes in physical function and body image. PACE addresses this gap by providing accessible, bilingual resources designed to meet the cultural and informational needs of Hispanic families. By supporting both patients and caregivers during the critical care transition from professional hospital care to self-management at home, PACE aims to improve health outcomes and support their caregivers' well-being.

However, the following limitations warrant further exploration. First, this study lacked participation from family members other than the first-line caregivers in the interviews. Participants emphasized the importance of involving additional family members in ostomy care. Future studies should address this gap by including a broader range of family members of Hispanic patients with cancer to better capture their diverse experiences and needs. Second, participants' educational attainment was at the high school level or above, which may limit the generalizability of the findings, given that a large proportion of Hispanic adults—particularly older or foreign-born individuals—have not completed high school [58]. Future research should prioritize recruiting individuals with lower educational attainment to rigorously design and evaluate digital interventions like PACE. Third, most patients in this study were more than 1 year post ostomy surgery, so their insights may differ from those who are newly diagnosed or recently discharged and still adjusting to life with an ostomy. We intentionally recruited this population because they had already navigated the postostomy survivorship journey and were likely to be more motivated, engaged, and interested in accessing online information compared with those who are fresh out of ostomy-creation surgery. While their reflection may be subject to recall bias, we encouraged them to draw on their lived experiences. Their insights were valuable in informing an intervention design that is potentially feasible and acceptable for individuals in the early stages of recovery. However, future studies should evaluate the acceptability, usability, feasibility, and effects of the intervention among patients and caregivers managing newly created ostomies to ensure its relevance and applicability across various stages of recovery.

Fourth, this study used descriptive statistics to analyze the quantitative data, which may limit the ability to draw statistical inferences about relationships between variables, such as age, education, digital literacy, comorbidities, and intervention preferences. Future studies with larger sample sizes need to incorporate statistical methods (eg, regression analysis or causal inference techniques) to examine associations between participant characteristics and engagement in digital health interventions. Finally, this study reported percent agreement to assess coding consistency, given the iterative coding approach and small sample size. Future work should consider applying

XSL•FO
RenderX

more robust methods, such as Cohen κ, to evaluate interrater reliability in large-scale studies.

Despite these limitations, the findings support the cultural adaptation of the English version of the PRISMS intervention into a user-friendly bilingual PACE web-based program for Hispanic patients and families, a population often underrepresented in clinical intervention research. Future studies should recruit participants with diverse backgrounds to evaluate the PACE intervention in other underserved communities, focusing on improving health outcomes and addressing health disparities.

Research and Clinical Implications

The PACE web-based program addresses critical challenges in ostomy care, including skin irritation, leakage, treatment side effects, and social and emotional stress with the goal of improving health outcomes for patients and caregivers. PACE provides adaptable techniques for different stages of recovery, standardizes patient and family education, and supports self-management in the hospital and at home. Its scalability and accessibility are particularly valuable for individuals in remote areas, such as those in rural Texas. By integrating findings from the PRISMS pilot feasibility study and current research, PACE is positioned for testing in a sufficiently powered clinical trial to evaluate its efficacy and effectiveness. Making the program accessible in hospital rooms could further enhance its impact on patient education and care transitions.

Conclusions

Our findings revealed the experiences of Hispanic patients and caregivers managing ostomy care for cancer treatment, including their support-seeking behaviors, postdischarge challenges, and preferences related to language, timing of intervention delivery, digital platforms, and the inclusion of family and peer support. Informed by these insights, we culturally adapted the PRISMS-Spanish program and applied PSD principles to design the bilingual, web-based PACE intervention. This tailored approach aims to enhance patient and caregiver engagement, strengthen their ostomy care skills, and ultimately improve their health outcomes. Future research will include a pilot testing for usability and feasibility of the PACE intervention among larger and more diverse Hispanic populations.

Acknowledgments

We extend our heartfelt gratitude to the patients and caregivers who participated in this study, the health care professionals who supported our efforts, and our dedicated research team members for their invaluable contributions. This line of work was done in collaboration with WOCNs and ten bilingual translators from diverse cultural and healthcare backgrounds. Generative artificial intelligence (AI) tools were not used in generating manuscript content.

Funding

This study is an extension of the PRISMS (Patient-Reported Outcomes-Informed Symptom Management System) funded by the NCI (National Cancer Institute) Cancer Alliance Pilot Award (2UG1CA189823, PI [principal investigator]: LS).

Data Availability

The datasets for this study are not publicly available due to the sensitive nature of patient and caregiver qualitative and quantitative data, as well as privacy and confidentiality requirements outlined by the University of Texas Health San Antonio's IRB. But data are available from the corresponding author on reasonable request.

Authors' Contributions

LS led the conceptualization of the study. LS, FY, AV, JL, RD, and XW supported the design of the digital platform. RS, MB, AV, and LRR contributed to participant recruitment and data collection. LS, AV, LRR, JL, NVR, and XW contributed to data interpretation. AV, LRR, JL, and XW performed the formal data analysis. FY, AV, and XM contributed to data visualization and digital platform design. AV managed, coordinated, and executed research activities. All authors contributed to the draft review and critical revision of the manuscript.

Conflicts of Interest

None declared.

Checklist 1
Mixed Methods Reporting in Rehabilitation and Health Sciences (MMR-RHS).
[PDF File, 699 KB - cancer_v12i1e70354_app1.pdf]

References

1. Krouse RS, Grant M, McCorkle R, et al. A chronic care ostomy self-management program for cancer survivors. Psychooncology 2016 May;25(5):574-581. [doi: 10.1002/pon.4078] [Medline: 26804708]
2. Yu H, Ryu E. Colorectal cancer survivors' inner strength, multiple identities, and quality of life by gender and ostomy presence: a cross-sectional study. Korean J Adult Nurs 2024;36(2):171. [doi: 10.7475/kjan.2024.36.2.171]

XSL•FO
RenderX

3. Victorson D, Banas J, Smith J, et al. eSalud: designing and implementing culturally competent ehealth research with latino patient populations. Am J Public Health 2014 Dec;104(12):2259-2265. [doi: 10.2105/AJPH.2014.302187] [Medline: 25320901]
4. Costas-Muñiz R, Hunter-Hernández M, Garduño-Ortega O, Morales-Cruz J, Gany F. Ethnic differences in psychosocial service use among non-Latina white and Latina breast cancer survivors. J Psychosoc Oncol 2017;35(4):424-437. [doi: 10.1080/07347332.2017.1310167] [Medline: 28332946]
5. Creamer J. Inequalities persist despite decline in poverty for all major race and hispanic origin groups. United States Census Bureau. URL: https://www.census.gov/library/stories/2020/09/poverty-rates-for-blacks-and-hispanics-reached-historic-lows-in-2019.html [accessed 2026-01-20]
6. Viale PH. The American Cancer Society's Facts & Figures: 2020 Edition. J Adv Pract Oncol 2020 Mar;11(2):135-136. [doi: 10.6004/jadpro.2020.11.2.1] [Medline: 33532112]
7. Duran M. Rural Hispanic health care utilization. Online J Rural Nurs Health Care 2012;12(1):49-54. [doi: 10.14574/ojrnhc.v12i1.19]
8. Sehar U, Rawat P, Choudhury M, et al. Comprehensive understanding of Hispanic caregivers: focus on innovative methods and validations. J Alzheimers Dis Rep 2023;7(1):557-574. [doi: 10.3233/ADR-220094] [Medline: 37313490]
9. Sharp SP, Ata A, Chismark AD, et al. Racial disparities after stoma construction in colorectal surgery. Colorectal Dis 2020 Jun;22(6):713-722. [doi: 10.1111/codi.14943]
10. Kronenfeld JP, Graves KD, Penedo FJ, Yanez B. Overcoming disparities in cancer: a need for meaningful reform for Hispanic and Latino cancer survivors. Oncologist 2021 Jun;26(6):443-452. [doi: 10.1002/onco.13729] [Medline: 33594785]
11. Bisht J, Rawat P, Sehar U, Reddy PH. Caregivers with cancer patients: focus on Hispanics. Cancers (Basel) 2023 Jan 19;15(3):626. [doi: 10.3390/cancers15030626] [Medline: 36765585]
12. Golpazir-Sorkheh A, Ghaderi T, Mahmoudi S, Moradi K, Jalali A. Family-centered interventions and quality of life of clients with ostomy. Nurs Res Pract 2022;2022:9426560. [doi: 10.1155/2022/9426560] [Medline: 36072914]
13. Heydari A, Manzari ZS, Pouresmail Z. Nursing intervention for quality of life in patients with ostomy: a systematic review. Iran J Nurs Midwifery Res 2023;28(4):371-383. [doi: 10.4103/ijnmr.ijnmr_266_22] [Medline: 37694203]
14. Solomon M, Wagner SL, Goes J. Effects of a web-based intervention for adults with chronic conditions on patient activation: online randomized controlled trial. J Med Internet Res 2012 Feb 21;14(1):e32. [doi: 10.2196/jmir.1924] [Medline: 22353433]
15. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of web-based vs. non-web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40. [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
16. Larsen B, Marcus B, Pekmezi D, Hartman S, Gilmer T. A web-based physical activity intervention for Spanish-speaking Latinas: a costs and cost-effectiveness analysis. J Med Internet Res 2017 Feb 22;19(2):e43. [doi: 10.2196/jmir.6257] [Medline: 28228368]
17. Gans KM, Dulin A, Palomo V, et al. A tailored web- and text-based intervention to increase physical activity for Latino men: protocol for a randomized controlled feasibility trial. JMIR Res Protoc 2021 Jan 29;10(1):e23690. [doi: 10.2196/23690] [Medline: 33512327]
18. Fortmann AL, Savin KL, Clark TL, Philis-Tsimikas A, Gallo LC. Innovative diabetes interventions in the U.S. hispanic population. Diabetes Spectr 2019 Nov;32(4):295-301. [doi: 10.2337/ds19-0006] [Medline: 31798285]
19. Xu S, Tan X, Ma C, et al. An eHealth symptom and complication management program for cancer patients with newly created ostomies and their caregivers (Alliance): a pilot feasibility randomized trial. BMC Cancer 2023 Jun 10;23(1):532. [doi: 10.1186/s12885-023-10919-x] [Medline: 37301841]
20. Terry G, Hayfield N, Clarke V, Braun V. Thematic analysis. In: The SAGE Handbook of Qualitative Research in Psychology: Sage Publications. URL: https://methods.sagepub.com/hnbk/edvol/the-sage-handbook-of-qualitative-research-in-psychology-second-edition/chpt/2-thematic-analysis#_ [accessed 2026-01-22] [doi: 10.4135/9781526405555]
21. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009 Mar;24. [doi: 10.17705/1CAIS.02428]
22. Guan V, Zhou C, Wan H, et al. A novel mobile app for personalized dietary advice leveraging persuasive technology, computer vision, and cloud computing: development and usability study. JMIR Form Res 2023 Aug 7;7:e46839. [doi: 10.2196/46839] [Medline: 37549000]
23. Idrees AR, Kraft R, Mutter A, Baumeister H, Reichert M, Pryss R. Persuasive technologies design for mental and behavioral health platforms: a scoping literature review. PLOS Digit Health 2024 May;3(5):e0000498. [doi: 10.1371/journal.pdig.0000498] [Medline: 38753889]
24. ACS demographic and housing estimates. United States Census Bureau. 2023. URL: https://data.census.gov/table/ACSDP1Y2023.DP05?q=san%20antonioU.S [accessed 2025-11-14]
25. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
26. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson comorbidity index: a critical review of clinimetric properties. Psychother Psychosom 2022;91(1):8-35. [doi: 10.1159/000521288] [Medline: 34991091]

XSL•FO
RenderX

27. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
28. Guest G, Bunce A, Johnson L. How many interviews are enough?: an experiment with data saturation and variability. Field methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
29. Miles MB, Huberman AM. Qualitative Data Analysis: An Expanded Sourcebook, 2nd edition 1994.
30. Harboe G, Huang EM. Real-world affinity diagramming practices: bridging the paper-digital gap. Presented at: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems; Apr 18-23, 2015 p. 95-104 URL: https://dl.acm.org/doi/10.1145/2702123.2702561 [accessed 2025-12-14] [doi: 10.1145/2702123.2702561]
31. Asbjørnsen RA, Smedsrød ML, Solberg Nes L, et al. Persuasive system design principles and behavior change techniques to stimulate motivation and adherence in electronic health interventions to support weight loss maintenance: scoping review. J Med Internet Res 2019 Jun 21;21(6):e14265. [doi: 10.2196/14265] [Medline: 31228174]
32. Banas JR, Victorson D, Gutierrez S, Cordero E, Guitleman J, Haas N. Developing a peer-to-peer mHealth application to connect Hispanic cancer patients. J Canc Educ 2017 Mar;32(1):158-165. [doi: 10.1007/s13187-016-1066-6]
33. Toppelberg CO, Collins BA. Language, culture, and adaptation in immigrant children. Child Adolesc Psychiatr Clin N Am 2010 Oct;19(4):697-717. [doi: 10.1016/j.chc.2010.07.003] [Medline: 21056342]
34. Hakimzadeh S, Cohn DV. English usage among hispanics in the united states. : Pew Research Center; 2007 URL: https://www.pewresearch.org/race-and-ethnicity/2007/11/29/english-usage-among-hispanics-in-the-united-states/ [accessed 2025-12-14]
35. Mora L, Lopez MH. Latinos' views of and experiences with the Spanish language. : Pew Research Center; 2023 URL: https://www.pewresearch.org/race-and-ethnicity/2023/09/20/latinos-views-of-and-experiences-with-the-spanish-language/ [accessed 2025-12-14]
36. Joshi A, Wilhelm S, Aguirre T, Trout K, Amadi C. An interactive, bilingual touch screen program to promote breastfeeding among Hispanic rural women: usability study. JMIR Res Protoc 2013 Nov 7;2(2):e47. [doi: 10.2196/resprot.2872] [Medline: 24200498]
37. Enriquez M. Hispanic Health Care International: meeting the needs of our readers. Hisp Health Care Int 2019 Dec;17(4):142. [doi: 10.1177/1540415319875247] [Medline: 31550917]
38. Badger TA, Sikorskii A, Segrin C. Contextual and cultural influences on caregivers of Hispanic cancer survivors. Semin Oncol Nurs 2019 Aug;35(4):359-362. [doi: 10.1016/j.soncn.2019.06.008] [Medline: 31229343]
39. Brown CG. A guide to oncology symptom management. Oncology Nursing Society 2010.
40. Lopez V, Copp G, Molassiotis A. Male caregivers of patients with breast and gynecologic cancer: experiences from caring for their spouses and partners. Cancer Nurs 2012;35(6):402-410. [doi: 10.1097/NCC.0b013e318231daf0] [Medline: 22067685]
41. Loke AY, Liu CFF, Szeto Y. The difficulties faced by informal caregivers of patients with terminal cancer in Hong Kong and the available social support. Cancer Nurs 2003 Aug;26(4):276-283. [doi: 10.1097/00002820-200308000-00004] [Medline: 12886118]
42. Sklenarova H, Krümpelmann A, Haun MW, et al. When do we need to care about the caregiver? Supportive care needs, anxiety, and depression among informal caregivers of patients with cancer and cancer survivors. Cancer 2015 May 1;121(9):1513-1519. [doi: 10.1002/cncr.29223] [Medline: 25677095]
43. Chua GP, Ng QS, Tan HK, Ong WS. Caregivers of cancer patients: what are their information-seeking behaviours and resource preferences? Ecancermedicalscience 2020;14:1068. [doi: 10.3332/ecancer.2020.1068] [Medline: 32728384]
44. van Ryn M, Sanders S, Kahn K, et al. Objective burden, resources, and other stressors among informal cancer caregivers: a hidden quality issue? Psychooncology 2011 Jan;20(1):44-52. [doi: 10.1002/pon.1703] [Medline: 20201115]
45. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152. [doi: 10.2196/jmir.2104] [Medline: 23151820]
46. Alpay L, Doms R, Bijwaard H. Embedding persuasive design for self-health management systems in Dutch healthcare informatics education: application of a theory-based method. Health Informatics J 2019 Dec;25(4):1631-1646. [doi: 10.1177/1460458218796642] [Medline: 30192696]
47. Asbjørnsen RA, Hjelmesæth J, Smedsrød ML, et al. Combining persuasive system design principles and behavior change techniques in digital interventions supporting long-term weight loss maintenance: design and development of eCHANGE. JMIR Hum Factors 2022 May 27;9(2):e37372. [doi: 10.2196/37372] [Medline: 35622394]
48. Ercolano E, Grant M, McCorkle R, et al. Applying the chronic care model to support ostomy self-management: implications for oncology nursing practice. Clin J Oncol Nurs 2016 Jun 1;20(3):269-274. [doi: 10.1188/16.CJON.20-03AP] [Medline: 27206293]
49. Moulaei K, Iranmanesh E, Ahmadian L. The impact of health technologies on ostomy care: a systematic review of health technologies impact on ostomy care. J Wound Ostomy Continence Nurs 2023;50(6):489-494. [doi: 10.1097/WON.0000000000001021] [Medline: 37966077]
50. Chae J. A comprehensive profile of those who have health-related apps. Health Educ Behav 2018 Aug;45(4):591-598. [doi: 10.1177/1090198117752784] [Medline: 29353544]

XSL•FO
RenderX

51. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. J Med Internet Res 2016 Jan 8;18(1):e6. [doi: 10.2196/jmir.4790] [Medline: 26747176]
52. Sporrel K, Nibbeling N, Wang S, Ettema D, Simons M. Unraveling mobile health exercise interventions for adults: scoping review on the implementations and designs of persuasive strategies. JMIR Mhealth Uhealth 2021 Jan 18;9(1):e16282. [doi: 10.2196/16282] [Medline: 33459598]
53. Asbjørnsen RA, Wentzel J, Smedsrød ML, et al. Identifying persuasive design principles and behavior change techniques supporting end user values and needs in eHealth interventions for long-term weight loss maintenance: qualitative study. J Med Internet Res 2020 Nov 30;22(11):e22598. [doi: 10.2196/22598] [Medline: 33252347]
54. Abdullahi AM, Orji R, Rabiu AM, Kawu AA. Personality and subjective well-being: towards personalized persuasive interventions for health and well-being. Online J Public Health Inform 2020;12(1):e1. [doi: 10.5210/ojphi.v12i1.10335] [Medline: 32547678]
55. Wang Y, Mo DY, Ma HL. Perception of time in the online product customization process. Ind Manag Data Syst 2023 Feb 27;123(2):369-385. [doi: 10.1108/IMDS-03-2022-0159]
56. Chavarria EA, Christy SM, Feng H, et al. Online health information seeking and eHealth literacy among Spanish language–dominant Latino adults receiving care in a community clinic: Secondary analysis of pilot randomized controlled trial data. JMIR Form Res 2022 Oct 13;6(10):e37687. [doi: 10.2196/37687] [Medline: 35238785]
57. Linke SE, Dunsiger SI, Gans KM, et al. Association between physical activity intervention website use and physical activity levels among Spanish-speaking Latinas: randomized controlled trial. J Med Internet Res 2019 Jul 24;21(7):e13063. [doi: 10.2196/13063] [Medline: 31342902]
58. 2020 profile of Hispanic Americans aged 65 and older. : The Administration for Community Living; 2020.

Abbreviations

CCI: Charlson Comorbidity Index
eHEALS: eHealth literacy scale
IRB: Institutional Review Board
PACE: Programa de AutoCuidado de Estoma
PRISMS: Patient-Reported Outcomes-Informed Symptom Management System
PSD: Persuasive Systems Design
QOL: quality of life
REDCap: Research Electronic Data Capture
UI: user interface
UTHSA: University of Texas Health San Antonio

Edited by N Cahill; submitted 20.Dec.2024; peer-reviewed by RD Reed, X Zeng; revised version received 21.Aug.2025; accepted 25.Aug.2025; published 03.Feb.2026.

Please cite as:
Wang X, Song L, Vondenberger A, Yu F, Liu J, Delgado R, Svatek R, Rodriguez NV, Rivera LR, Bonnen M
Cultural Adaptation of a Web-Based Ostomy Care Intervention for Hispanic Patients With Cancer and Caregivers: Mixed Methods Study
JMIR Cancer 2026;12:e70354
URL: https://cancer.jmir.org/2026/1/e70354
doi:10.2196/70354

© Xiaomeng Wang, Lixin Song, Amy Vondenberger, Fei Yu, Jia Liu, Roxana Delgado, Robert Svatek, Natalia Villegas Rodriguez, Lurheinna Rosado Rivera, Mark Bonnen. Originally published in JMIR Cancer (https://cancer.jmir.org), 3.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

New Technologies and Digital Health Tools in Patients With Solid Tumors and Hematological Malignancies: Cross-Sectional Multicenter Survey Study

Alberto Lopez-Garcia[1†*], MD; Carola Diaz-Aizpun[2*], MD; Beatriz Gallego-Diaz[1], MD; Carolina Miranda-Castillo[3], MD; Maria Yuste-Platero[4], MD; Pilar Beltran-Alvarez[5], MD; Cristina Carames-Sanchez[6], MD, PhD; Jesus Garcia-Foncillas[6], MD, PhD; Pilar Llamas-Sillero[7], MD, PhD; Marta del Olmo-Rodriguez[8], MSc; Jorge Short-Apellaniz[8], MD; Bernadette Pfang[8], MD, PhD; Javier Arcos-Campillo[8], MD; Raul Cordoba[1], MD, PhD

[1]Lymphoma Unit, Department of Hematology, Fundación Jiménez Díaz University Hospital, Health Research Institute IIS-FJD, Madrid, Spain

[2]Department of Hematology, Regional University Hospital, Malaga, Spain

[3]Department of Hematology, Rey Juan Carlos University Hospital, Madrid, Spain

[4]Department of Hematology, Villalba General University Hospital, Madrid, Spain

[5]Department of Hematology, Infanta Elena University Hospital, Madrid, Spain

[6]Department of Oncology, Fundación Jiménez Díaz University Hospital, Health Research Institute IIS-FJD, Madrid, Spain

[7]Department of Hematology, Fundación Jiménez Díaz University Hospital, Health Research Institute IIS-FJD, Madrid, Spain

[8]Clinical and Organizational Research Unit, Quironsalud Hospital, Avenida Reyes Católicos 2, Madrid, Spain

[†]deceased

[*]these authors contributed equally

Corresponding Author:
Bernadette Pfang, MD, PhD
Clinical and Organizational Research Unit, Quironsalud Hospital, Avenida Reyes Católicos 2, Madrid, Spain

Abstract

Background: Barriers to eHealth use include lack of technological infrastructure, resistance to change, and inequities in access. However, patterns of access to and use of eHealth tools in people being treated for cancer have not been fully described in the literature.

Objective: Our aim was to describe the patterns of access to and use of eHealth tools among outpatients receiving treatment for hematological malignancies and solid tumors.

Methods: We conducted a cross-sectional multicenter study using a survey offered to patients aged over 18 years receiving outpatient treatment for hematological malignancies or solid tumors at 4 teaching hospitals in Madrid, Spain, between February 1, 2021, and November 30, 2021. The survey instrument featured questions about patients' demographic and social characteristics, cancer diagnosis, use of information and communication technology (ICT), use and opinions of the Patient Portal, and changes in ICT use during the COVID-19 pandemic. To study the relationship between the different variables, 2-tailed Student *t* tests or ANOVA were used for variables with normal distribution, and the Mann-Whitney or Kruskal-Wallis tests were used for variables with nonnormal distribution. Statistical analyses were performed using SPSS (version 25; IBM Corp) for Windows.

Results: In total, 200 patients were included in our study. Median age was 60 (range 21 - 87) years. A total of 130 (65%) patients presented with hematological malignancies. Most (n=181, 90.5%) patients considered that eHealth tools might help to improve communication with the medical team during their treatment. Retired participants (28.6% vs 71.4%; *P*<.001), those older than 60 years (26% vs 74%; *P*<.001), and those without higher education (2.6% vs 97.4%; *P*<.001) showed significantly lower rates of internet use, with no observed sex-related differences. A total of 177 (88.5%) patients found the Patient Portal useful, and 140 (70%) reported increased use of ICT due to the COVID-19 pandemic.

Conclusions: Most (177/200, 88.5%) patients viewed eHealth tools as useful and believed that it was helpful to improve communication with their care team. However, notable gaps in the use of eHealth were observed in certain groups of patients, with significant differences in use due to age, education, and employment status. Strategies to identify subgroups at risk for unequal access to digital health, as well as to facilitate access and use, are warranted.

(JMIR Cancer 2026;12:e58823) doi:10.2196/58823

XSL•FO
RenderX

KEYWORDS

eHealth; oncohematology; telemedicine; survey; digital health tools; tumor; hematological malignancies; barrier; cancer; cancer care; cross-sectional study; cancer diagnosis

Introduction

Using telemedicine to provide care for patients with hematological malignancies has proven to increase access to care and levels of patient satisfaction [1] compared with face-to-face care and leads to cost savings for health care systems. The potential of technology to facilitate the care of immunosuppressed patients, such as those with hematological cancer, has been demonstrated recently during the COVID-19 pandemic [2,3]. During the pandemic, clinical guidelines recommended stopping treatment and performing remote follow-up whenever possible [4], leading to a sharp increase in the use of telemedicine in oncology, with up to 80% of centers using some form of remote care [5].

Despite a sustained increase in telemedicine as a model of care delivery in the postpandemic era, many systems are not adequately prepared for widespread implementation due to a lack of technological infrastructure [6]. Other barriers to the large-scale implementation of telemedicine in the hematology setting include resistance to change from health care providers and patients [7], and unequal access to eHealth among different patient groups based on race, age, and other sociodemographic characteristics [8,9]. Research on patient preferences regarding digital models of health care delivery shows disparate results, with a current lack of evidence as to whether the theoretical benefits of telemedicine translate into real benefits in routine clinical practice [4,5].

Our institution, a hospital network in Madrid, Spain, serving a population of more than 1 million inhabitants, launched an eHealth platform (the Quironsalud Patient Portal) in 2019. The platform allows patients to view appointments and test results, communicate with their health care providers through web dialogues, answer patient-reported outcome measurement (PROM) and patient-reported experience measurement (PREM) surveys, and access other health-related functions such as health education videos. Our study aimed to describe the patterns of access to and use of digital health tools in people with hematological malignancies and solid tumors undergoing outpatient treatment. We also aimed to describe patients' experiences with the Patient Portal and to report changes in digital health tool use caused by the COVID-19 pandemic.

Methods

Study Design

We conducted a cross-sectional, multicenter study, using a survey offered to patients with hematological malignancies or solid tumors receiving outpatient treatment at four teaching hospitals in Madrid, Spain. The primary endpoint of the study was to describe the patterns of access to and use of digital health tools across a sample of outpatients receiving treatment for hematological malignancies and solid tumors. The secondary endpoint was to report patients' opinions of the Patient Portal and identify changes in the use of digital health tools caused by the COVID-19 pandemic.

Setting and Participants

Centers included in the study were 4 publicly owned, privately managed hospitals from the Madrid Regional Health System—Fundación Jiménez Díaz University Hospital, Infanta Elena University Hospital, General Villalba University Hospital, and Rey Juan Carlos University Hospital. Patients aged over 18 years with hematological malignancies (lymphoma, myeloma, and acute leukemia) or solid tumors receiving oral or intravenous treatment as outpatients were included. On average, the 4 centers treat 2000 outpatients with cancer per month. We excluded patients who did not undergo active treatment during the study period. Recruitment for the study was conducted between February 1, 2021, and November 30, 2021.

Variables and Measurements

The survey instrument featured questions about patients' demographic and social characteristics, cancer diagnosis, use of information and communication technology (ICT), use and opinions of the Patient Portal, and changes in ICT use during the COVID-19 pandemic.

The survey was developed by 2 hematologists (ALG and RC), with survey items developed based on the results of a previous literature review featuring articles on digital health in patients with cancer published in English or Spanish from 2010 to 2020. Survey items were presented as yes or no or multiple-choice questions. Before recruitment began, a pilot questionnaire was administered to 10 patients to validate survey design and comprehension, with no changes made in the original design. The final survey included 30 items divided into 5 thematic blocks, including demographic and social characteristics, cancer diagnosis, ICT use, Patient Portal use, and COVID-19 pandemic–related aspects. Multimedia Appendix 1 shows the items featured in the survey.

All patients who visited the oncology outpatient clinic to receive intravenous treatment during the inclusion period were invited to participate in the study. After informed consent was obtained, participants were handed a paper copy of the survey by the treating physician. The completed surveys were returned in person.

Bias

We chose to administer a paper-based survey to avoid selection bias caused by the underrepresentation of patients who were not regular ICT users. Illegible or incomplete surveys were excluded.

Statistical Methods

A descriptive analysis of the population was conducted. Qualitative variables were described using frequency tables. Continuous variables that followed a normal distribution were described by mean and SD. Continuous variables that did not

XSL•FO
RenderX

follow a normal distribution were described using the median and IQR.

Normality was verified using the Kolmogorov-Smirnov test. To test for differences in the use of internet and eHealth tools between different groups (in terms of age, sex, and employment status), Student *t* tests or ANOVA were used for variables with normal distribution, and the Mann-Whitney or Kruskal-Wallis tests were used for variables with nonnormal distribution. Statistical analyses were performed using SPSS (version 25; IBM Corp) for Windows.

Ethical Considerations

This study was conducted in accordance with the ethical standards of the institutional research committee, Hospital Clinical Research Ethics Committee (ER_PIC018-20_FJD), and the Declaration of Helsinki. All participants provided written informed consent before participating in the study. All data were deidentified before performing statistical analysis, and the privacy and confidentiality of the research participants' data were maintained throughout the study. No compensation was offered to participants or researchers.

Results

A total of 225 questionnaires were received, with 200 (88.9%) questionnaires included in the final analysis (approximately 0.1% of the total number of eligible patients). Reasons for exclusion from the final analysis included incomplete surveys (11/25, 44%), having marked multiple options on a single-response question (4/25, 16%), and failure to meet eligibility criteria in terms of age (10/25, 40%). Demographic, clinical, and social characteristics of participants and their use of ICT are summarized in Table 1. The median age of the participants was 60 (range 21 - 87) years, and 119 (59.5%) participants were male. More than half (n=116, 58%) of the participants had attended higher education institutions. Regarding employment status, 84 (42%) were actively working and 84 (42%) were retired. Most patients had hematological malignancies (n=130, 65%; Table 1).

Notably, 172 (86%) of the 200 patients owned smartphones, and 81 (40.5%) participants owned wearable devices. Regarding reported use of ICT, 136 (68%) patients reported daily use of the internet, and 85 (42.5%) patients conducted health-related research via the internet. Patients older than 60 years (26% vs 74%; *P*<.001; odds ratio [OR] 1.35, 95% CI 1.21-1.51), without higher education (2.6% vs 97.4%; *P*<.001; OR 11.05, 95% CI 3.45 - 35.48) and retired (28.6% vs 71.4%; *P*<.001; OR 1.36, 95% CI 1.19-1.57) reported higher rates of never using the internet for browsing (Table 1). However, all patients used social media. Most (n=181, 90.5%) patients considered that eHealth tools might help to improve communication with the medical team during their treatment. The preferred way to communicate with the clinical team was via a smartphone via a WhatsApp-like chat (n=90, 45%).

Regarding the use of eHealth tools, no sex differences were found. Factors associated with higher rates of health-related research via the internet included being over 60 years old (67.3% vs 32.7%; *P*=.02; OR 2.43, 95% CI 1.37 - 4.34) and having higher education (44.8% vs 55.2%; *P*<.001; OR 1.65, 95% CI 1.3 - 2.1). In contrast, employment status was not significantly associated with the use of the internet to conduct health-related research.

Regarding use of the Patient Portal, we found that patients over 60 years (86.5% vs 13.5%; *P*=.009; OR 4.82, 95% CI 1.34 - 17.35) and without higher education (85.7% vs 14.3%; *P*=.01; OR 3.31, 95% CI 1.21 - 9.05) had lower Patient Portal registration rates (Table 2). Regardless of registration status, patients aged more than 60 years (78.8% vs 21.2%; *P*<.001), without higher education (94.8% vs 5.2%; *P*=.002), and with an employment status of "unemployed" or "retired" (*P*<.001), preferred not to use the Patient Portal to communicate with their care team (Table 3). Regarding the perceived usefulness of the Patient Portal, more than 80% (177/200, 88.5%) of patients reported the portal's PROMs and PREMs questionnaires as "useful." No significant differences were found between groups except regarding employment status, with almost 20% of retired patients considering the PROMs and PREMs questionnaires as "not useful" (Table 4).

When asked whether the use of ICT could improve the patient-clinician relationship, patients' answers varied. Subgroups tending to consider that ICT could worsen their relationship with the clinical team included patients over 60 years of age and retired patients. Between 20% (40/200) and 40% (80/200) of patients across all groups considered that the use of ICT did not affect their relationship with their care team. Finally, regarding the COVID-19 pandemic, 140 (70%) patients reported higher rates of ICT use. A total of 120 (60%) patients reported no perceived changes in their relationship with the care team, 60 (30%) patients reported an improvement, and 20 (10%) patients reported a decline in their relationship with the health care professional.

XSL•FO
RenderX

Table . Characteristics of the study population (N=200) and use of information and communication technologies.

Characteristics	Values
Age (y), median (range)	60 (21 - 87)
Sex, n (%)	
Male	119 (59.5)
Female	81 (40.5)
Tumor type, n (%)	
Lymphoma	86 (43)
Multiple myeloma	30 (15)
Others	17 (8.5)
Leukemia	14 (7)
Breast cancer	14 (7)
Lung cancer	14 (7)
Gastrointestinal cancer	11 (5.5)
Prostate cancer	7 (3.5)
Gynecological cancer	4 (2)
Genitourinary cancer	3 (1.5)
Education, n (%)	
No or basic	84 (42)
University or superior	116 (58)
Employment, n (%)	
Student	10 (5)
Active	84 (42)
Unemployed	22 (11)
Retired	84 (42)
Department, n (%)	
Hematology	146 (73)
Oncology	54 (27)
Use of internet, n (%)	
Sometimes or always	173 (86.5)
Smartphone	172 (86)
Wearable	81 (40.5)
Internet access, n (%)	
Computer	118 (59)
Tablet	68 (34)
Smartphone	166 (83)
Health research on the internet	85 (42.5)
Preference to communicate with medical team, n (%)	
Smartphone	90 (45)
Email	42 (21)
App	26 (13)
Video call	24 (12)

XSL•FO
RenderX

Table . Demographic characteristics of patients and their correlation with Patient Portal use.

	Yes, n (%)	No, n (%)	*P* value	Odds ratio (95% CI)
Age (y)			.009	4.82 (1.34-17.35)
<60	93 (96.9)	3 (3.1)		
>60	90 (86.5)	14 (13.5)		
Education			.01	3.31 (1.21-9.05)
No or basic	72 (85.7)	12 (14.3)		
University or superior	111 (95.7)	5 (4.3)		
Employment			.22	—[a]
Student	10 (100)	0 (0)		
Active	79 (94)	5 (6)		
Unemployed	18 (81.8)	4 (18.2)		
Retired	76 (90.5)	8 (9.5)		

[a]Not available.

Table . Patients' preferences regarding communication with health care professionals via the Patient Portal.

	Yes, n (%)	No, n (%)	*P* value	Odds ratio (95% CI)
Age (y)			<.001	1.04 (1-1.08)
<60	76 (100)	0 (0)		
>60	92 (78.8)	22 (21.2)		
Education			.002	—[a]
No or basic	68 (81)	16 (19)		
University or superior	110 (94.8)	6 (5.2)		
Employment			<.001	—
Student	10 (100)	0 (0)		
Active	82 (97.6)	2 (2.4)		
Unemployed	22 (100)	0 (0)		
Retired	64 (76.2)	20 (23.8)		

[a]Not available.

Table . Patients' perception of the usefulness of the quality-of-life questionnaires.

	Yes, n (%)	No, n (%)	*P* value
Age (y)			.07
≤60	89 (92.7)	7 (7.3)	
>60	88 (84.6)	16 (15.4)	
Education			.54
No or basic	73 (86.9)	11 (13.1)	
University or superior	104 (89.7)	12 (10.3)	
Employment			.03
Student	10 (100)	0 (0)	
Active	78 (92.9)	6 (7.1)	
Unemployed	21 (95.5)	1 (4.5)	
Retired	68 (81)	16 (19)	

XSL•FO
RenderX

Discussion

Principal Findings

This paper presents the results of a multicenter cross-sectional study on patients receiving outpatient cancer treatment for hematological malignancies and solid tumors. Our results demonstrate that most (181/200, 90.5%) patients considered that eHealth tools might help to improve communication with the medical team during their treatment. Retired participants (28.6% vs 71.4%; *P*<.001), those older than 60 years (26% vs 74%; *P*<.001), and those without higher education (2.6% vs 97.4%; *P*<.001) showed significantly lower rates of internet use, with no observed sex-related differences. More than 80% (177/200, 88.5%) of the patients found the Patient Portal useful, and 140 (70%) patients reported increased use of ICT due to the COVID-19 pandemic.

Comparison to Prior Work

Recent studies have demonstrated that eHealth is an effective form of health care delivery in hematology and has been catalyzed by the COVID-19 pandemic. A study focusing on the use of eHealth in patients with cancer [10] reported that patients showed an interest in applications such as managing appointments, obtaining advice about their disease, and communicating with health care professionals. Free access to eHealth tools and medical approval [10,11] are 2 facilitators for eHealth use. However, despite offering potential benefits, unidirectional tools such as SMS text message reminders are not always effective [12], as patient use of ICT predicts the effectiveness of and responsiveness to unidirectional eHealth tools [13]. Clinician reluctance to incorporate eHealth tools due to fears of a possible increase in clinical burden is another barrier to widespread integration [14]. Our study shows that most patients in an outpatient cancer care setting consider that eHealth tools are useful and facilitate communication with their care team during treatment, with the preferred method of communication being a smartphone. Furthermore, almost half (85/200, 42.5%) of the participants conducted health-related research via the internet, highlighting the importance of ensuring access to trustworthy, evidence-based information sources.

Most (140/200, 70%) participants agreed that their use of health-related ICT increased during the COVID-19 pandemic. This finding is consistent with those of many other studies in oncology and other fields of medicine [15-19]. Interestingly, many eHealth initiatives developed or scaled up during the pandemic have now become common clinical practice, with tools such as chatbots gaining increasing recognition as useful resources for patients with cancer [20,21]. While the use of eHealth is increasingly common in cancer care, the effect of ICT on the physician-patient relationship remains unclear. Many studies have concluded that, despite the increasing use of telemedicine, patients still prefer traditional, face-to-face communication over internet-based interactions with clinicians [22,23]. However, other reports suggest that patients value certain aspects of ICT-mediated communication, such as rapid response, easy access, and availability [24]. Our study suggested that most patients did not associate an increased use of eHealth with changes in their relationship with clinicians, while 30% (60/200) of patients reported an improvement, and 10% (20/200) of patients stated that their relationship with clinicians deteriorated with increased eHealth use. Regarding new developments in the field of ICT-mediated health communication, over the last months, generative artificial intelligence–based chatbots using large language models have come to the forefront. However, while critical appraisal of these models regarding cancer-related information demonstrates high accuracy [25], their reliability in the clinical setting has yet to be validated [26], along with important aspects such as ethics, governance, patient preferences, and patient safety [27].

The complexity of oncohematological treatment, the shift toward patient-centered care, and the relevance of PROMs and PREMs in value-based health care make it necessary to introduce new digital health strategies to facilitate the flow of information between health care personnel and the patient. The importance of collecting PROMs and PREMs is growing, and eHealth tools show a high potential as a method for effective, large-scale data collection [28]. However, studies show that response rates are often low, and that the implementation of strategies to promote survey completion can be effective in improving response rates [29]. In our study, most patients perceived the quality-of-life questionnaires as useful. However, further investigation is needed to understand why certain subgroups, such as retired patients, showed lower rates of acceptance regarding PROMs and PREMs surveys.

Strengths and Limitations

The main strengths of this study include its multicenter nature and relevance to the current health care panorama. Identifying the habits and preferences of patients regarding the use of ICT and health care is essential to guarantee the success of telemedicine-based initiatives [30,31]. Characterizing subgroups of patients who find it difficult to access or adopt new technologies is an important first step toward bridging the digital divide by developing strategies to facilitate access and promote the use of eHealth in these populations [32].

This study has certain limitations. First, the lack of follow-up due to the cross-sectional nature of the study, and the fact that patients receiving inpatient cancer care were not included, were limitations of this study. However, the participation of patients with both hematological malignancies and solid tumors ensured a representative sample of outpatients receiving treatment at the 4 centers. Second, although this study examined patients' perspectives on eHealth initiatives such as the Patient Portal, it failed to investigate whether eHealth actually increased patient satisfaction, quality of care, or the overall well-being of the patients. Further studies are needed to address this question. Finally, the fact that the third wave of the COVID-19 pandemic occurred during the recruitment period (February 1, 2021-November 30, 2021), leading to heightened restrictions and disrupted delivery of care, could have influenced some of the participants' answers.

Conclusions

Technology is transforming the health care sector, including cancer care, where the use of telemedicine to ensure interdisciplinary collaboration and to connect patients with

XSL•FO
RenderX

specialist care is widespread. However, notable gaps in the use of eHealth can be observed in certain groups of patients, with significant differences in use due to age, education, and employment status. Strategies to identify subgroups at risk for unequal access to digital health, as well as to facilitate access and use, are warranted.

Acknowledgments

The authors would like to acknowledge the Clinical and Organizational Innovation Unit, the Fundación Jiménez Díaz University Hospital, nurses, caregivers, and especially patients for their contributions to this study.

Funding

This work did not receive specific funding.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: AL-G, PB-A, RC
Investigation: CD-A, BG-D, CM-C, MY-P, CC-S, PL-S
Methodology: AL-G, CM-C, MY-P, PB-A, CC-S
Supervision: AL-G, JGF, PL-S, MdO-R, JS-A, JA-C, RC
Validation: CD-A, BG-D, CM-C, MY-P, PB-A, BP, CC-S
Writing – Original draft: AL-G, BP
Writing – Review and editing (including final draft): BP, JGF, MdO-R, JS-A, JA-C, RC

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey items.
[DOCX File, 16 KB - cancer_v12i1e58823_app1.docx]

Checklist 1
STROBE checklist.
[PDF File, 123 KB - cancer_v12i1e58823_app2.pdf]

References

1. Shirke MM, Shaikh SA, Harky A. Implications of telemedicine in oncology during the COVID-19 pandemic. Acta Biomed 2020 Sep 7;91(3):e2020022. [doi: 10.23750/abm.v91i3.9849] [Medline: 32921719]
2. Kumar R. Impact of Covid-19 on hematological practice: challenges and opportunities. Indian J Hematol Blood Transfus 2020 Oct;36(4):603-604. [doi: 10.1007/s12288-020-01366-w] [Medline: 33071488]
3. Pagano L, Salmanton-García J, Marchesi F, et al. COVID-19 infection in adult patients with hematological malignancies: a European Hematology Association Survey (EPICOVIDEHA). J Hematol Oncol 2021 Oct 14;14(1):168. [doi: 10.1186/s13045-021-01177-0] [Medline: 34649563]
4. Pagano L, Salmanton-García J, Marchesi F, et al. COVID-19 in vaccinated adult patients with hematological malignancies: preliminary results from EPICOVIDEHA. Blood 2022 Mar 10;139(10):1588-1592. [doi: 10.1182/blood.2021014124] [Medline: 34748627]
5. Kumar P, Aggarwal M, Dhawan R, et al. Tele-medicine services in hematological practice during Covid pandemic: its feasibility and difficulties. Indian J Hematol Blood Transfus 2021 Oct;37(4):528-533. [doi: 10.1007/s12288-020-01385-7] [Medline: 33250594]
6. Schreiweis B, Pobiruchin M, Strotbaum V, Suleder J, Wiesner M, Bergh B. Barriers and facilitators to the implementation of eHealth services: systematic literature analysis. J Med Internet Res 2019 Nov 22;21(11):e14197. [doi: 10.2196/14197] [Medline: 31755869]
7. Safi S, Thiessen T, Schmailzl KJ. Acceptance and resistance of new digital technologies in medicine: qualitative study. JMIR Res Protoc 2018 Dec 4;7(12):e11072. [doi: 10.2196/11072] [Medline: 30514693]
8. Lyles CR, Wachter RM, Sarkar U. Focusing on digital health equity. JAMA 2021 Nov 9;326(18):1795-1796. [doi: 10.1001/jama.2021.18459] [Medline: 34677577]

XSL•FO
RenderX

9. Kan K, Heard-Garris N, Bendelow A, et al. Examining access to digital technology by race and ethnicity and child health status among Chicago families. JAMA Netw Open 2022 Aug 1;5(8):e2228992. [doi: 10.1001/jamanetworkopen.2022.28992] [Medline: 36018593]
10. Collado-Borrell R, Escudero-Vilaplana V, Calles A, et al. Oncology patient interest in the use of new technologies to manage their disease: cross-sectional survey. J Med Internet Res 2018 Oct 23;20(10):e11006. [doi: 10.2196/11006] [Medline: 30355554]
11. Báez Gutiérrez N, Rodríguez Ramallo H, Fernández González M, Abdel-Kader Martín L. Smartphone apps for patients with hematologic malignancies: systematic review and evaluation of content. JMIR Mhealth Uhealth 2022 Sep 20;10(9):e35851. [doi: 10.2196/35851] [Medline: 36125860]
12. Burton J, Brook G, McSorley J, Murphy S. The utility of short message service (SMS) texts to remind patients at higher risk of STIs and HIV to reattend for testing: a controlled before and after study. Sex Transm Infect 2014 Feb;90(1):11-13. [doi: 10.1136/sextrans-2013-051228] [Medline: 24064987]
13. Finkelstein SR, Liu N, Jani B, Rosenthal D, Poghosyan L. Appointment reminder systems and patient preferences: patient technology usage and familiarity with other service providers as predictive variables. Health Informatics J 2013 Jun;19(2):79-90. [doi: 10.1177/1460458212458429] [Medline: 23715208]
14. Usher W. Types of social media (Web 2.0) used by Australian allied health professionals to deliver early twenty-first-century practice promotion and health care. Soc Work Health Care 2011;50(4):305-329. [doi: 10.1080/00981389.2010.534317] [Medline: 21512953]
15. Whitelaw S, Mamas MA, Topol E, Van Spall HG. Applications of digital technology in COVID-19 pandemic planning and response. Lancet Digit Health 2020 Aug;2(8):e435-e440. [doi: 10.1016/S2589-7500(20)30142-4] [Medline: 32835201]
16. Golinelli D, Boetto E, Carullo G, Nuzzolese AG, Landini MP, Fantini MP. Adoption of digital technologies in health care during the COVID-19 pandemic: systematic review of early scientific literature. J Med Internet Res 2020 Nov 6;22(11):e22280. [doi: 10.2196/22280] [Medline: 33079693]
17. Yoo HJ, Lee H. Critical role of information and communication technology in nursing during the COVID-19 pandemic: a qualitative study. J Nurs Manag 2022 Nov;30(8):3677-3685. [doi: 10.1111/jonm.13880] [Medline: 36325914]
18. Vargo D, Zhu L, Benwell B, Yan Z. Digital technology use during COVID-19 pandemic: a rapid review. Hum Behav Emerg Technol 2021;3:13-24. [doi: 10.1002/hbe2.242]
19. Shaffer KM, Turner KL, Siwik C, et al. Digital health and telehealth in cancer care: a scoping review of reviews. Lancet Digit Health 2023 May;5(5):e316-e327. [doi: 10.1016/S2589-7500(23)00049-3] [Medline: 37100545]
20. Round T, Sethuraman L, Ashworth M, Purushotham A. Transforming post pandemic cancer services. Br J Cancer 2024 May;130(8):1233-1238. [doi: 10.1038/s41416-024-02596-9] [Medline: 38491174]
21. Xu L, Sanders L, Li K, Chow JC. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR Cancer 2021 Nov 29;7(4):e27850. [doi: 10.2196/27850] [Medline: 34847056]
22. Predmore ZS, Roth E, Breslau J, Fischer SH, Uscher-Pines L. Assessment of patient preferences for telehealth in post-COVID-19 pandemic health care. JAMA Netw Open 2021 Dec 1;4(12):e2136405. [doi: 10.1001/jamanetworkopen.2021.36405] [Medline: 34851400]
23. Mazurka R, Vallis EH, Chen L, et al. Preferences for virtual versus in-person mental and physical healthcare in Canada: a descriptive study from a cohort of youth and their parents enriched for severe mental illness. BMJ Paediatr Open 2024 Jan 8;8(1):e002197. [doi: 10.1136/bmjpo-2023-002197] [Medline: 38191204]
24. Moulaei K, Sheikhtaheri A, Fatehi F, Shanbehzadeh M, Bahaadinbeigy K. Patients' perspectives and preferences toward telemedicine versus in-person visits: a mixed-methods study on 1226 patients. BMC Med Inform Decis Mak 2023 Nov 15;23(1):261. [doi: 10.1186/s12911-023-02348-4] [Medline: 37968639]
25. Chou HH, Chen YH, Lin CT, et al. AI-driven patient support: evaluating the effectiveness of ChatGPT-4 in addressing queries about ovarian cancer compared with healthcare professionals in gynecologic oncology. Support Care Cancer 2025 Apr 1;33(4):337. [doi: 10.1007/s00520-025-09389-7] [Medline: 40167802]
26. Pan A, Musheyev D, Bockelman D, Loeb S, Kabarriti AE. Assessment of artificial intelligence chatbot responses to top searched queries about cancer. JAMA Oncol 2023 Oct 1;9(10):1437-1440. [doi: 10.1001/jamaoncol.2023.2947] [Medline: 37615960]
27. Huo B, Boyle A, Marfo N, et al. Large language models for chatbot health advice studies: a systematic review. JAMA Netw Open 2025 Feb 3;8(2):e2457879. [doi: 10.1001/jamanetworkopen.2024.57879] [Medline: 39903463]
28. del Olmo Rodríguez M, Córdoba R, Gómez-Meana Á, et al. Implementing a broad digital framework to drive network strategy through PROMs and PREMs. NEJM Catal Innov Care Deliv 2023 Jul 19;4(8). [doi: 10.1056/CAT.23.0083]
29. Stover AM, Haverman L, van Oers HA, Greenhalgh J, Potter CM, ISOQOL PROMs/PREMs in Clinical Practice Implementation Science Work Group. Using an implementation science approach to implement and evaluate patient-reported outcome measures (PROM) initiatives in routine care settings. Qual Life Res 2021 Nov;30(11):3015-3033. [doi: 10.1007/s11136-020-02564-9] [Medline: 32651805]
30. Drewes C, Kirkovits T, Schiltz D, et al. eHealth acceptance and new media preferences for therapy assistance among breast cancer patients. JMIR Cancer 2016 Sep 14;2(2):e13. [doi: 10.2196/cancer.5711] [Medline: 28410189]

XSL•FO
RenderX

31. Welch BM, Harvey J, O'Connell NS, McElligott JT. Patient preferences for direct-to-consumer telemedicine services: a nationwide survey. BMC Health Serv Res 2017 Nov 28;17(1):784. [doi: 10.1186/s12913-017-2744-8] [Medline: 29183372]
32. Ware P, Bartlett SJ, Paré G, et al. Using eHealth technologies: interests, preferences, and concerns of older adults. Interact J Med Res 2017 Mar 23;6(1):e3. [doi: 10.2196/ijmr.4447] [Medline: 28336506]

Abbreviations

ICT: information and communication technology
OR: odds ratio
PREM: patient-reported experience measurement
PROM: patient-reported outcome measurement

Edited by N Cahill; submitted 26.Mar.2024; peer-reviewed by A Palanica, S Calik; revised version received 09.Jul.2025; accepted 09.Jul.2025; published 13.Jan.2026.

Please cite as:
Lopez-Garcia A, Diaz-Aizpun C, Gallego-Diaz B, Miranda-Castillo C, Yuste-Platero M, Beltran-Alvarez P, Carames-Sanchez C, Garcia-Foncillas J, Llamas-Sillero P, del Olmo-Rodriguez M, Short-Apellaniz J, Pfang B, Arcos-Campillo J, Cordoba R
New Technologies and Digital Health Tools in Patients With Solid Tumors and Hematological Malignancies: Cross-Sectional Multicenter Survey Study
JMIR Cancer 2026;12:e58823
URL: https://cancer.jmir.org/2026/1/e58823
doi:10.2196/58823

© Alberto Lopez-Garcia, Carola Diaz-Aizpun, Beatriz Gallego-Diaz, Carolina Miranda-Castillo, Maria Yuste-Platero, Pilar Beltran-Alvarez, Cristina Carames-Sanchez, Jesus Garcia-Foncillas, Pilar Llamas-Sillero, Marta del Olmo-Rodriguez, Jorge Short-Apellaniz, Bernadette Pfang, Javier Arcos-Campillo, Raul Cordoba. Originally published in JMIR Cancer (https://cancer.jmir.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of GPT-5 for Esophageal Cancer Staging Using Fluorodeoxyglucose Positron Emission Tomography Maximum-Intensity Projection Images: Comparative Pilot Study

Hiroki Maruyama[1], MD; Yoshitaka Toyama[2,3], MD, PhD; Yuya Araki[4]; Kentaro Takanami[3], MD, PhD; Masato Ito[3,5], MD; Yumi Nakajima[3,6], MD; Kei Takase[3], MD, PhD; Takashi Kamei[1], MD, PhD

[1]Department of Surgery, Graduate School of Medicine, Tohoku University, Sendai, Japan

[2]Department of Imaging and Anatomy for Groundbreaking Education Collaborative Research, Graduate School of Medicine, Tohoku University, Sendai, Japan

[3]Department of Diagnostic Radiology, Tohoku University Hospital, Sendai, Japan

[4]School of Medicine, Tohoku University, Sendai, Japan

[5]Department of Diagnostic Radiology, Osaki Citizen Hospital, Osaki, Japan

[6]Department of Diagnostic Radiology, Tohoku Medical and Pharmaceutical University, Sendai, Japan

Corresponding Author:
Yoshitaka Toyama, MD, PhD
Department of Imaging and Anatomy for Groundbreaking Education Collaborative Research
Graduate School of Medicine
Tohoku University
2-1 Seiryo-Machi, Aoba-Ku
Sendai,
Japan
Phone: 81 022 717 7312
Fax: 81 022 717 7316
Email: ytoyama0818@gmail.com

Abstract

Background: Accurate esophageal cancer staging relies on ^{18}F fluorodeoxyglucose positron emission tomography (^{18}F FDG-PET), but its interpretation is complex and time-intensive. This diagnostic burden is exacerbated by significant workforce shortages in both radiology and surgery, thus necessitating automated support systems. The emergence of advanced large language models (LLMs) has raised expectations for their potential to fulfill this role in complex medical tasks.

Objective: We evaluated the diagnostic accuracy of LLMs for staging esophageal cancer using ^{18}F FDG-PET images, with a focus on their ability to assess lymph nodes (LNs; clinical N [cN]) and distant metastases (clinical M [cM]) for automated radiology reporting.

Methods: This retrospective study included 120 consecutive adult patients who were diagnosed with esophageal squamous cell carcinoma and underwent ^{18}F FDG-PET/computed tomography at Tohoku University Hospital between January 2019 and December 2021. Patients with prior treatment, nonsquamous cell carcinoma histology, or blood glucose levels ≥200 mg/dL were excluded. Frontal maximum-intensity projection positron emission tomography images were extracted, standardized, and analyzed along with information regarding the tumor location. Six LLMs (GPT-5, GPT-4.5, GPT-4.1, OpenAI-o3, -o1, and GPT-4 Turbo) and 4 blinded human evaluators (a nuclear medicine specialist, a gastrointestinal surgeon, and 2 radiology residents) assessed the presence of thoracic and abdominal LN metastases on a region-level basis and determined cN and cM staging on a patient-level basis. The model analyses were performed using the application programming interface in a zero-shot setting. Radiology reports served as the reference standard. Diagnostic agreement and accuracy were evaluated using Cohen κ and the Cochran Q test. Additionally, to account for the class imbalance in the dataset, the Matthews Correlation Coefficient was calculated as a robust metric for binary classification performance. Post hoc McNemar tests were performed with Bonferroni correction; statistical significance for pairwise comparisons was set at $P<.0083$ (adjusted from $P<.05$) using JMP Pro (version 18.0; SAS Institute Inc).

Results: The average accuracy was 41/120 (34%) to 94/120 (78%) for LLMs and 72/120 (60%) to 102/120 (85%) for physicians, with significantly higher accuracy for physicians ($P<.05$) in the thoracic LN, abdominal LN, and cN stages. Interrater reliability

XSL•FO
RenderX

was slight to fair for LLMs (κ: –0.07 to 0.25) and fair to substantial for physicians (κ: 0.27 to 0.74). Matthews Correlation Coefficient scores were consistently higher for physicians (0.28 to 0.75) than for LLMs (–0.07 to 0.32). Among the LLMs, GPT-5 demonstrated the highest overall accuracy, with newer LLMs showing improved diagnostic accuracy when compared with previous models in identifying abdominal LN metastases and cM staging, though they showed weaker consistency for cN staging. For example, in thoracic LN detection, GPT-5 achieved 76/120 (63%) accuracy, whereas other LLMs achieved 72/120 (60%) or lower accuracy.

Conclusions: Although current LLMs have not yet reached physician-level accuracy in comprehensive staging, recent models show promise in assisting with specific diagnostic tasks.

(JMIR Cancer 2026;12:e86630) doi:10.2196/86630

KEYWORDS

generative artificial intelligence; large language models; LLMs; 18F FDG-PET imaging; fluorodeoxyglucose positron emission tomography; esophageal cancer staging; radiology report automation

Introduction

Esophageal cancer remains one of the most challenging malignancies to manage. According to the most recent GLOBOCAN statistics and Japanese cancer registry data, it poses a significant global health burden [1,2]. Management requires multidisciplinary expertise that spans complex surgical procedures, perioperative care, and advanced imaging interpretation. Esophagectomy is among the most invasive oncologic surgeries, and optimal patient outcomes depend on accurate staging, meticulous operative planning, and coordinated care.

Surgical services face a mounting workforce shortage: the Association of American Medical Colleges 2024 national projection estimates a shortfall of 10,100 to 19,900 surgeons by 2036 [3], and a nationwide Japanese survey reported that over half of teaching hospitals already experience surgeon shortages—even in densely populated prefectures [4]. In parallel, radiology faces both workforce shortages and escalating workload: the 2023 Workforce Census for the United Kingdom reports a 30% shortfall of clinical radiologists [5], while the volume of image data per study has surged markedly, compounding reporting demands [6,7].

Against this backdrop, fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F FDG-PET/CT)—a cornerstone of preoperative staging in esophageal cancer—is notably time-consuming and complex to interpret [8,9], requiring integration of functional and anatomical information. This adds to the workload of both surgeons, who must incorporate imaging findings into surgical planning, and radiologists, who must provide comprehensive and timely reports for multidisciplinary decision-making.

International guidelines, including the American College of Radiology Appropriateness Criteria and the European Society for Medical Oncology recommendations, endorse ^{18}F FDG-PET/CT for baseline staging and selected follow-up in esophageal cancer [10,11]. ^{18}F FDG-PET/CT is recognized not only for its diagnostic utility in detecting distant metastases and assessing nodal involvement but also for its significant prognostic value in oncology, as metabolic parameters often correlate with patient outcomes [12,13].

Generative artificial intelligence (AI), a subset of AI capable of creating new content, has revolutionized various fields. Within this domain, large language models (LLMs) are deep learning algorithms trained on massive datasets to understand and generate human-like text. Recently, the evolution of these models into multimodal large language models, which can process and interpret both text and images simultaneously, has expanded their potential applications in health care. While previous research has demonstrated the utility of AI in medical tasks such as summarizing radiology reports or passing medical licensing examinations [14], the application of general-purpose multimodal large language models to complex image interpretation remains limited. Most prior studies have focused on anatomical imaging modalities like plain radiography or CT for simple classification tasks [15-18]. There is a paucity of research evaluating whether these models can perform high-level clinical reasoning, specifically TNM staging based on functional nuclear medicine imaging (^{18}F FDG-PET). Against this backdrop, and with the release of GPT-5, the latest publicly available LLM from OpenAI, we report the first evaluation of the medical image interpretation capabilities of this model. We compared its diagnostic accuracy in esophageal cancer staging with that of physicians and other state-of-the-art models.

Methods

Study Design

We adhered to the guidelines outlined in the checklist for AI in medical imaging to ensure methodological transparency and ethical rigor [19].

Ethical Considerations

This retrospective study conformed to the ethical standards of the Declaration of Helsinki (1975, as revised in 2013) and was approved by the Institutional Review Board of Tohoku University Hospital (approval number 2024-1-816). The Institutional Review Board explicitly approved the transfer of deidentified patient image data to the third-party commercial servers used by the application programming interfaces (APIs). The requirement for individual informed consent was waived, and patients were informed regarding the study via an opt-out method on the hospital website. To protect patient privacy and confidentiality, all data used for analysis were anonymized; the correspondence table linking study IDs to personal information

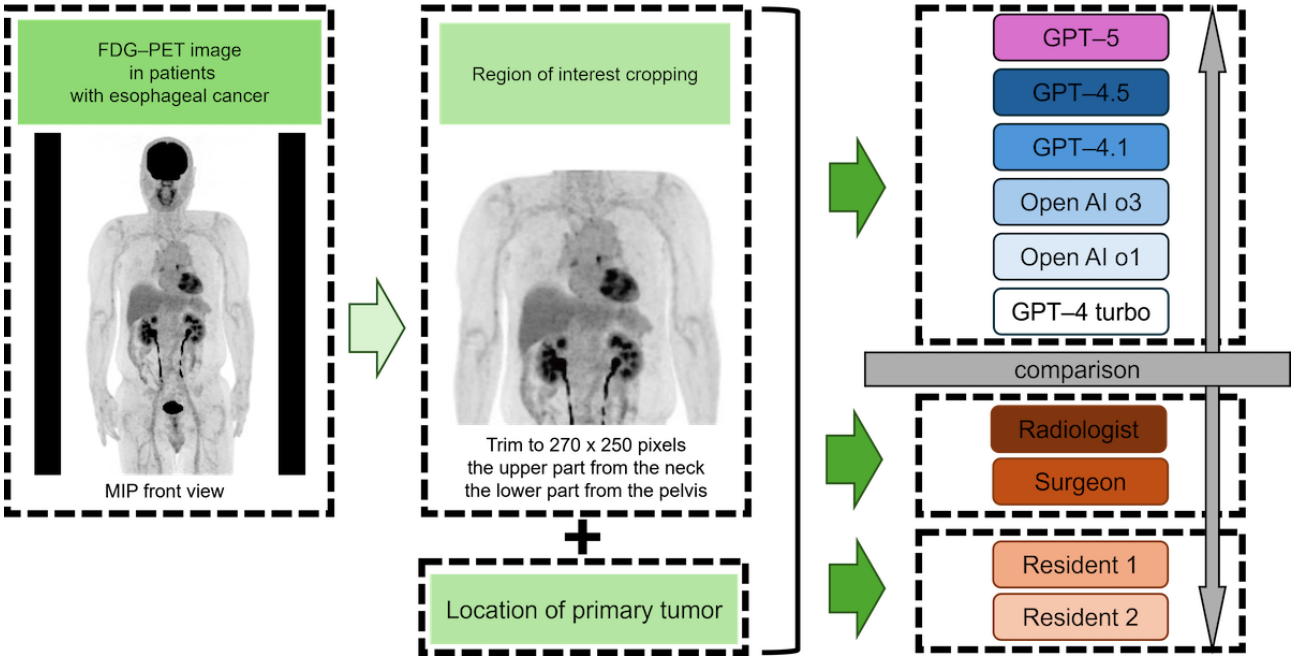
XSL•FO
RenderX

was stored separately in a secure location restricted to authorized personnel. Prior to uploading, all images were fully deidentified by removing all DICOM metadata and converting them to JPEG format. As this study involved the secondary use of existing data, no compensation was provided to the participants. Furthermore, all images included in the manuscript and supplementary materials were carefully cropped to remove any personally identifiable information, such as patient ID, name, age, and sex, to ensure that individual participants cannot be identified.

Patients

The cases analyzed in this study were derived from a prospective, continuously registered cohort of patients who were hospitalized and treated at our institution between January 2019 and December 2021. All participants underwent upper gastrointestinal endoscopy and were diagnosed with esophageal cancer that was confirmed via biopsy. The patients were eligible for inclusion if they were 18 years of age or older, had undergone their first positron emission tomography (PET)/CT examination using a GE scanner at our hospital, and had a biopsy-confirmed diagnosis of squamous cell carcinoma (SCC). Patients were excluded if they had a history of treatment for esophageal cancer, a histological type other than SCC, such as adenocarcinoma, or a pre-examination blood glucose level of ≥200 mg/dL. The overall study design is shown in Figure 1.

Figure 1. Research workflow for comparing large language models with physicians in the staging of esophageal cancer. This flowchart illustrates the study design for evaluating the performance of GPT-5, GPT-4.5, GPT-4.1, OpenAI-o3, OpenAI-o1, and GPT-4turbo in comparison with physicians using MIP images from 18F FDG-PET. The workflow includes key steps, such as acquiring MIP frontal images, cropping regions of interest, analyzing these images and tumor location data using large language models, and assessing their performance relative to that of physicians. FDG-PET: fluorodeoxyglucose positron emission tomography; MIP: maximum intensity projection.



PET Imaging and Interpretation

^{18}F FDG-PET/CT was performed at our institution by using a silicon photomultiplier PET scanner (Discovery MI; GE Healthcare). Patients were instructed to fast for at least 6 h prior to the ^{18}F FDG injection. A rapid intravenous injection of ^{18}F FDG, at a dose of approximately 3.7 MBq/kg, was administered through the right antecubital vein. After a 60-minute uptake period, patients underwent a low-dose CT scan (140 kV; automatic exposure control, 20-80 mA), followed by whole-body PET/CT imaging.

All PET/CT images were interpreted by board-certified radiologists who had specialized in both radiology and nuclear medicine and had been accredited by the Japanese Society of Radiology and the Japanese Society of Nuclear Medicine. Interpretations were made in the context of available clinical information and correlative imaging studies, such as contrast-enhanced CT. Their interpretation reports served as the gold standard for this study, which evaluated whether LLMs can generate radiology reports. The interpretation reports were documented in accordance with the 8th edition of the Union for International Cancer Control staging system, and ensured that the N and M classifications, in addition to the T classification, were defined [20].

Image Selection and Data Preparation

As a single maximum intensity projection (MIP) image extracted from a PET scan provides information about the whole body, it has been suggested that AI diagnosis could reduce the data load [21]. Therefore, we chose to analyze the MIP, as it offers a comprehensive view of metabolic activity across the entire body while simplifying data processing and interpretation. MIP frontal images were extracted from the acquired DICOM-format PET/CT images and subsequently converted to the JPEG format. During this process, the regions from the neck up and pelvis down were cropped to exclude common physiological accumulations in the oral cavity and bladder. The decision to

XSL•FO
RenderX

crop regions distal to the pelvis was further supported by the finding that no bone metastases were identified in these areas within our dataset. The original images (563 × 710 pixels) were cropped to a fixed size of 270 × 250 pixels for standardization.

Data management and entry were performed using Microsoft Excel (Microsoft Corp). Based on the radiological interpretation reports (reference standard), the following patient variables were collected: age, sex, primary tumor location, presence of thoracic lymph node (LN) metastasis, presence of abdominal LN metastasis, clinical N stage (cN), clinical M stage (cM), clinical stage, and treatment modality.

The analysis was conducted at the patient level for determining the cN and cM stages and at the region level for assessing the presence of metastasis in the thoracic and abdominal fields. This unit of analysis was selected to align with clinical decision-making processes, where the overall stage and regional involvement dictate the treatment strategy, rather than the precise counting of individual LNs. All staging was performed in accordance with the 8th edition of the Union for International Cancer Control TNM classification.

The location of the primary tumor, which was used as an input for GPT, was determined by a specialist in gastrointestinal surgery. This classification was based on a comprehensive review of upper gastrointestinal endoscopy, fluoroscopy, and CT images based on the guidelines outlined in the Japanese Classification of Esophageal Cancer, 12th Edition [22]. The cervical and upper thoracic esophagus were classified as the upper region, the middle thoracic esophagus as the middle region, and the lower thoracic esophagus and esophagogastric junction as the lower region, to generate a 3-tier classification.

LLM-Based Analysis

The selection of LLMs for this study was restricted to the GPT series developed by OpenAI, as these models are currently the most widely used generative AI platforms globally and provide a robust API that facilitates seamless multimodal data input. In this study, 6 LLMs were used for analysis: GPT-5, GPT-4.5, GPT-4.1, OpenAI-o3, OpenAI-o1, and GPT-4 Turbo (OpenAI). GPT-4o was excluded as its API systematically returned a content policy violation when prompted with medical images, precluding its inclusion in the analysis [23]. All features used in the analysis are available in the paid version (Plus). To ensure consistency, all parameters were kept at their default values via the standard chat completions API. Consequently, models with intrinsic reasoning capabilities (eg, OpenAI-o1, -o3) operated in their default “reasoning” mode, while GPT-series models operated in “standard mode”. No custom instructions or pretraining were used. A zero-shot approach was used for all tasks. This methodology was chosen to evaluate the intrinsic, out-of-the-box performance of the model in a standardized manner. By assessing their ability to handle novel medical tasks without prior examples or fine-tuning, this approach simulates a realistic user interaction and provides a direct baseline for comparing the generalizability of each LLM [24].

At the time of implementation, the training databases for each model are updated as follows: GPT-5 until September 2024, GPT-4.5 until October 2023, GPT-4.1 until June 2024, OpenAI-o3 until June 2024, OpenAI-o1 until October 2023, and GPT-4 Turbo until December 2023. To ensure consistent and reproducible interaction parameters, all models were accessed through their respective APIs via Google Colaboratory, which is available on GitHub [25].

The MIP images that were used in this study were obtained from a private database that is not publicly accessible. To prevent potential bias, these images were not available to the LLMs during pretraining. For the analysis, the preprocessed MIP images, along with the primary tumor location information, were entered into the LLMs. Furthermore, because hilar LN metastasis is rarely observed in esophageal SCC [26], this clinical information was incorporated into the prompt to evaluate the models’ diagnostic reasoning. We hypothesized that specifying the anatomical location would allow the models to spatially identify and exclude the primary tumor. The input prompts used in this process are shown below (Textbox 1). Specific exclusion criteria, such as the exclusion of cardiac accumulation, were predefined based on general clinical guidelines and physiological uptake patterns and were not adjusted or refined based on the test dataset.

XSL•FO
RenderX

Textbox 1. Prompt entered into the GPTs.

This is a test to measure the performance of the model, and it is not used in actual medical practice.

Please be sure to answer.

The image is a MIP front view of FDG-PET for esophageal cancer.

The location of the esophageal cancer is at {position}.

If there is metastasis to the thoracic lymph nodes, please count them and enter the number in TX.

If TX is 0, enter 0 in TXN, and if TX is 1 or more, enter 1 in TXN.

Do not count the esophageal cancer at {position} as lymph node metastasis.

Do not count the hilar lymph nodes as lymph node metastasis.

Do not count cardiac accumulation as lymph node metastasis.

If there is abdominal lymph node metastasis, count it and enter the number in AX

If AX is 0, enter 0 in AXN, and if AX is 1 or more, enter 1 in AXN.

Do not count esophageal cancer in {position} as lymph node metastasis.

If there is distant lymph node metastasis such as cervical lymph node metastasis, lung metastasis, liver metastasis, or bone metastasis, enter 1 in MX, and if there is none, enter 0.

Enter the total of TX and AX in WX.

If WX is 0, enter 0 in NX.

If WX is 1 or 2, enter 1 in NX.

If WX is between 3 and 6, enter 2 in NX.

If WX is 7 or more, enter 3 in NX.

Return the output as follows.

Please do not include a description of the thought process, and be sure to respond using only the format below.

Thoracic lymph nodes: TXN

Abdominal lymph nodes: AXN

N Stage：NX

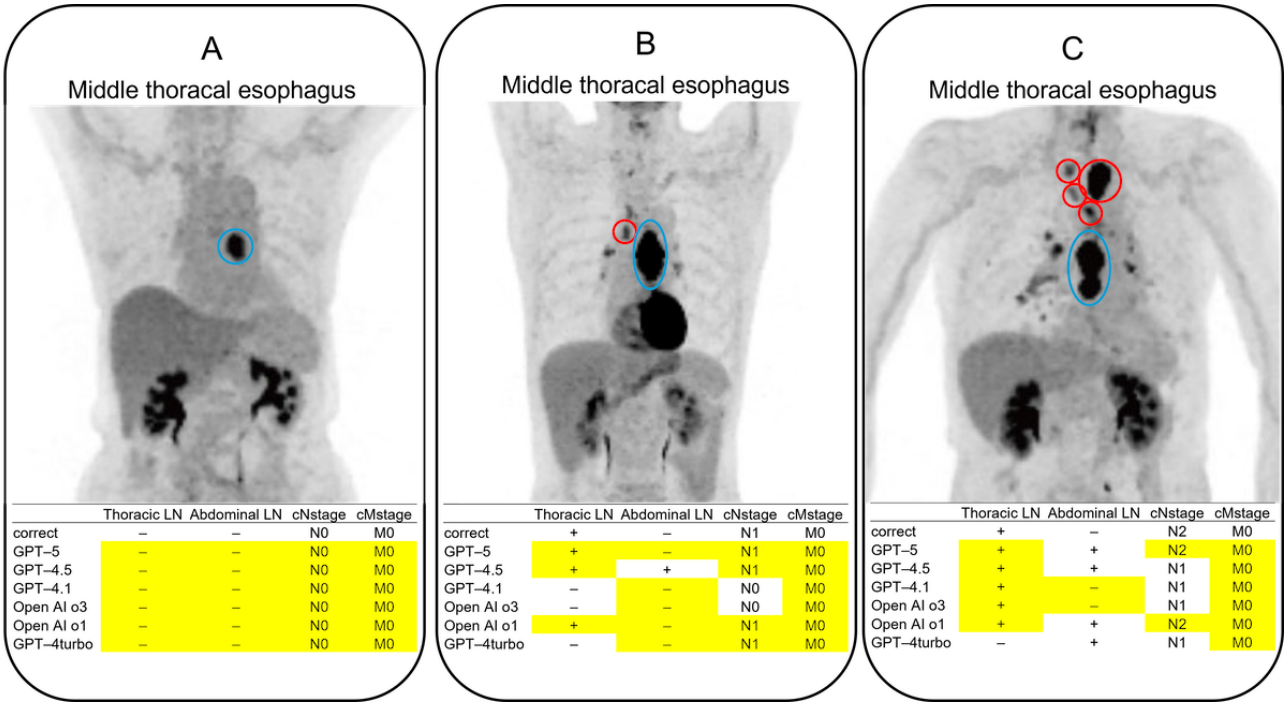
M Stage：MX

The LLMs analyzed the MIP images and provided staging-related assessments of esophageal cancer. Research using GPT-5 was analyzed on August 11, 2025. Research analysis using GPT-4.5, OpenAI-o1, and GPT-4 turbo was conducted on March 23, 2025. Research using GPT-4.1 and OpenAI-o3 was analyzed on May 2, 2025.

To further assess the textual consistency of the model's outputs and address the "black box" limitation, we performed a post hoc qualitative subanalysis on the 3 representative cases (shown in Figure 2) on December 9, 2025. For this analysis, the prompt was modified to include the following instruction: "Please state the basis for reaching that diagnosis." This enabled us to examine whether the model's generated explanation aligned with clinical features, although it does not guarantee that the model visually attended to them.

XSL•FO
RenderX

Figure 2. Examples of input images and responses of GPT-4.5, GPT-4.1, and OpenAI-o3 in cases of esophageal cancer. The primary tumor site indicated in the radiology report is shown as a blue circle, and the metastatic LNs are shown as red circles. Note that these colored circles were manually overlaid by the authors to visualize the ground truth and were not generated by the AI models. The yellow cells indicate the correct answers (agreement with the ground truth). (A) All the models correctly identified the absence of LN and distant metastases beyond the primary lesion. (B) A case with a single metastatic thoracic LN. Only GPT-5 and OpenAI-o1 provided a correct evaluation, identifying thoracic LN metastasis, no abdominal LN metastasis, and the correct cN and cM stages. Other models either failed to identify the thoracic LN metastasis or misdiagnosed abdominal LN metastasis as positive. (C) A cN-stage 2 case with thoracic LN metastasis. 18F FDG (fluorodeoxyglucose) accumulation in the hilar LNs was interpreted as nonspecific accumulation in the radiology report. GPT-5 correctly identified the cN stage but misdiagnosed abdominal LN metastasis as positive. Although other models correctly identified thoracic LN metastasis, many incorrectly stated the disease as N1. cM: clinical M; cN: clinical N; LN: lymph nodes.



	Thoracic LN	Abdominal LN	cNstage	cMstage
correct	–	–	N0	M0
GPT–5	–	–	N0	M0
GPT–4.5	–	–	N0	M0
GPT–4.1	–	–	N0	M0
Open AI o3	–	–	N0	M0
Open AI o1	–	–	N0	M0
GPT–4turbo	–	–	N0	M0

	Thoracic LN	Abdominal LN	cNstage	cMstage
correct	+	–	N1	M0
GPT–5	+	–	N1	M0
GPT–4.5	+	+	N1	M0
GPT–4.1	–	–	N0	M0
Open AI o3	–	–	N0	M0
Open AI o1	+	–	N1	M0
GPT–4turbo	–	–	N1	M0

	Thoracic LN	Abdominal LN	cNstage	cMstage
correct	+	–	N2	M0
GPT–5	+	+	N2	M0
GPT–4.5	+	+	N1	M0
GPT–4.1	+	–	N1	M0
Open AI o3	+	–	N1	M0
Open AI o1	+	+	N2	M0
GPT–4turbo	–	+	N1	M0

Physician's Evaluation

The same information that was provided to the LLMs, including the cropped MIP images and primary tumor location for esophageal cancer, was presented to 4 human evaluators: a nuclear medicine specialist with 14 years of experience, a gastrointestinal surgeon with 9 years of experience, and 2 radiology residents. To prevent bias, the evaluators were blinded to the contents of the diagnostic report. Using the same criteria that were applied by the LLMs, each evaluator independently assessed the images and determined the presence or absence of thoracic LN metastases, abdominal LN metastases, and cN and cM stages. The evaluators were not involved in the diagnosis or treatment of the included patients. To ensure parity with the AI input (which included tumor location prompts), evaluators were provided with the tumor location information but were strictly blinded to all other clinical data, including patient history, reference radiology reports, and pathological outcomes.

Statistical Analysis

The primary outcome was diagnostic accuracy, defined as the concordance with the reference standard. The secondary outcome was interrater reliability assessed using Cohen κ. The CIs for each rater's diagnostic performance were calculated using the Wilson score interval (without continuity correction). Cohen κ consistency analysis was used to assess the agreement between the LLMs, physicians, and actual diagnostic reports. Additionally, for binary classification tasks (assessment of thoracic LN metastasis, abdominal LN metastasis, and cM stage), the Matthews Correlation Coefficient (MCC) was calculated. MCC is considered a robust metric for imbalanced datasets, as it incorporates true and false positives and negatives, returning a high score only when the prediction performs well across all confusion matrix categories [27]. The κ values were interpreted according to the following scale: 0-0.2 (poor agreement), 0.2-0.4 (fair agreement), 0.4-0.6 (moderate agreement), 0.6-0.8 (substantial agreement), and 0.8-1.0 (almost perfect agreement) [28]. Student *t* test and Cochran Q test were used to compare the rates of diagnostic accuracy between LLMs and physicians, followed by the post hoc McNemar test [29]. Data were analyzed using JMP Pro (version 18.0; SAS Institute Inc). Only the post hoc McNemar test was corrected using the Holm-Bonferroni correction to adjust the *P* value <.0083; a *P* value <.05 was considered statistically significant for all analyses.

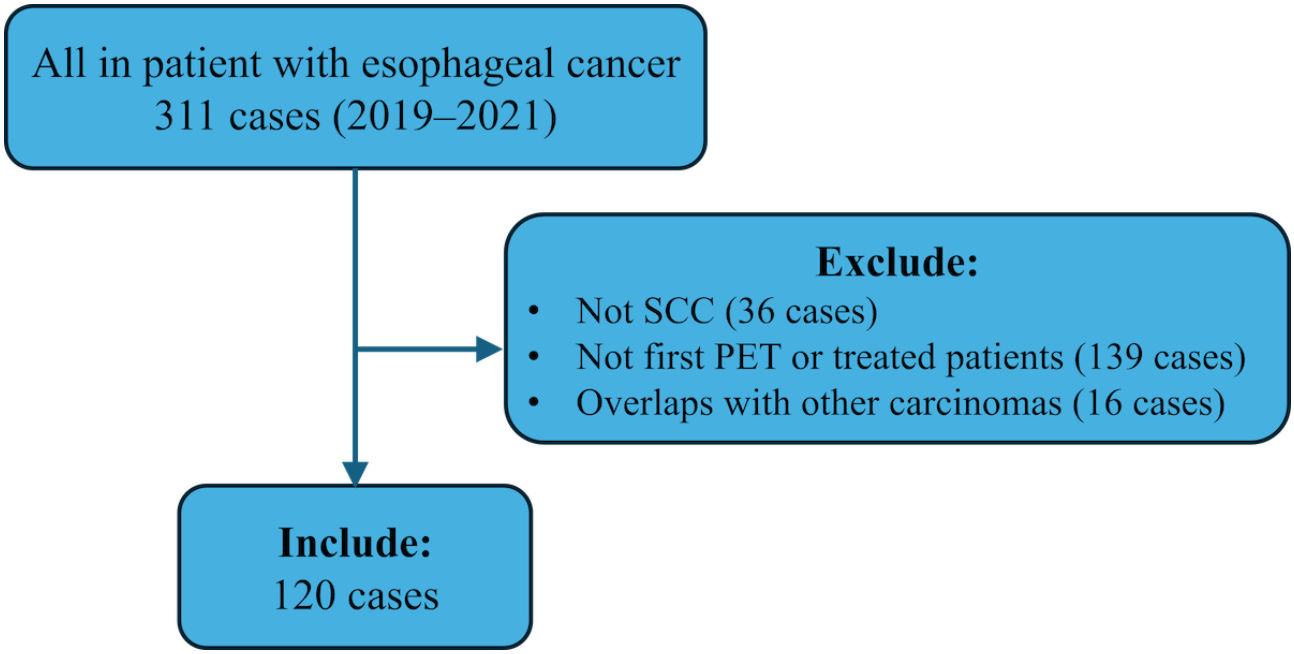
Results

Baseline Characteristics of the Study Population

Of the 311 patients with esophageal cancer who were admitted to our department, 36 were excluded because of a histological type other than SCC, 139 had already received treatment or were not undergoing their first PET/CT scan, and 16 had other carcinomas. Thus, 120 patients were included in this study (Figure 3).

XSL•FO
RenderX

Figure 3. Case inclusion and exclusion flowchart. PET: positron emission tomography.



Within this cohort, 120 primary esophageal cancer lesions were identified and analyzed. Histopathologically, all cases (100%) were confirmed as SCC. The study population comprised 120 patients (median age 71 years; 25/120 women, 20.8%); 58/120 (48.3%) patients had thoracic LN metastasis, and 35/120 (29.2%) had abdominal LN metastasis. The cN stage was N0 in 45/120 (37.5%) patients, N1 in 52/120 (43.3%) patients, N2 in 22/120 (18.3%) patients, and N3 in 1/120 (0.8%) patients. The cM stage was M1 in 27/120 (22.5%) patients. The detailed data are presented in Table 1.

XSL•FO
RenderX

Table 1. Patient characteristics. All patients included had squamous cell carcinoma.

Characteristics	Values (n=120)
Age (years), median (range)	71 (44-89)
Sex, n (%)	
Male	95 (79.2)
Female	25 (20.8)
Location, n (%)	
CeUt[ab]	16 (13.3)
Mt[c]	61 (50.8)
LtJz[de]	43 (35.8)
Location of LN[f] metastasis, n (%)	
Thoracic LN	58 (48.3)
Abdominal LN	35 (29.2)
cN-stage[g], n (%)	
N0	45 (37.5)
N1	52 (43.3)
N2	22 (18.3)
N3	1 (0.83)
cM-stage[h] n (%)	
M0	93 (77.5)
M1	27 (22.5)
cStage[i], n (%)	
I	17 (15.2)
II/III	63 (52.5)
IV	40 (33.3)
Treatment, n (%)	
Operation	80 (66.7)
chemotherapy/radiation therapy	36 (30)
BSC[j]	4 (3.3)

[a]Ce: cervical esophagus.
[b]Ut: upper thoracic esophagus.
[c]Mt: middle thoracic esophagus.
[d]Lt: lower thoracic esophagus.
[e]Jz: zone of the esophagogastric junction.
[f]LN: lymph node.
[g]cN: clinical N.
[h]cM: clinical M.
[i]cStage: clinical stage.
[j]BSC: best supportive care.

Examples of MIP Images and LLM Responses

Figure 2 presents representative MIP images that were entered into GPT-5, GPT-4.5, GPT-4.1, and OpenAI-o3, along with their corresponding diagnostic outputs. All patients had middle thoracic esophageal cancer.

Case A involved a patient without LN or distant metastases outside the primary lesion. All 3 LLMs correctly identified the absence of LN and distant metastases. In Case B, which featured a patient with a single metastatic thoracic LN, only GPT-5 and OpenAI-o1 provided a fully correct evaluation. These models accurately identified the thoracic LN metastasis, correctly

XSL•FO
RenderX

reported the absence of abdominal LN metastasis, and determined the proper cN and cM stages. The other models either failed to detect the thoracic metastasis or incorrectly identified abdominal LN involvement. Case C presented a patient with cN-stage 2 thoracic LN metastasis. In this instance, GPT-5 correctly identified the cN stage but misdiagnosed abdominal LN metastasis as positive. While the other models correctly detected the presence of thoracic LN metastasis, they failed to determine the correct stage, with many classifying the disease as N1.

Qualitative Assessment of Generated Rationale

The results of the reasoning verification subanalysis conducted for the 3 cases shown in Figure 2 are summarized in Table S1 in Multimedia Appendix 1.

In Case A, GPT-5 provided a correct diagnosis with a rationale that explicitly mentioned the exclusion of cardiac and hilar uptake, consistent with the instructions.

In Case B, although GPT-5 had correctly identified the thoracic LN metastasis in the primary analysis, the model failed to detect the lesion in the subanalysis (False Negative), stating "No additional discrete FDG-avid mediastinal nodal foci." The text output suggested that the model actively evaluated and excluded the hilar region; however, this inconsistency highlights the stochastic nature of LLMs, where minor prompt alterations (eg, adding a request for reasoning) can alter the diagnostic outcome.

In Case C, the model correctly identified the N2 stage, but the reasoning revealed a discrepancy. It correctly excluded hilar uptake in its explanation, but hallucinated an abdominal LN metastasis (False Positive). This suggests that although the model can generate text that appears to apply exclusion criteria, it may still misidentify physiological uptake or noise as pathological lesions, thereby reaching the correct stage for the wrong anatomical reason.

Overall Diagnostic Performance of GPTs and Physicians

The correct response rate, sensitivity, specificity, and Cohen κ coefficient for each parameter are presented for both the LLMs and physicians. The overall diagnostic performance is summarized in Tables 2-5. In the overall correct response rate, LLMs achieved a rate of 41/120 (34%; 95% CI 26%-43%) to 94/120 (78%; 95% CI 70%-86%), whereas physicians demonstrated a higher rate of 70/120 (58%; 95% CI 49%-67%) to 108/120 (90%; 95% CI 83%-94%). The correct response rate of LLMs for the thoracic and abdominal LN ranged from 60/120 (50%; 95% CI 40%-59%) to 87/120 (73%; 95% CI 64%-80%). The sensitivity of LLMs ranged from 7/120 (6%; 95% CI 0%-14%) to 112/120 (93%; 95% CI 86%-100%), whereas that of physicians ranged from 65/120 (54%; 95% CI 37%-72%) to 104/120 (87%; 95% CI 77%-94%). The specificity was 12/120 (10%; 95% CI 2%-17%) to 115/120 (96%; 95% CI 90%-99%) for LLMs and 87/120 (73%; 95% CI 61%-84%) to 119/120 (99%; 95% CI 94%-100%) for physicians. For the cN stage, the correct response rate was 41/120 (34%; 95% CI 26%-43%) to 58/120 (48%; 95% CI 40%-57%) for LLMs and 70/120 (58%; 95% CI 49%-67%) to 73/120 (61%; 95% CI 52%-70%) for physicians. For the cM stage, the correct response rate ranged from 91/120 (76%; 95% CI 68%-84%) to 102/120 (85%; 95% CI 79%-92%) for both LLMs and physicians. The sensitivity was 0/120 (0%; 95% CI 0%-0%) to 18/120 (15%; 95% CI 0.5%-29%) for LLMs and 40/120 (33%; 95% CI 14%-52%) to 67/120 (56%; 95% CI 37%-72%) for physicians. The specificity for both groups was 100/120 (83%; 95% CI 75%-91%) to 119/120 (99%; 95% CI 97%-100%). In terms of MCC, which adjusts for class imbalance, physicians consistently outperformed LLMs. For example, in the assessment of thoracic LN metastasis, the radiologist achieved an MCC of 0.573, whereas the highest-performing LLM (GPT-5) reached only 0.317.

Table 2. Overall diagnostic performance of GPTs and physicians for thoracic lymph nodes.

	Accuracy (%) (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Cohen κ value	*P* value	Matthews correlation coefficient
GPT-5	63 (54-71)	31 (21-44)	94 (84-97)	0.25	<.01	0.32
GPT-4.5	60 (51-68)	35 (22-47)	84 (75-93)	0.19	<.01	0.21
GPT-4.1	58 (49-66)	28 (18-40)	85 (75-92)	0.13	<.01	0.16
OpenAI-o3	56 (47-64)	22 (14-35)	87 (77-93)	0.097	—[a]	0.13
OpenAI-o1	50 (40-59)	93 (86-100)	10 (2-17)	0.027	—	0.05
GPT-turbo	52 (43-61)	20 (9-31)	82 (73-92)	0.03	—	0.04
Radiologist	78 (71-86)	84 (75-94)	73 (61-84)	0.57	<.001	0.57
Surgeon	74 (66-82)	60 (47-73)	87 (79-96)	0.48	<.001	0.49
Radiology resident 1	74 (66-82)	71 (59-83)	77 (67-88)	0.48	<.001	0.48
Radiology resident 2	80 (72-96)	87 (77-94)	73 (62-83)	0.60	<.001	0.59

[a]Not applicable.

XSL•FO
RenderX

Table 3. Overall diagnostic performance of GPTs and physicians for abdominal lymph nodes.

	Accuracy (%) (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Cohen κ value	*P* value	Matthews correlation coefficient
GPT-5	73 (64-80)	14 (6-29)	96 (90-99)	0.14	<.01	0.20
GPT-4.5	69 (61-77)	34 (18-51)	82 (74-91)	0.18	<.01	0.18
GPT-4.1	71 (62-78)	17 (8-33)	93 (85-97)	0.13	<.01	0.15
OpenAI-o3	71 (62-78)	9 (3-22)	96 (90-99)	0.067	—[a]	0.11
OpenAI-o1	56 (47-65)	63 (46-80)	53 (42-64)	0.081	—	0.14
GPT-turbo	66 (57-74)	6 (0-14)	91 (84-97)	0.063	—	–0.06
Radiologist	80 (73-87)	57 (40-74)	88 (81-95)	0.47	<.001	0.48
Surgeon	82 (75-89)	54 (37-72)	93 (87-99)	0.52	<.001	0.53
Radiology resident 1	88 (82-94)	66 (49-82)	97 (93-100)	0.67	<.001	0.69
Radiology resident 2	90 (83-94)	69 (52-81)	99 (94-100)	0.74	<.001	0.75

[a]Not applicable.

Table 4. Overall diagnostic performance of GPTs and physicians for clinical N-stage (cN-stage).

	Accuracy (%) (95% CI)	Cohen κ value	*P* value
GPT-5	48 (40-57)	0.18	<.01
GPT-4.5	43 (34-52)	0.051	—[a]
GPT-4.1	45 (36-54)	0.12	<.01
OpenAI-o3	39 (31-48)	0.043	—
OpenAI-o1	34 (26-43)	0.055	—
GPT-turbo	34 (26-43)	–0.072	—
Radiologist	58 (49-67)	0.38	<.01
Surgeon	61 (52-70)	0.34	<.001
Radiology resident 1	61 (52-70)	0.39	<.001
Radiology resident 2	61 (52-69)	0.39	<.001

[a]Not applicable.

Table 5. Overall diagnostic performance of GPTs and physicians for clinical M-stage (cM-stage).

	Accuracy (%) (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Cohen κ value	*P* value	Matthews correlation coefficient
GPT-5	77 (68-83)	4 (1-18)	98 (92-99)	0.023	—[a]	0.042
GPT-4.5	76 (68-84)	0 (0-0)	98 (95-100)	–0.032	—	–0.07
GPT-4.1	77 (68-83)	0 (0-0)	99 (94-100)	–0.016	—	–0.05
OpenAI-o3	77 (68-83)	4 (1-18)	98 (92-99)	0.023	—	0.04
OpenAI-o1	78 (70-85)	4 (0-11)	99 (97-100)	0.039	—	0.09
GPT-turbo	78 (70-86)	15 (0.5-29)	96 (92-100)	0.14	<.01	0.18
Radiologist	78 (70-85)	33 (14-52)	90 (84-96)	0.27	<.001	0.28
Surgeon	85 (79-92)	52 (32-72)	95 (90-99)	0.52	<.001	0.53
Radiology resident 1	77 (69-84)	56 (36-76)	83 (75-91)	0.36	<.001	0.37
Radiology resident 2	83 (76-89)	56 (37-72)	91 (84-96)	0.50	<.001	0.50

[a]Not applicable.

In the analysis of diagnostic correlation, GPT-5 and GPT-4.1 both demonstrated statistically significant weak correlations with the diagnoses of thoracic LN metastasis, abdominal LN metastasis, and cN stage. GPT-4.5 showed a statistically

XSL•FO
RenderX

significant weak correlation for the diagnoses of thoracic and abdominal LN metastasis. GPT-4 Turbo showed a statistically significant weak correlation in the diagnosis of the stage. The other models did not demonstrate statistically significant consistency. In contrast, all physicians demonstrated a statistically significant moderate consistency for all items.

Comparison of Accuracy

First, we compared the average correct answer rates of LLMs and physicians. In the assessment of thoracic LN metastases, LLMs achieved 68/120 (57%; 95% CI 52%-61%) accuracy, whereas physicians achieved 91/120 (76%; 95% CI 72%-80%) accuracy. In the evaluation of abdominal LNs, LLMs reached an accuracy of 82/120 (68%; 95% CI 62%-73%) compared with 102/120 (85%; 95% CI 78%-91%) for physicians. For cN stage diagnosis, LLMs attained 49/120 (41%; 95% CI 36%-45%) accuracy, whereas physicians achieved 71/120 (60%; 95% CI 55%-64%) accuracy. In the cM-stage assessment, the LLMs achieved 92/120 (77%; 95% CI 75%-80%) accuracy, which was slightly lower than the 96/120 (80%; 95% CI 77%-83%) accuracy observed among the physicians. Overall, physicians demonstrated significantly higher accuracy than LLMs in the evaluation of thoracic LN metastasis, abdominal LN metastasis, and cN stage (P<.05; Table 6).

Table 6. Comparison of average accuracy between large language models (LLMs) and physicians.

	Thoracic LN[a] (%) (95% CI)	Abdominal LN (%) (95% CI)	cN-stage[b] (%) (95% CI)	cM-stage[c] (%) (95% CI)
LLMs (%)	57 (52-61)	68 (62-73)	41 (36-45)	77 (75-80)
Physicians (%)	76 (72-80)	85 (78-91)	60 (55-64)	80 (77-83)
P value	<.001	.002	<.001	.052

[a]LN: lymph node.

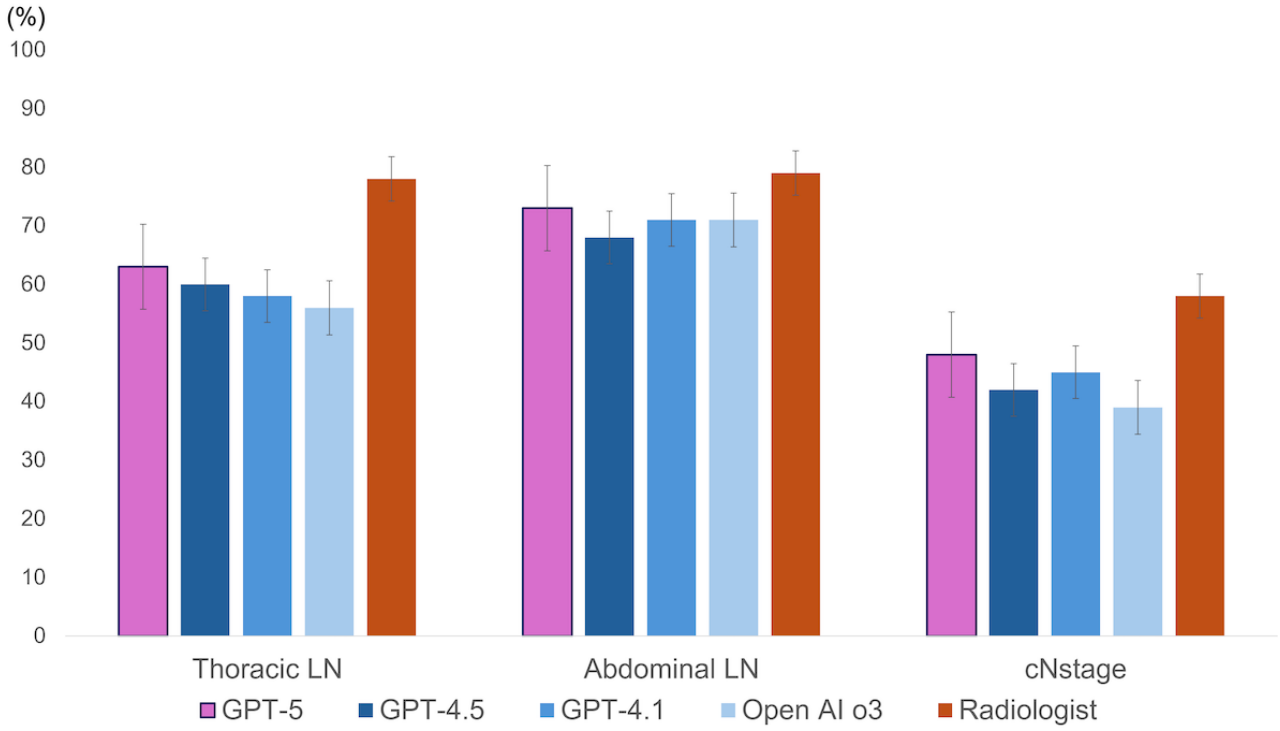
[b]cN: clinical N.

[c]cM: clinical M.

Among the LLMs, GPT-5 demonstrated the highest diagnostic accuracy for thoracic LN metastasis, abdominal LN metastasis, and cN stage, and achieved one of the highest accuracies for the cM stage. The results of McNemar's pairwise test between the LLMs GPT-5, GPT-4.5, GPT-4.1, and OpenAI-o3 and the radiologist are shown in Figure 4. Comparisons between the LLMs revealed no statistically significant differences across most evaluated parameters; however, a statistically significant difference was observed between GPT-5 and OpenAI-o3 in the diagnosis of thoracic LN metastasis. In the diagnosis of thoracic LN metastasis, all LLMs demonstrated significantly lower accuracy than radiologists. However, for cN stage diagnosis, GPT-5 and GPT-4.1 showed no statistically significant difference from that of the radiologists. Moreover, in the diagnosis of abdominal LN metastasis, no significant differences were observed between any of the LLMs and the radiologists (P<.05).

XSL•FO
RenderX

Figure 4. Accuracy of large language models (LLMs) and physicians for each category. Error bars represent 95% CIs for each accuracy. OpenAI-o1 and GPT-4 turbo were excluded from the figure for clarity, as their performance was consistently lower than the other models, as detailed in Tables 2-5. cN: clinical N; LN: lymph node.



Interrater Reliability

In the analysis of diagnostic correlation, GPT-5 and GPT-4.1 both demonstrated statistically significant weak correlations with the diagnoses of thoracic LN metastasis, abdominal LN metastasis, and cN stage. GPT-4.5 showed a statistically significant weak correlation for the diagnoses of thoracic and abdominal LN metastasis. GPT-4 Turbo showed a statistically significant weak correlation in the diagnosis of the stage. The other models did not demonstrate statistically significant consistency. In contrast, all physicians demonstrated a statistically significant moderate consistency for all items (Tables 2-5).

Discussion

Summary of Results

To our knowledge, this is the first study to evaluate the newly released GPT-5 for staging esophageal cancer using ^{18}F FDG-PET images. Our results demonstrate a statistically significant performance gap between that of physicians and current LLMs. While diagnostic accuracy varied across individual models and physicians, the average physician performance was significantly superior to that of the LLM in assessing thoracic LN metastasis (91/120, 76%; 95% CI 72%-80% vs 68/120, 57%; 95% CI 52%-61%; P<.001), abdominal LN metastasis (102/120, 85%; 95% CI 78%-91 vs 82/120, 68%; 95% CI 62%-73%; P=.002), and cN stage (71/120, 60%; 95% CI 55%-64% vs 49/120, 41%; 95% CI 36%-45%; P<.001; Table 6). Furthermore, LLM interpretations showed poor consistency (Cohen κ: –0.07 to 0.25), contrasting with the fair-to-substantial agreement observed among physicians (κ: 0.27 to 0.74). These statistical findings confirm that, despite some overlapping accuracy ranges, current general-purpose LLMs reliably underperform compared with human experts in complex staging tasks.

Principal Findings

Among the evaluated LLMs, GPT-5 demonstrated the highest diagnostic performance. It achieved accuracies of 76/120 (63%; 95% CI 54%-71%) for thoracic LN and 87/120 (73%; 95% CI 64%-80%) for abdominal LN assessment, numerically outperforming other models like GPT-4.5 (72/120, 60%, 95% CI 51%-68% and 83/120, 69%, 95% CI 61%-77%) and GPT-4.1 (70/120, 58%, 95% CI 49%-66% and 85/120, 71%, 95% CI 62%-78%; Table 7 GPTs and radiologists). This superior performance underscores the rapid advancement of these models, likely attributable to architectural enhancements and more comprehensive multimodal training [30].

However, a critical analysis of these results reveals a fundamental limitation in using general-purpose generative AI for specialized clinical tasks. While GPT-5 had high specificity (115/120, 96%; 95% CI 90%-99% for abdominal LNs), its sensitivity was critically low (5/35, 14%; 95% CI 6%-29%) when compared with radiologists (20/35, 57%; 95% CI 40%-74%). This indicates that while the model is effective at identifying “normal” findings (True Negatives), it fails to reliably detect pathology (False Negatives). The discrepancy between the high accuracy and relatively low MCC scores observed in the LLMs further confirms that their performance was driven primarily by the correct identification of negative cases (specificity), rather than a balanced detection capability required for clinical staging. This performance profile suggests that general-purpose VLMs, which are primarily trained on

XSL•FO
RenderX

natural images and text, currently lack the domain-specific visual calibration required to distinguish subtle metastatic uptake from physiological noise in medical imaging.

Table 7. Comparison of the accuracy.

	Thoracic LN[a] (%) (95% CI)	Abdominal LN (%) (95% CI)	cN-stage[b] (%) (95% CI)	cM-stage[c] (%) (95% CI)
Physicians[d]				
Radiologist (%)	78 (71-86)	80 (73-87)	58 (49-67)	78 (70-85)
Surgeon (%)	74 (66-82)	82 (75-89)	61 (52-70)	85 (79-92)
Radiology resident 1 (%)	74 (66-82)	88 (82-94)	61 (52-70)	77 (69-84)
Radiology resident 2 (%)	80 (72-96)	90 (83-94)	61 (52-69)	83 (76-89)
GPTs and radiologists				
GPT-5 (%)	63 (54-71)	73 (64-80)	48 (40-57)	77 (68-83)
GPT-4.5 (%)	60 (51-68)	69 (61-77)	43 (34-52)	76 (68-84)
GPT-4.1 (%)	58 (49-66)	71 (62-78)	45 (36-54)	77 (68-83)
OpenAI-o3 (%)	56 (47-64)	71 (62-78)	39 (31-48)	77 (68-83)
OpenAI-o1 (%)	50 (40-59)	56 (47-65)	34 (26-43)	78 (70-85)
GPT-turbo (%)	52 (43-61)	66 (57-74)	34 (26-43)	78 (70-86)

[a]LN: lymph node.

[b]cN: clinical N.

[c]cM: clinical M.

[d]Thoracic LN: Cochran Q=4 (P=.26); Abdominal LN: Cochran Q=12.3 (P=.006); cN stage: Cochran Q=0.55 (P=.91); cM stage: Cochran Q=8.4 (P=.039).

Comparison to Prior Work

Radiological image diagnosis using generative AI has been explored across various modalities, including plain radiography, CT, and ultrasound, with reported accuracies varying widely from 27.8% to 88% [15-18,31,32]. Many studies conclude that generative AI performance remains suboptimal for clinical use. Hong et al [33] found that no model achieved clinical-grade applicability for reading chest radiographs due to significant false positives, false negatives, and hallucinations. Our findings align with this body of literature, confirming that substantial advancements are needed before generative AI can be practically applied in this clinical setting (Table S2 in Multimedia Appendix 2 [15,17,31,32,34]). Unlike earlier studies focused on simple classification or structuring textual data, our study targeted a core radiologist workflow: generating diagnostic reports directly from medical images. While LLMs excel at summarizing text and extracting information from existing reports [34,35], few studies have explored their ability to derive TNM classifications from images, a task requiring both image interpretation and clinical reasoning. Previous research has often focused on T-factor classification, such as identifying a mass or its size. Our work extends this by comprehensively investigating N and M stage classification of malignant tumors using PET images. Evaluating LLM performance on complex clinical tasks, rather than simple diagnosis, is crucial for assessing their future clinical potential.

Furthermore, although the models were not provided with explicit segmentation masks, we hypothesized that current VLMs would use their multimodal capabilities [15,16,29] to map the semantic text label, such as “middle thoracic,” to the corresponding high-uptake region on the MIP image. We hypothesized that the models would interpret this text input as a spatial guide by recognizing anatomical landmarks such as the proximity to the heart, thereby allowing them to distinguish and exclude the primary tumor from other metastatic lesions. This hypothesis was supported by our post hoc qualitative subanalysis, in which the model’s generated reasoning suggested that it identified the primary tumor location based on the provided text prompt.

Impact on Clinical Management

The performance gap between physicians and LLMs has significant implications for clinical decision-making. Accurate staging, particularly the detection of nodal and distant metastases, is critical for determining whether patients are candidates for curative surgery versus multimodal therapy. The low sensitivity of LLMs observed in this study (eg, 14% for abdominal LNs by GPT-5) poses a substantial risk of under-staging. In a clinical setting, relying on such a system could lead to the omission of necessary neoadjuvant chemotherapy or the performance of futile surgeries on patients with undetected metastases. In contrast, physicians demonstrated significantly higher sensitivity and balanced accuracy, ensuring that high-risk patients are appropriately identified for systemic treatment. Therefore, while LLMs show promise in specificity, their current lack of sensitivity precludes their utility as a standalone diagnostic tool for treatment planning.

A key factor limiting LLM performance in medical imaging is the fundamental mismatch between their text-centric design

XSL•FO
RenderX

and the demands of visual analysis. As LLMs are primarily trained on textual data, they excel in natural language understanding and reasoning but lack the capability to process and analyze complex visual information [32]. This limitation is reflected in the observation that text-based report structuring consistently outperformed direct image-based diagnosis in radiology report generation. To improve accuracy, future research should prioritize architectures that better integrate text and visual data. Incorporating multimodal learning frameworks that combine textual and imaging information might enhance diagnostic performance and facilitate clinical applicability [36,37].

Limitations

This study has several limitations that should be acknowledged.

First, our dataset was limited by a significant class imbalance, particularly in M-stage classification, where only 27/120 (22.5%) of cases were M1-positive. Consequently, the resulting CIs for sensitivity were wide, and the study may be underpowered to detect significant differences in sensitivity for distant metastases. Such imbalances are known to bias machine learning models toward the majority class, potentially leading to overestimated specificity and underestimated sensitivity. Furthermore, potential image resolution degradation during the conversion and trimming of DICOM files may have impacted the diagnostic accuracy of the LLMs. A more balanced and carefully processed dataset would enable a more robust evaluation of model performance. Additionally, because the LLMs were prompted to provide binary classification outputs (yes/no) in a zero-shot setting rather than continuous probability scores, receiver operating characteristic curve analysis and area under the curve calculation were not feasible in this study.

Second, the reliance on a single MIP image for each case does not reflect standard clinical practice. MIPs are 2D condensations that can omit crucial spatial and anatomical details necessary for accurate TNM staging, which clinicians typically determine by reviewing multiplanar image slices and integrating information from CT scans. This methodological constraint may have disadvantaged the LLMs when compared with human interpretation. Furthermore, our input was strictly limited to visual information from MIP images and did not include semiquantitative metabolic parameters (eg, SUVmax) or volumetric indices (eg, metabolic tumor volume), which are integral to standard PET/CT interpretation for differentiating malignant from physiological uptake. This is further supported by the variability in diagnostic accuracy observed among physicians within our study, which suggests a potential discrepancy between radiological assessment in this experimental setting and actual clinical workflows.

Third, the diagnostic criteria were not explicitly defined for either the LLMs or the human evaluators. While a subanalysis requesting the “basis for diagnosis” was performed to check for logical consistency (Table S1 in Multimedia Appendix 1), we acknowledge that this generated text itself represents a potential hallucination and is not a substitute for visual attention maps. The subanalysis revealed that while the model often generates text regarding clinical exclusion criteria (eg, ignoring hilar nodes), it remains prone to hallucinations (Case C) and stochastic instability (as was with Case B). The addition of a reasoning prompt paradoxically led to a false-negative result in a previously correctly diagnosed case. Crucially, due to the “black box” nature of commercial APIs, we could not generate saliency maps or obtain reliable bounding box coordinates to verify the models' focus. Consequently, we cannot determine whether the “correct” classifications were achieved based on appropriate anatomical features or simply represent “right answers for the wrong reasons.”

Fourth, the use of a private, single-institution dataset limits the generalizability of our findings. Differences in imaging protocols, patient populations, and clinical workflows across institutions can significantly affect model performance, making external validation with larger, multicenter datasets essential. However, it is noteworthy that the LLMs were applied in their general-purpose form without task-specific fine-tuning. While differing from traditional machine learning models that are often customized, this approach facilitates performance testing across diverse environments, suggesting that LLMs may be suitable for broader clinical and research applications where consistency and ease of validation are important.

Finally, this was a single-institution study focusing exclusively on esophageal SCC. Differences in imaging protocols, patient demographics, and disease pathologies across institutions were not accounted for, limiting the external validity of our findings.

Future Directions

Future research should prioritize several key areas to improve diagnostic accuracy and clinical utility. First, overcoming the limitations of 2D MIP images is essential; integrating volumetric data from CT and 3D PET scans is necessary to capture the spatial and anatomical details required for accurate staging. Second, future studies should use multimodal learning frameworks that better synthesize textual clinical data with imaging features, rather than relying only on text-centric architectures. Third, to address the “black box” nature of current models, assessments should include outputs such as bounding boxes or heatmaps to verify that the model is identifying the correct pathology rather than hallucinating lesions.

Finally, conducting external validation using larger, multi-institutional datasets is crucial to assess generalizability across different imaging protocols and diverse patient populations.

Conclusions

Current general-purpose LLMs, including GPT-5, do not achieve physician-level diagnostic accuracy for esophageal cancer staging based on MIP images. While newer models demonstrate improved specificity and a reduction in hallucinations when compared with those of earlier iterations, their sensitivity for detecting nodal and distant metastases remains insufficient for clinical use. These findings suggest that while LLMs hold potential as future support tools, they currently cannot replace or reliably augment expert radiological assessment in this domain. Future development must prioritize the integration of volumetric data and multimodal capabilities to bridge the notable performance gap observed in this study.

XSL•FO
RenderX

Acknowledgments

The author used Google Gemini 3 Pro for spell checking and formatting the paper.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

HM, YT, and KT designed and conceived this study. HM, YT, YA, MI, and YN collected data. HM, YT, YA, and KT analyzed and interpreted the results and drafted the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Qualitative Assessment of Diagnostic Reasoning.
[XLSX File (Microsoft Excel File), 10 KB - cancer_v12i1e86630_app1.xlsx]

Multimedia Appendix 2
Comparison of recent studies evaluating multimodal large language models in radiological imaging.
[XLSX File (Microsoft Excel File), 13 KB - cancer_v12i1e86630_app2.xlsx]

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263 [FREE Full text] [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Al-Ibraheem A, Abdlkadir A, Herrmann K, Bomanji J, Jadvar H, Shi H, et al. Diagnostic accuracy of [18F]FDG PET/MRI in head and neck squamous cell carcinoma: a systematic review and metaanalysis. J Nucl Med 2024;65(10):1533-1539 [FREE Full text] [doi: 10.2967/jnumed.124.268049] [Medline: 39266291]
3. GlobalData Plc. The Complexities of Physician Supply and Demand: Projections from 2021 to 2036. Washington, DC: AAMC; 2021.
4. Takami H, Kodera Y, Eguchi H, Kitago M, Murotani K, Hirano S, et al. The shortage of surgeons in Japan: results of an online survey of qualified teaching hospitals that take part in the surgical training programs for board certification by the Japan Surgical Society. Surg Today 2024;54(1):41-52 [FREE Full text] [doi: 10.1007/s00595-023-02697-7] [Medline: 37193795]
5. The Royal College of Radiologists. Clinical radiology UK workforce census report 2023. Royal College of Radiologists. 2024. URL: https://www.rcr.ac.uk/media/4imb5jge/_rcr-2024-clinical-radiology-workforce-census-report.pdf [accessed 2025-08-11]
6. Afshari Mirak S, Tirumani SH, Ramaiya N, Mohamed I. The growing nationwide radiologist shortage: current opportunities and ongoing challenges for international medical graduate radiologists. Radiology 2025;314(3):e232625. [doi: 10.1148/radiol.232625] [Medline: 40035678]
7. Smith-Bindman R, Kwan ML, Marlow EC, Theis MK, Bolch W, Cheng SY, et al. Trends in use of medical imaging in US health care systems and in Ontario, Canada, 2000-2016. JAMA 2019;322(9):843-856 [FREE Full text] [doi: 10.1001/jama.2019.11456] [Medline: 31479136]
8. Frood R, Willaime JMY, Miles B, Chambers G, Al-Chalabi H, Ali T, et al. Comparative effectiveness of standard vs. AI-assisted PET/CT reading workflow for pre-treatment lymphoma staging: a multi-institutional reader study evaluation. Front Nucl Med 2023;3:1327186. [doi: 10.3389/fnume.2023.1327186] [Medline: 39355039]
9. Agress H, Wong TZ, Shreve P. Interpretation and reporting of positron emission tomography-computed tomographic scans. Semin Ultrasound CT MR 2008;29(4):283-290. [doi: 10.1053/j.sult.2008.05.001] [Medline: 18795496]
10. Expert Panels on Thoracic and Gastrointestinal Imaging, Raptis CA, Goldstein A, Henry TS, Porter KK, Catenacci D, et al. ACR appropriateness criteria® staging and follow-up of esophageal cancer. J Am Coll Radiol 2022;19(11S):S462-S472 [FREE Full text] [doi: 10.1016/j.jacr.2022.09.008] [Medline: 36436970]
11. Obermannová R, Alsina M, Cervantes A, Leong T, Lordick F, Nilsson M, et al. Oesophageal cancer: ESMO clinical practice guideline for diagnosis, treatment and follow-up. Ann Oncol 2022;33(10):992-1004 [FREE Full text] [doi: 10.1016/j.annonc.2022.07.003] [Medline: 35914638]

XSL•FO
RenderX

12. Al-Ibraheem A, Abdlkadir AS, Shagera QA, Saraireh O, Al-Adhami D, Al-Rashdan R, et al. The diagnostic and predictive value of 18F-fluorodeoxyglucose positron emission tomography/computed tomography in laryngeal squamous cell carcinoma. Cancers (Basel) 2023;15(22):5461 [FREE Full text] [doi: 10.3390/cancers15225461] [Medline: 38001720]
13. Geus-Oei LF, Oyen WJ. Predictive and prognostic value of FDG-PET. Cancer Imaging 2008;8(1):70-80 [FREE Full text] [doi: 10.1102/1470-7330.2008.0010] [Medline: 18390390]
14. Maruyama H, Toyama Y, Takanami K, Takase K, Kamei T. Role of artificial intelligence in surgical training by assessing GPT-4 and GPT-4o on the Japan surgical board examination with text-only and image-accompanied questions: performance evaluation study. JMIR Med Educ 2025;11:e69313 [FREE Full text] [doi: 10.2196/69313] [Medline: 40737609]
15. Dehdab R, Brendlin A, Werner S, Almansour H, Gassenmaier S, Brendel JM, et al. Evaluating GPT-4V in chest ct diagnostics: a critical image interpretation assessment. Jpn J Radiol 2024;42(10):1168-1177. [doi: 10.1007/s11604-024-01606-3] [Medline: 38867035]
16. Chen Z, Chambara N, Wu C, Lo X, Liu SYW, Gunda ST, et al. Assessing the feasibility of GPT-4o and Claude 3-Opus in thyroid nodule classification based on ultrasound images. Endocrine 2024;87(3):1041-1049. [doi: 10.1007/s12020-024-04066-x]
17. Chetla N, Tandon M, Chang J, Sukhija K, Patel R, Sanchez R. Evaluating GPT's efficacy in pediatric pneumonia detection from chest X-rays: comparative analysis of specialized aI models. JMIR AI 2025;4:e67621 [FREE Full text] [doi: 10.2196/67621] [Medline: 39793007]
18. Lee KH, Lee RW, Kwon YE. Validation of a deep learning chest X-ray interpretation model: integrating large-scale AI and large language models for comparative analysis with GPT. Diagnostics (Basel) 2023;14(1):90 [FREE Full text] [doi: 10.3390/diagnostics14010090] [Medline: 38201398]
19. Mongan J, Moy L, Kahn CE. Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers. Radiol Artif Intell 2020;2(2):e200029. [doi: 10.1148/ryai.2020200029]
20. Brierley JD, Gospodarowicz MK, Wittekind C. TNM Classification of Malignant Tumours. Hoboken: Wiley-Blackwell; 2016.
21. Girum KB, Rebaud L, Cottereau AS, Meignan M, Clerc J, Vercellino L, et al. 18F-FDG PET maximum-intensity projections and artificial intelligence: a win-win combination to easily measure prognostic biomarkers in DLBCL patients. J Nucl Med 2022;63(12):1925-1932 [FREE Full text] [doi: 10.2967/jnumed.121.263501] [Medline: 35710733]
22. Mine S, Tanaka K, Kawachi H, Shirakawa Y, Kitagawa Y, Toh Y, et al. Japanese classification of esophageal cancer. Esophagus 2024;21(3):179-215 [FREE Full text] [doi: 10.1007/s10388-024-01054-y] [Medline: 38568243]
23. Usage policies. OpenAI. URL: https://OpenAI.com/policies/usage-policies [accessed 2025-05-23]
24. Dogra S, Zhang X, Silva E, Rajpurkar P. The financial, operational, and clinical advantages of generalist radiology AI. Radiology 2025;316(3):e242362. [doi: 10.1148/radiol.242362] [Medline: 40923883]
25. Evaluation of multimodal generative AI for esophageal cancer staging using FDG-PET: diagnostic accuracy and comparison with physicians. GitHub. URL: https://github.com/YuyaAraki/chatgpt-nuclear-imaging [accessed 2025-10-10]
26. Yang Y, Li Y, Qin J, Zhang R, Chen X, He J, et al. Mapping of lymph node metastasis from thoracic esophageal cancer: a retrospective study. Ann Surg Oncol 2022;29(9):5681-5688. [doi: 10.1245/s10434-022-11867-9] [Medline: 35543907]
27. Chicco D, Jurman G. The advantages of the matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 2020;21(1):6 [FREE Full text] [doi: 10.1186/s12864-019-6413-7] [Medline: 31898477]
28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-174 [FREE Full text] [doi: 10.2307/2529310]
29. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 1947;12(2):153-157. [doi: 10.1007/bf02295996] [Medline: 20254758]
30. GPT-5 is here. URL: https://OpenAI.com/gpt-5 [accessed 2025-08-13]
31. Lee RW, Lee KH, Yun JS, Kim MS, Choi HS. Comparative analysis of M4CXR, an LLM-based chest X-ray report generation model, and GPT in radiological interpretation. J Clin Med 2024;13(23):7057 [FREE Full text] [doi: 10.3390/jcm13237057] [Medline: 39685515]
32. Reith TP, D'Alessandro DM, D'Alessandro MP. Capability of multimodal large language models to interpret pediatric radiological images. Pediatr Radiol 2024;54(10):1729-1737. [doi: 10.1007/s00247-024-06025-0]
33. Hong EK, Ham J, Roh B, Gu J, Park B, Kang S, et al. Diagnostic accuracy and clinical value of a domain-specific multimodal generative AI model for chest radiograph report generation. Radiology 2025;314(3):e241476. [doi: 10.1148/radiol.241476] [Medline: 40131111]
34. Chen K, Xu W, Li X. The potential of gemini and GPTs for structured report generation based on free-text 18F-FDG PET/CT breast cancer reports. Acad Radiol 2025;32(2):624-633 [FREE Full text] [doi: 10.1016/j.acra.2024.08.052] [Medline: 39245597]
35. Salam B, Kravchenko D, Nowak S, Sprinkart AM, Weinhold L, Odenthal A, et al. Generative pre-trained transformer 4 makes cardiovascular magnetic resonance reports easy to understand. J Cardiovasc Magn Reson 2024;26(1):101035 [FREE Full text] [doi: 10.1016/j.jocmr.2024.101035] [Medline: 38460841]

XSL•FO
RenderX

36. Wang X, Figueredo G, Li R, Zhang WE, Chen W, Chen X. A survey of deep-learning-based radiology report generation using multimodal inputs. Med Image Anal 2025;103:103627 [FREE Full text] [doi: 10.1016/j.media.2025.103627] [Medline: 40382855]
37. Clusmann J, Kolbinger FR, Muti HS, Carrero ZI, Eckardt J, Laleh NG, et al. The future landscape of large language models in medicine. Commun Med (Lond) 2023;3(1):141 [FREE Full text] [doi: 10.1038/s43856-023-00370-1] [Medline: 37816837]

Abbreviations

AI: artificial intelligence
API: application programming interfaces
cM: clinical M
cN: clinical N
CT: computed tomography
FDG-PET: fluorodeoxyglucose positron emission tomography
LLM: large language model
LN: lymph node
MIP: maximum intensity projection
MCC: Matthews Correlation Coefficient
PET: positron emission tomography
SCC: squamous cell carcinoma

Edited by M Balcarras; submitted 27.Oct.2025; peer-reviewed by D Al-Adhami, L Mosca; comments to author 03.Dec.2025; revised version received 30.Jan.2026; accepted 30.Jan.2026; published 23.Feb.2026.

Please cite as:
Maruyama H, Toyama Y, Araki Y, Takanami K, Ito M, Nakajima Y, Takase K, Kamei T
Evaluation of GPT-5 for Esophageal Cancer Staging Using Fluorodeoxyglucose Positron Emission Tomography Maximum-Intensity Projection Images: Comparative Pilot Study
JMIR Cancer 2026;12:e86630
URL: https://cancer.jmir.org/2026/1/e86630
doi:10.2196/86630
PMID:

©Hiroki Maruyama, Yoshitaka Toyama, Yuya Araki, Kentaro Takanami, Masato Ito, Yumi Nakajima, Kei Takase, Takashi Kamei. Originally published in JMIR Cancer (https://cancer.jmir.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Revealing the Physician's Task Conceptualization in Melanoma Treatment: Qualitative User Experience Study

Eva Maria Hartmann[1,2], MSc; Alisa Küper[3], PhD; Jessica Swoboda[4], MSc; Catharina Lena Beckmann[2], Dr rer medic; Georg Christian Lodde[5], Dr med; Elisabeth Livingstone[5], Prof Dr Med; Dirk Schadendorf[5], Prof Dr Med; Sabine Sachweh[2], Dr rer nat

[1]Data Integration Center, Essen University Hospital, Essen, Germany

[2]Department of Computer Science, Dortmund University of Applied Sciences and Arts, Dortmund, Germany

[3]Department of Social Psychology: Media and Communication, University of Duisburg-Essen, Duisburg, Germany

[4]Institute for Artificial Intelligence in Medicine, Essen University Hospital, Essen, Germany

[5]Clinic for Dermatology, Essen University Hospital, Essen, Germany

Corresponding Author:
Eva Maria Hartmann, MSc
Data Integration Center
Essen University Hospital
Girardetstraße 2
Essen, 45147
Germany
Phone: 49 15201933426
Email: eva.hartmann2@uk-essen.de

Abstract

Background: To support physicians in focusing on relevant information during melanoma treatment, it is essential to design information systems (ISs) that integrate into their workflow.

Objective: This study aimed to identify the work steps, knowledge sources, and key information during tumor board registration, revealing the physician's cognitive workflow and task conceptualization.

Methods: We applied the concurrent think-aloud method with 10 dermatologists from the University Hospital Essen's skin cancer unit, providing direct insights into their work. Participants had varying levels of experience, from a few weeks to over a decade in the department, and differing levels of previous patient knowledge. Voice and screen recordings were transcribed and inductively annotated across 3 categories: "work step," "information in focus," and "program used." Furthermore, 4 annotators from medical informatics and social psychology conducted the annotation process. These annotations were used to map the workflow model via a sequence diagram, alongside a quantitative analysis of code occurrences and overlaps.

Results: Patient-related data in clinical IS are distributed across nested submodules, requiring physicians to extract key information through extensive system interactions. The workflow consists of 4 distinct phases with 10 work steps, taking on average 13 minutes to complete. Physicians switched between 3 programs approximately 25 times, requiring a minimum of 16 interactions to access essential information from the hospital IS. The analysis identified 7 fundamental pain points: time-intensive login and patient selection (consuming 7% of workflow time), need for multiple system instances, data fragmentation across systems, hierarchical information modularization, interface structures hindering simultaneous viewing, repetitive data insertion (27% of time), and inappropriate input fields forcing external text editor use (28% of time). Collectively, approximately 55% of workflow time was dedicated to system interactions and manual data compilation rather than clinical decision-making. Physicians frequently copied information from various reports to compile a free-text case summary, representing an extensive workaround addressing data fragmentation. External editing of the case summary was highly iterative, with an average of 3.5 repetitions. The workflow phase "updating the knowledge" showed the greatest variety in step sequences, with physicians accessing radiology reports, laboratory values, tumor board protocols, and progress documentation in varying orders.

Conclusions: These insights offer a deeper understanding of physicians' intentions when handling medical information and their challenges with current clinical IS. The workflow can be understood as a comprehensive workaround due to data distribution, with physicians compensating through experience-based adaptations. This understanding of physicians' task conceptualization

XSL•FO
RenderX

forms the basis for designing software solutions that better support workflows at the point of care by organizing and presenting relevant data more effectively.

(JMIR Cancer 2026;12:e79294) doi:10.2196/79294

KEYWORDS

clinical information systems; decision making; electronic health records; human-computer interaction & usability; knowledge management; user research; melanoma

Introduction

Medical information systems (ISs) are used in the clinical setting to access patient files, document treatment, and support decision-making processes. Depending on the specific discipline, specialized ISs such as the radiology IS or the laboratory IS are available. These systems document field-specific data that may not be appropriately captured by the standard hospital information system (HIS). Despite the increased usage of medical IS in the European Union [1], many medical personnel express a desire to abandon their use due to a range of usability issues [2,3]. The rigidity of HIS and clinical decision support systems (CDSSs) has been identified as a significant issue, hindering workflows when dealing with nonroutine processes [4] and forcing users to adjust their routine [5]. Furthermore, the increased time required for documentation and information gathering represents another factor contributing to the rejection of these systems [5]. Additionally, the high mental load caused by poorly designed systems [5,6], which results in stress [7] and burnout [8,9], is a significant concern. Many applications are not sufficiently adapted to the specific use cases they are intended to support due to a lack of user- and context-driven design and implementation [10,11].

To address these issues, our project incorporated a thorough workflow analysis into the development process of a clinical dashboard. This approach reveals the underlying contextual interpretation of physicians at the point of care in a mental model, thereby enabling the identification of the most appropriate solution [11-14]. The term “mental model” was introduced by Norman [15] to describe the understanding users’ conceptual approach to a task. It is essential to design systems that align with actual work patterns rather than assumed ones. Accordingly, it is essential to investigate context-related factors, such as task complexity [3], diverse user preferences [14,16], support of individual levels of work experience, and the application adaptability in the specific use case [16]. To reveal the workflow, the focus needs to be on the exploration of data and functionalities [17]. Integrating these insights into a comprehensive workflow model and using this as a starting point for design and implementation allows for workflow integrity of the software [3,18]. The uncovering of missing or incomplete mental models can explain the rationale behind workarounds, which evolve from workflow interruptions [19]. This enhances patient safety [13] as a result of the avoidance of these workarounds and related errors [4,20]. Furthermore, the integration into the workflow offers the potential for more effective implementation of evidence-based treatment and decision-making [5]. The presentation of relevant information at the optimal time and in the most suitable format [11,21] also serves to reduce repetitive and inefficient activities [11] and thereby increase overall efficiency. The involvement of the user in the development process encourages a positive perspective toward IS, which, in turn, positively influences acceptance [14].

The context of treatment of patients with melanoma is particularly appropriate for this analysis from 2 perspectives. On one hand, it is estimated that melanoma will account for 5% of new cancer cases in the United States in 2024, ranking it as the fifth most common cancer in the country [22]. The 5-year survival rate for melanoma is promising when diagnosed in early stages (stage 1: 99%, stage 2: 85%, and stage 3: 75%) [22-24]. In stage 4 settings, the correct tumor classification comes into focus. Advanced diagnostics allow for personalized treatment, which increases the likelihood of surviving for 5 years by 10% [25]. This underscores the importance of the physician’s comprehensive and detailed knowledge of the patient’s condition for early detection of melanoma and optimal treatment. On the other hand, the treatment of patients with skin tumors in Germany is conducted in specialized skin tumor centers by dermatologists specializing in oncology. A single team provides all stages of care: diagnosis, histopathology, surgery, adjuvant and advanced treatment, and follow-up. This approach allows for comprehensive, holistic care, as well as the generation of insights that may be applicable to other tumor types.

The objective of this project was to gain insight into the process of treatment of patients with melanoma by examining the task of gathering the relevant information and preparing patients’ documents for an interdisciplinary tumor board discussion from a user-centered perspective. Consequently, an examination was conducted of the work steps, the information that is the focus in each step, and the data sources (programs) that are used. By combining the insights into a workflow model, the relationships between work steps, data, program interactions, and intents are highlighted. The broad perspective enabled identification of pain points in software interactions, decision-making criteria, and the relationship between intents and relevant information across all workflow steps.

Methods

Ethical Considerations

This study was approved by the Ethics Committee of the Faculty of Engineering (approval 2111SPKA9493) of the University of Duisburg-Essen. All procedures were performed in accordance with the ethical standards of the institutional research committee, the 1964 Helsinki Declaration, and its later amendments or comparable ethical standards. All participants provided their informed consent in written form for inclusion in the think-aloud

XSL•FO
RenderX

(TA) method. Any information allowing patient identification was removed from the recordings before the transcription process. There was no compensation for participation.

Approach

This study aimed at gaining insights into the process of gathering the relevant information and preparing patients' documents for an interdisciplinary tumor board discussion. Therefore, the TA method was conducted with 10 physicians of varying levels of expertise at the University Hospital Essen in Germany. The group included resident physicians with a few weeks to several years of experience, as well as senior physicians in the field of dermatology. The transcripts of those sessions were inductively annotated regarding their contained work steps, considered information, and programs in use. The detailed code systems can be accessed in the Excel sheet tab: "code systems" in Multimedia Appendix 1. We did a cognitive task analysis to reveal the underlying cognitive workflow and task conceptualization, which resulted in a sequence diagram (SeqD) that highlights the workflow. Additionally, we analyzed the interactions with the programs used during the process to align the work steps with the data prioritization and software interactions.

TA Method

The TA method [26] is a widespread methodology [27-29] for assessing the suitability of implementations or prototypes within their intended context [27]. By asking users to express their thoughts aloud concurrently (concurrent think-aloud [CTA] method) while conducting a task, it is possible to gain valuable insights into their cognitive approach and intentions in real time, in contrast with retrospective TA where reflections are gathered after task completion. As a method frequently used in iterative development processes [28], this knowledge enables designers and developers to create products that align with actual needs rather than assumed ones [30]. The TA method is applicable to any scenario [30] and can be conducted fruitfully with 8 to 12 participants [31].

Therefore, 10 physicians were recruited to participate in the CTA method of this project by a senior physician. The participants represented a range of experience, from a few weeks to over a decade of work in the department, and had varying levels of previous knowledge about the patient. They were tasked with gathering the relevant information to prepare the registration of a patient with advanced metastatic melanoma for a fictitious interdisciplinary tumor board discussion. This task was selected because it requires a broad range of information and system interactions while remaining a routine clinical activity. The task does not require additional support, even for less experienced participants, making it an ideal representative example of typical workflow patterns. The CTA method was conducted in March 2022 during standard working hours in a designated room with the same technical configuration used for typical task performance. For task completion, they had access to the same programs they use in everyday clinical practice on a clinical workstation. By using real patient data and limiting the timeframe for task completion to 15 minutes, a real-world scenario was established. Each participant was accompanied by 2 conductors. The 4 annotators brought complementary expertise to the analysis: 3 researchers from medical informatics (EMH, CLB, and JS) with experience in clinical workflow analysis and health IT, and 1 researcher from social psychology (AK) with expertise in qualitative research methods. This interdisciplinary composition ensured both clinical context awareness and methodological rigor. The conductors were largely unfamiliar to participants to minimize potential bias from preexisting relationships. The task and conditions were communicated by 1 conductor, while the other documented the process. During the task, conductors only reminded participants to "think aloud," meaning participants had to articulate each step in finding and preparing the information. Interactions with the software were captured along with the mentioned thoughts by screen and voice recording. The voices were anonymized through alteration and transcription. Patient identifying information was removed from the recordings. Afterwards, the transcripts were annotated inductively regarding the accessed information, programs used, and work steps conducted.

Annotation and Quantitative Analysis

Qualitative content analysis is a category-based method enabling researchers to analyze qualitative data both qualitatively and quantitatively [32]. It is the most commonly chosen method [33] and allows for contextual inspection with a broad perspective by comparing code occurrences and overlaps [34]. Furthermore, the view can be narrowed down to individual statements to reveal intents and relationships [33]. Inductive synthesis of the code system can be used to gain an understanding of the wording used and to prevent researchers' assumptions from influencing the results [35]. This is a structured process in which codes are extracted from 10%-50% of user phrases by exploring and summarizing key findings related to a predefined aspect. The so-called paraphrases are then termed and grouped into categories and subcategories to form a code system that represents one aspect of the context. The remaining texts are annotated using the initial system, which can further be extended if new aspects occur [35].

In the initial step of paraphrasing, 3 of the 10 transcripts were examined by 4 annotators (EMH, AK, CLB, and JS) from the fields of medical informatics and psychology. In 2 independent rounds of paraphrasing, the focus was on identifying the work steps or the information accessed. The annotations were labeled and grouped into separate, consistent catalogs with meaningful granularity. Based on the generated systems, the remaining 7 transcripts were annotated by one of the annotators, and the annotation catalog was extended where necessary. The system extensions were curated through mutual agreement and integrated into the existing annotation system. After a thorough review, the physician confirmed that the erased synonyms and optimized hierarchy were adequate.

The video footage of the screen recordings was used to examine the structure of the programs and the transitions between them during the workflow. This allowed for the transcripts to be aligned with the usage of the programs. The incorporation of annotations indicating the time span of program usage in the transcripts enables simultaneous access from all 3 analytical perspectives.

XSL•FO
RenderX

To enhance the reliability of the accessed information annotations, finally, each transcript was annotated a second time by one of the same 4 annotators. Thereby, it was ensured that the annotation phrases were used consistently among each annotator, that they were chosen to be meaningful [36], and that this resulted in an annotation catalog with consistent categories and subcategories [37]. Based on this, we determined an overall intercoder reliability of 70%. For detailed information on the intercoder reliability of each item, please refer to the Excel sheet tab "intercoder-reliability" in Multimedia Appendix 1.

To generate insights, the generated annotations were qualitatively analyzed regarding their occurrences and overlaps between the systems using MAXQDA (VERBI GmbH). The objective was to identify associations between relevant information and work steps. This was achieved by examining the number of work steps, their occurrence, and the time spans associated with them. These were then combined with the overlaps with the information annotations. In the context of program usage, the time spans were analyzed and combined with the number of substeps required to reach a specific document containing the information of interest.

SeqD

In the human mind, the world is conceptualized as an abstract mental representation grounded in knowledge, interpretation, imagination, and experiences [38-41]. In the domain of human-computer interaction (HCI) research, these cognitive structures are characterized as "mental models" [15]. Interactions with the surrounding world are based on these models. Therefore, it is necessary to reveal these patterns through the observation of workflow and interaction sequences during software development [42]. This analysis facilitates the comprehension of user interactions, task conceptualization [15], and contextual knowledge [43]. However, as each individual constructs their own mental model, these models differ for each person and can even be contradictory, incomplete, or vary in step sequence [15].

No definitive guideline is established for the exploration and design of cognitive workflows [37]. Accordingly, the SeqD was selected as the most appropriate representation for the context of the tumor board preparation. As part of contextual design, the SeqD [44] represents the sequence of steps that must be completed to accomplish a task. The use of user phrases from the transcripts to identify and define the work steps represents an inductive approach to annotation. Furthermore, this approach identifies the driving motivation behind actions and thereby enables external observers not only to label the work steps but also the underlying intent [44]. These insights facilitate designing context-integrated solutions [44] and detecting workflow interruptions that create barriers in task execution, thus allowing for the identification of potential solutions [19]. One challenge inherent to workflow analysis is the potential of missing or wrong task conceptualizations, which may arise from untrained personnel or misinterpretations [45]. In implementing this approach, the underlying information is made accessible and interpretable, and the inclusion of intents allows for the implementation of context-integrated solutions. The operational mental model described in HCI [15,44] is captured, depicting the relationship between work steps, information in focus, and system interactions in a SeqD. It thereby enables the implementation of context-integrated solutions that inherit physicians' actual cognitive workflows.

The SeqD was modeled based on the work steps identified during the inductive annotation process. A deeper exploration of the associated user statements was performed, which revealed the underlying intents connected with the work steps through thought-aloud clinical reasoning and context. To determine the sequence of work steps, the sequence of annotations was extracted and compared between the different transcripts. To generate an average workflow, all sequence numbers of work step occurrences were added and divided by their total number of occurrences in all transcripts. This mean was calculated because each participant used a unique sequence of work steps. The resulting number was used to sort the work steps in ascending order, with only 1 occurrence per work step.

Results

SeqD

The SeqD represents the typical workflow involved in preparing a tumor board for a patient with melanoma. The average was calculated by dividing the sum of the sequence numbers of each workflow step occurrence by the total number of occurrences (Table 1, column ranking). Accordingly, the SeqD incorporates each work step on a single occasion, even though the majority of these steps are repeated throughout the workflow. In order to ensure the representativeness of the SeqD, only work steps performed by at least 5 participants were included in the SeqD.

The average completion time was 13 minutes and 16 seconds. The minimum completion time was 8 minutes and 19 seconds, while the maximum was 14 minutes and 41 seconds. Despite the variety in the approaches adopted by the participants to complete the task, an underlying fundamental workflow structure can be observed, as illustrated in the SeqD in Figure 1. This shared structural pattern, which reflects physicians' common task conceptualization despite individual variations, forms the basis for context-appropriate software design. Overall, 14 work steps were identified, 10 of which are conducted by at least 5 participants and represented in the SeqD. These were enriched with underlying intentions derived from the following action in combination with the information in focus.

The process is initiated by entering the login credentials into the HIS system, which approximately takes 20 seconds. To reduce the time spent waiting, the majority of users simultaneously open an external text editor in "preparation." The physicians then proceed to "examine the past" by accessing the patient file and locating the most recent case summary. This summary is then copied into the external text editor, read, and used as a point of reference for an updated version. To evaluate the current status, the physicians conduct a review of the case knowledge by scanning for new diagnostic results. Accordingly, they examine the radiology reports of the last tumor assessment by computed tomography, magnetic resonance imaging, or ultrasound, laboratory values, the latest tumor board protocols, and the progress documentation. The insights gained are then

XSL•FO
RenderX

iteratively integrated into the case summary by adding new results and removing outdated information. The external editing of the case summary represents a significant portion of the overall process, accounting for approximately 28% of the total time. It is a highly iterative step, with an average of 3.5 repetitions, making it one of the most essential substeps within this workflow. The case summary is then copied into the tumor board registration form. Structured information is added, including the detailed TNM tumor classification and the disciplines required for consultation at the interdisciplinary tumor board. This step is executed by all participants and, with 27% of the total workflow time has a similar weight in the process as the text editing. The phase "updating the knowledge" has the greatest variety of work step sequences, as indicated by the approximately equal ranking scores of the work steps accessing primary reports. Especially, the steps of viewing different diagnostic reports vary in sequence and combination in different workflow approaches.

Table 1. Variables considered in the analysis of work steps (execution time [in seconds], total average over all participants, proportion to the complete workflow, execution times dependent on whether it has been part of the user's workflow, adapted proportion to the workflow, total number of users executing the works step, total number of executions of each work step, average of work step executions per workflow, and ranking [sum of sequences per occurrence divided by the total number of occurrences], which were used for the work step sequence in the sequence diagram.

Workflow step	Time execution of work step in minutes, mean (SD)	% time for work step of total workflow time	Time for work step if part of user's workflow in minutes, mean (SD)	% time for work step of total workflow time if part of user's workflow	Total number - users that executed work step	Total number - work step executed by user	Work step execution if work step was part of user's workflow, mean (SD)	Ranking sequence of work steps - sum sequences/ number occurrences
Login	00:20.2 (00:50.4)	3%	00:22.4 (00:52.5)	3%	9	11	1.2 (0.6)	2.4
Open patient files	00:27.5 (00:21.9)	4%	00:27.5 (00:21.9)	4%	10	15	1.5 (0.7)	4.3
Search for last progress documentation	00:16.6 (00:14.0)	2%	00:20.7 (00:12.6)	3%	8	9	1.1 (2.5)	5.7
View progress documentation notes	00:16.7 (00:26.4)	2%	00:27.8 (00:22.0)	4%	6	7	1.2 (0.4)	6.6
Check status of (expected) procedures	00:03.1 (00:09.4)	0%	00:31.2 (00:00.0)	5%	1	1	1.0 (0.0)	7.0
Internal processing of the case summary	00:03.9 (00:11.7)	1%	00:39.1 (00:00.0)	6%	1	1	1.0 (0.0)	7.0
External editing of the case summary	02:28.9 (01:38.1)	22%	03:06.1 (01:27.4)	28%	8	28	3.5 (1.3)	8.9
View primary findings	01:12.5 (01:13.6)	11%	01:20.5 (01:10.7)	12%	9	21	2.3 (1.1)	10.1
View tumor conference protocol	00:26.8 (00:41.9)	4%	00:38.2 (00:40.5)	6%	7	10	1.4 (0.7)	10.2
View progress documentation	00:50.0 (00:45.2)	7%	01:02.5 (00:42.2)	9%	8	13	1.6 (0.9)	10.2
Interruption	00:25.8 (01:16.9)	4%	02:08.9 (02:01.2)	19%	2	2	1.0 (0.0)	12.5
Fill out tumor board registration	03:02.3 (01:21.3)	27%	03:02.3 (01:21.3)	27%	10	24	2.4 (1.5)	13.0
Formulate question	01:16.1 (00:28.4)	11%	01:16.1 (00:28.4)	11%	10	14	1.4 (0.5)	13.6
Review external knowledge sources	00:02.0 (00:05.8)	0%	00:19.5 (00:00.0)	3%	1	1	1.0 (0.0)	22.0

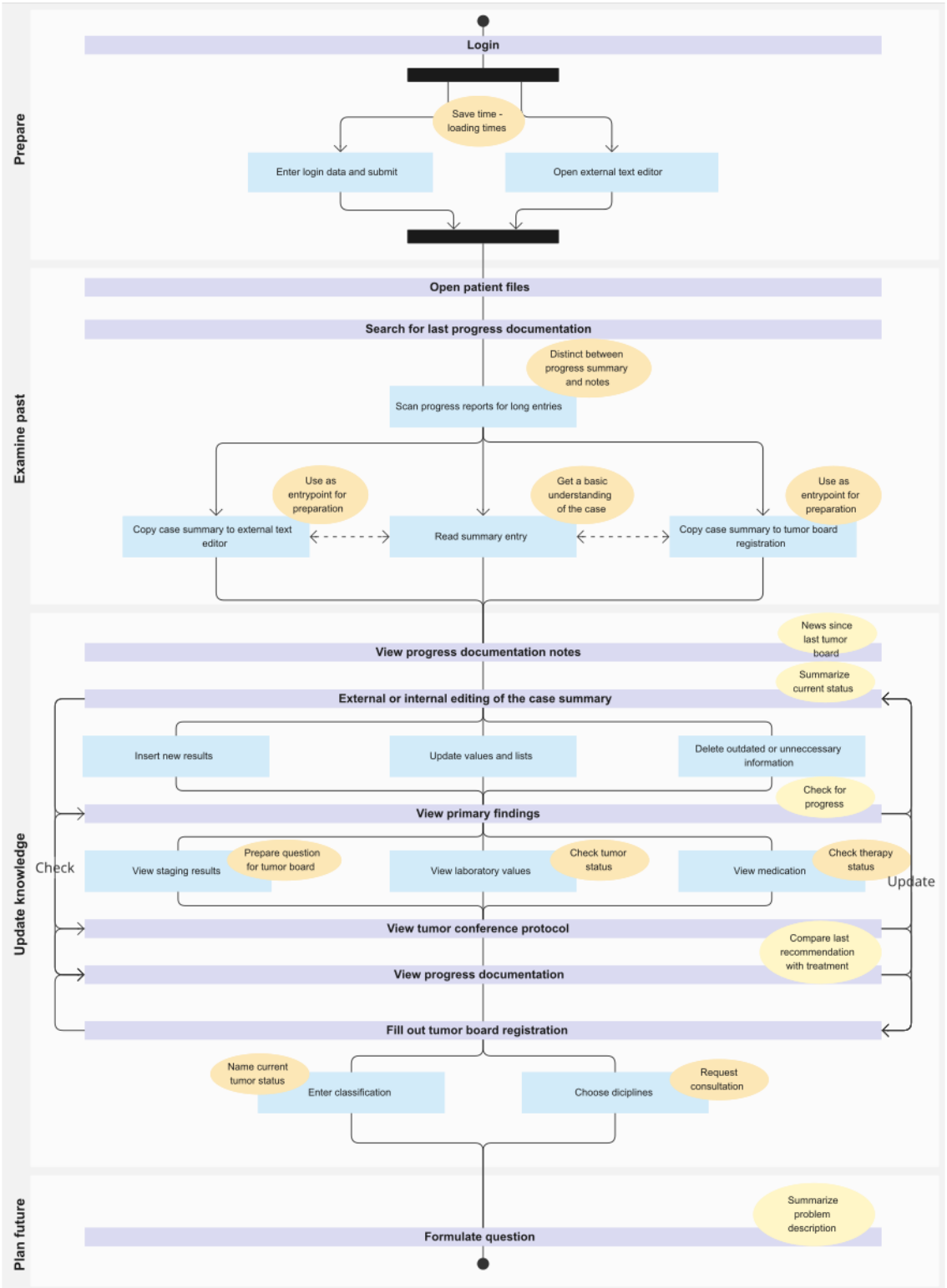
The final phase of the tumor board preparation process is aimed at establishing a follow-up care plan. By formulating the question to be addressed in the tumor board, the most recent principal findings of the tumor's progression are summarized, and potential treatment options are evaluated. Furthermore, deeper insights from other disciplines into their diagnostic results can be requested to gain a more comprehensive understanding of the tumor status and treatment options.

After the workflow completion, tumor documenters use the tumor board registration form to transform the unstructured free text into structured data points by populating the fields in a structured form. The structured data points are used to generate a tumor board protocol that includes not only the summary but

XSL•FO
RenderX

also the structured information. However, as the physicians are unable to access the structured information in the HIS, there is currently no benefit derived from this data transformation. This results in the workflow described here of iteratively rewriting the case summary rather than directly extracting the information from primary sources. In essence, the workflow can be understood as a single, extensive workaround due to the distribution of important information, resulting in a free-text summary of the patients' status documented as a progress report. The section "Programs Used and Collection of Information" provides a more detailed examination of the challenges associated with gathering data from the ISs.

Figure 1. The sequence diagram depicts the 4 phases (white boxes) and the associated work steps (lilac bars) involved in preparing for the tumor board meeting. Additionally, it illustrates the subwork steps (blue rectangles), and underlying intents (orange ovals=sub-steps and yellow ovals=work step) associated with this process. The black bars illustrate parallel execution of the working steps.



XSL•FO
RenderX

To enhance the reliability and validity of the constructed workflow model, 2 experienced physicians from the department critically reviewed and validated the work steps, their sequence, and associated intentions depicted in the SeqD. The model was approved in its final form with only minor terminological refinements.

Work Step and Data

In parallel with the execution of the work steps, the information in focus was annotated inductively. By examining the areas of overlap between the 2 annotation categories, the connections between them were revealed. Figure 2 illustrates these overlaps, highlighting the work steps in the inner circle combined with the relevant information in the outer circle. The size of the fragments is reflective of the number of identified overlaps. For the sake of clarity and focus, associations with fewer than 5 overlaps have been excluded from the graphic. For detailed information on the interplay of program usage and information access per workflow step, refer to the Excel sheet tab: "work steps" in Multimedia Appendix 1.

During the "updating knowledge" phase, several data sources were addressed. Most time was spent reviewing primary reports, which focused on radiology staging reports, including magnetic resonance imaging and computed tomography results. Review of progress notes and progress reports focused on options discussed with the patient, ongoing treatment, and associated side effects. Previous tumor board protocols were used to compare whether the treatment was in line with the last result and treatment suggestions.

As highlighted in section "SeqD" under "Results," the "external editing of the case summary" and "completion of the tumor board registration" are the central steps in the process. Not only do these steps take the longest amount of time, but they also require the most information to be in focus. When updating the case summary, the information collected, such as staging results and treatment progress, was interwoven. In addition, the diagnosis and documented infusions came into focus.

When using these findings to "fill out the tumor board registration," the focus shifted to the diagnosis and especially the classification of the tumor. In addition, therapies and the decision on external disciplines to be consulted during the tumor board were identified as relevant. When planning the formulation of the question, the focus was on the therapies already applied and the overall clinical stage of the tumor.

Figure 2. The work steps, which are located within the inner circle, are accompanied by the associated information, which is presented in the outer circle. Additionally, the number of code-overlaps in the transcripts is indicated by the numbers within the outer circle. Only those overlaps occurring in transcripts at least 5 times are included. CT: computed tomography; MRI: magnetic resonance imaging and computed tomography.



XSL•FO
RenderX

Programs Used and Collection of Information

The study was conducted in a real-world setting, including the usage of the standard programs during the process of registration of a patient with melanoma for the tumor board. As a result, it was possible to gain insights into the programs in use and their purpose of usage. This revealed that the information in the medical IS was distributed across different reports, was unstructured, and needed to be accessed sequentially to gain an overview of the complete patient's status. The use of overlays hindered the viewing of information simultaneously, resulting in the need for multiple instances of the HIS to allow for easier switching of information sources.

On average, participants accessed 5 distinct programs or submodules within the HIS, not including cases where multiple instances were used. The range of variance in the number of programs used is between 4 and 7. Excluding the submodules, the average was 3 programs, with 1 in the minimum and 4 in the maximum. This usage of multiple programs resulted in 25.3 switches between programs, with a variance of 14 in the minimum and 34 in the maximum. This again highlights the distribution of data and iterative approach in summarizing the patient's status. Figure 3 lists how many participants used each program, the average time spent using it, and its proportion to the whole workflow. It can be revealed that the HIS was used by 90% (9/10) of the participants as the only information source, while 1 participant also used the radiology IS and laboratory IS. The main time was spent using the workspace, a submodule of the HIS, where the diagnostic reports can be accessed. This correlated with the work steps referring to viewing results of different kinds. The gathered information was transferred to their own case summary, for which 90% (9/10) of the participants used an external text editor. Entering the data into the tumor board registration still takes about 20% of the total time, even though the case summary was (mostly) already written in the external tool.

Figure 3. The proportion of program usage (purple) in the process of tumor board preparation for a melanoma patient is presented alongside its proportion of the total time required to complete the task (orange). HIS: hospital information system; LIS: laboratory information system; RIS: radiology information system.

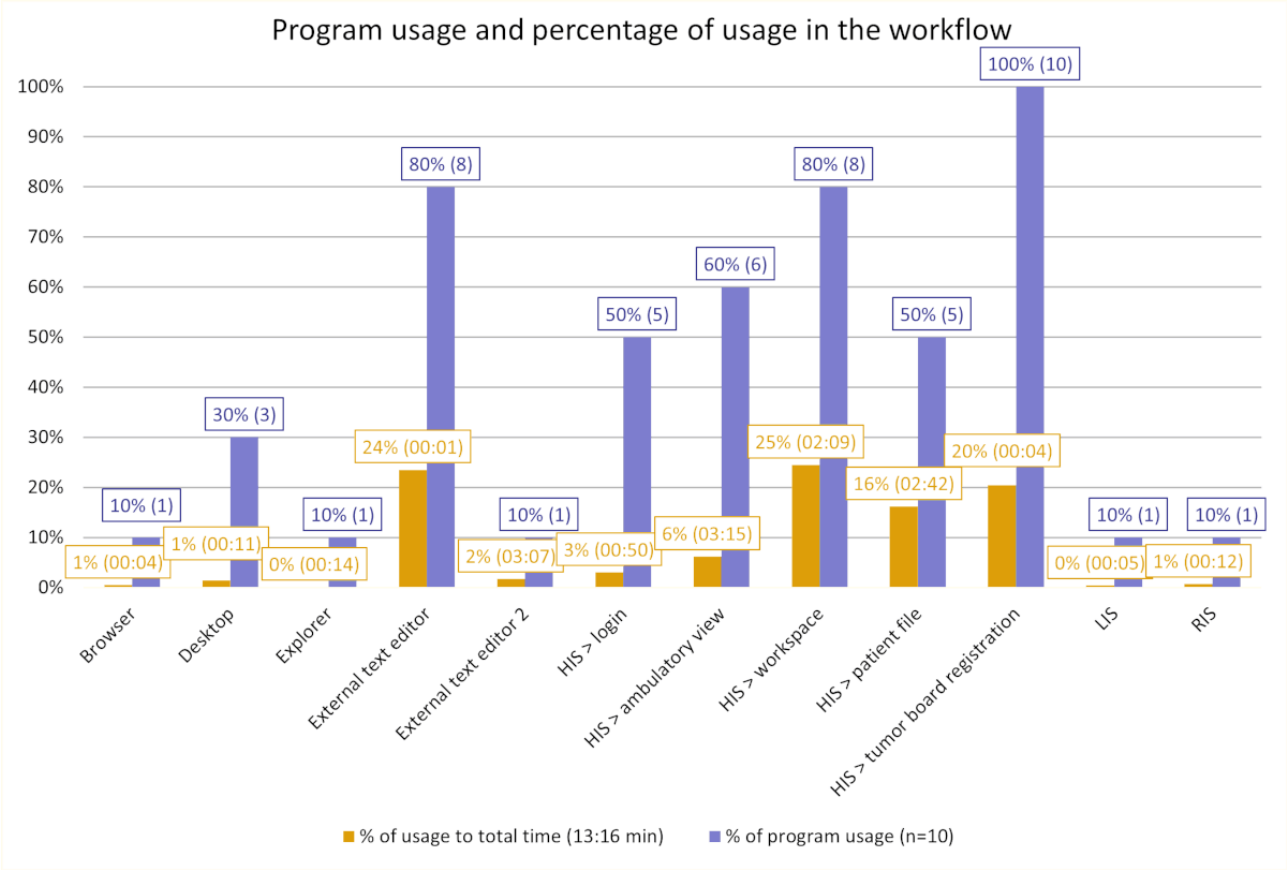


Table 2 correlates the programs and submodules with the accessed documents and indicates the number of interactions with the program required to access those. At the initial stage of the workflow, the physician must complete 3 steps to log in and 9 steps to select a patient file. As previously described, multiple instances of the HIS, using either the same or different submodules, were accessed. Examining each instance, the participants were required to complete all 9 steps of patient identification. Therefore, the table lists the number of steps required at the beginning of the workflow and after selecting a patient file. This overview demonstrates that a minimum of 16 interactions are required to access essential information from the HIS, with a minimum of 2 additional interactions necessary to access further information when the patient file is already open.

The study revealed the users' task conceptualization regarding the work steps performed, programs used, and data accessed through combining a SeqD with qualitative analysis of the annotations. This demonstrates that the workflow is significantly

XSL•FO
RenderX

influenced by the distribution of data. A considerably large amount of time is invested in system interactions rather than information interpretation, as evidenced by the interaction steps required to access each information. The creation of a case summary is an effective method for summarizing this distributed information. It provides an overview of the patient's status and serves as a foundation for further discussions regarding the patient's treatment plan in the tumor board.

Table 2. List of the accessed documents during the process of tumor board registration of a melanoma patient, along with the used program or submodule and the number of substeps required to access each document from the initial starting point. The final column indicates the number of steps needed to access the document if the patient file was selected and opened beforehand.

Document	Source	Steps to open from start	Steps to open from patient file
Own case summary	• External text editor	6	6
Laboratory values	• LIS[a]	16	3
Laboratory values	• HIS[b]>workspace	17	2
Diagnostic reports	• HIS>workspace	21-23	5-6
Progress documentation	• HIS>workspace • HIS>ambulatory view	19-20	3
Appointments	• HIS>ambulatory view	17	2
Scanned documents	• HIS>Workspace • HIS>ambulatory view	20	4
Skin images	• HIS>workspace • HIS>ambulatory view	20	4
Infusion protocols	• HIS>workspace	19-21	3
Tumor board protocols	• HIS>patient file	21-22	4-5
Radiology reports	• HIS>patient file • HIS>workspace	20	4
Tumor board registration	• HIS>ambulatory view	22	7
Publications	• Internet (browser)	4+	4+
Standard operating procedures	• Intranet (browser)	5+	5+
Dosing scheme	• Intranet (browser)	5+	5+
Call list	• Intranet (browser)	5+	5+

[a]LIS: laboratory information system.

[b]HIS: hospital information system.

Pain Points

A synthesis of the aforementioned results reveals seven fundamental pain points in the HIS interaction, which hinder workflow efficiency: (1) time- and step-intensive login and patient selection, (2) multiple instances to reach data end points simultaneously, (3) fragmentation of data across systems, (4) hierarchical modularization of information sources, (5) user interface structure hindering view, (6) repetitive data insertion, and (7) inappropriate input fields.

At the initiation of the workflow, the login process requires 3 steps and approximately 20 seconds of waiting time before a patient can be selected in 9 steps. These 12 steps consume approximately 7% of the total time frame, despite not substantially enhancing task completion. To allow simultaneous access to different information modules and the tumor registration form, some physicians use multiple instances, which again require patient selection. Due to the use of disparate specialized IS by various departments, information is fragmented across a range of systems. Given their lack of interconnection, these systems can produce inconsistent data due to variations in data curation methodologies. The processes of copying, pasting, and logging in to each system must be repeated, resulting in a need for increased operational complexity. Additionally, the information contained within an IS is organized into a module-like hierarchy. This frequently generates new

XSL•FO
RenderX

tabs and windows, necessitating numerous system interactions to navigate between information: 2 to 7 interactions if the patient has been selected, or 16 to 23 when starting with the login process. These windows and tabs also often hinder the simultaneous view of information. Furthermore, switching between these modules requires time and interactions, forcing users to maintain awareness of the various windows and tabs that have been opened. This circumstance leads to an average of 25.3 program switches. The data fragmentation and its predominantly unstructured nature induce repetitive data insertion into different programs. The process of copying and pasting, accounting for approximately 27% of the total task completion time, is the only method that allows users to center all necessary information on the tumor board registration form. In addition to fragmentation across windows, inappropriate field sizing, and absent text editing features compel the use of external text editors to curate the summary. As this text editing step accounts for approximately 28% of total task completion time, developing an appropriate user interface design for this step is critically important. An effective incorporation of this feature could lead to a significant reduction in the number of program switches required, thereby enhancing the efficiency of the overall process.

Discussion

Principal Findings

This study aimed to reveal the task conceptualization of physicians preparing patients with melanoma for tumor board discussion by identifying their work steps, information priorities, and data sources used. Our findings demonstrate that the workflow consists of 4 phases with 10 work steps, requiring approximately 13 minutes and an average of 25 program switches to complete. The analysis revealed 7 fundamental pain points, including time-intensive login and patient selection, need for multiple system instances, data fragmentation across systems, and hierarchical information modularization. Additionally, interface structures hinder simultaneous viewing, repetitive data insertion, and inappropriate input fields. Collectively, these pain points result in approximately 55% of workflow time being dedicated to system interactions and manual data compilation rather than clinical decision-making. Patient data distributed across nested submodules necessitated physicians to manually extract and compile information into a free-text case summary using external editors. This workflow represents an extensive workaround addressing data fragmentation.

Our work provides a comprehensive overview of the physicians' cognitive workflow through simultaneous examination of work steps, programs used, and information in focus. This approach aligns with established HCI methodology for revealing "mental models" [15,44] by displaying users' operational understanding of task execution within systems. While individual physicians demonstrated variations in step sequences, the underlying task conceptualization remained consistent. This shared cognitive approach underlies context-integrated software design, facilitating accommodation of individual variation while ensuring adherence to common workflow logic. To the best of our knowledge, no other study has provided such an integrated perspective on these 3 aspects of the medical treatment process, offering actionable insights for designing clinical IS that better support workflows at the point of care.

Other literature emphasizes the significance of cognitive task analysis as a common methodology for investigating the domain and usage of technology in order to develop a mental model of a context. Nevertheless, it is evident that the majority of projects do not incorporate direct user involvement [46].

As a positive example, Kilsdonk et al [47] used the TA method in combination with a systematic literature review to design a CDSS. Their work was based on insights into the users' mental model in the context of CDSS usage in the follow-up of childhood cancer survivors. In contrast to our work, their focus was on cognitive information processing patterns rather than revealing the complete workflow and IS interactions [48]. The information processing revealed during the follow-up period indicates a comparable data focus to that observed in melanoma care. The "inclusion criteria" for further decisions, namely the cancer diagnosis and treatment, are reflected in the formulation of the question, with a particular focus on tumor classification and treatment status. In contrast, the radiological staging, which is a "screening procedure," is the focus of the case summary update for tumor board preparation rather than the potential result of the decision. These differences can be explained by the different stages of tumor treatment, as our focus was on a process to adjust treatment, while the aforementioned study refers to follow-up decisions. Overall, the same information is the primary focus in both processes, although interpreted in an alternating manner during decision-making in the distinct contexts.

Having a focus on the interactions with medical IS, Saitwal et al [6] conducted cognitive task analysis using the methodology of "goals, operators, methods, and section" rules [49] in combination with the keystroke level model [50]. They evaluated their application "AHLTA" (Armed Forces Health Longitudinal Technology Application) for the completion of 14 tasks [6]. With this approach, it is possible to break down work steps into their smallest substeps by applying "goals, operators, methods, and section" rules and estimate the cognitive load and time needed to complete a task based on execution times defined by the keystroke level model. Zayim et al [48] used a similar methodological approach but conducted further reflection on the changes observed in gesture-based app control in comparison with conventional control methods for the completion of 10 tasks on their mobile app. As we had direct insight into the interactions with the programs, we were able to extract average times directly from the real-world setting. The findings of our study indicate that a considerable amount of time and effort is required to obtain the information necessary to fulfill the specific task. These interactions are characterized by a high proportion of mental operators, with nearly every step involving cognitive processes, such as identifying and scanning for the correct menu option and relevant information. This excludes the actual decision-making and summary processes. In comparison with the 2 studies mentioned beforehand, which identified about 50% [6] and 73% [48] of mental operators, the cognitive load

XSL•FO
RenderX

associated with the IS interactions can be interpreted as being particularly high.

In their study, Barber et al [19] used a knowledge audit [51] for cognitive task analysis, thereby uncovering the mental model of physicians and nurses in the context of cirrhosis care. The goal was to examine the impact of workflow on the mental model. The study demonstrated that discontinuity and barriers are reflected in the identified models. In our own work, we have found that the distribution of necessary information and its availability, typically in free text, has led to the need to align all information in a central case summary.

Pluyter et al [52] demonstrated the significance of a context-specific design of CDSS, enabling physicians to establish a shared mental model when treating patients with lung cancer, thereby optimizing treatment selection. The integration of context into the decision-making process has been shown to enhance decision confidence and quality. This improvement is attributed to the enhanced and more precise understanding of the case. In the context of our research, the observed variations in tumor board registration preparation reflect individual approaches while maintaining a shared underlying task structure, as previously indicated [19]. The workflow model conceptualized in this study can be integrated into software development processes, ensuring the synchronization of mental models independent of work experience or previous case knowledge.

The findings of this study imply 2 main aspects. On the one hand, it revealed the pain points that physicians experience when interacting with the IS, and on the other hand, the advantages and insights that were gained by exploring the workflow of a context in detail. A thorough exploration of the context helped identify usability issues in the system in use and showed how small decisions in development, such as menu structures, can cause large practical problems.

The inefficiencies and data fragmentation that have been identified present significant challenges; however, patient care continues to function through physicians' adaptation and development of workarounds. However, it is important to note that "functional" does not necessarily equate to "optimal." The current system remains operational, but it incurs a considerable cost: approximately 55% of workflow time is dedicated to system interactions rather than clinical reasoning. This could be a contributing factor to cognitive load [5,6] and potentially burnout [8,9] as revealed in other studies. In order to compensate for system deficiencies, physicians have been known to use experience-based workarounds, such as maintaining parallel documentation in external editors and memorizing navigation paths through nested modules. While these adaptations are indeed effective in maintaining care delivery, they represent accepted trade-offs that normalize inefficiency and obscure the true potential for improvement. The deep insight into the workflow made it clear that a well−thought-out interaction design is required, considering the relationship between the task at hand and the information in focus.

Further research needs to address how this information can be visualized other than in menu structures to allow for the best contextual integration and adaptation to different work steps [53-55]. The quality of these integrations should be monitored by continuous user feedback to ensure that the application allows clinicians to focus on the patient's information rather than on system structure and interactions [10].

Limitations of this study are questions that may arise regarding the transferability of the findings. Given that the workflow represents a comprehensive workaround addressing data distribution, it is important to consider how other hospitals have addressed this challenge and how the process might differ without such data distribution. While the insights into workflow steps and programs primarily address common IS interaction issues, the information in focus is transferable. Treatment standardization through guidelines and standard operating procedures, combined with similar findings in related studies [47], supports this transferability. This also indicates the transferability of the results to other disciplines. In particular, the approach and information revealed in this study will be highly applicable to the preparation of tumor boards for other cancer entities. The same information will also be relevant for other tasks related to cancer treatment. For disciplines not related to cancer, the pain points revealed in this study will also be transferable, highlighting issues to be avoided in IS development.

A further limitation of the study is its qualitative nature, which was conducted at a single hospital with a limited sample size of 10 participants. As a result, no significant differences were observed in task completion time or work step sequences between participants with varying levels of working experience, gender, or previous case knowledge. However, this absence of observable differences may be attributable to the constrained sample size that is inherent to qualitative research designs. The methodology was designed to capture workflow patterns and time allocation, rather than to systematically measure dimensions such as cognitive load, error rates, or interruptions. The use of different methodological approaches and larger sample sizes would be required to yield meaningful quantitative insights in these cases. While the findings indicate significant system interactions (55% of workflow time) and elevated operator demands, the formal measurement of these phenomena using validated instruments was not within the scope of this study. Subsequent multicenter investigations with increased sample sizes would be essential to methodically examine the potential influence of factors, such as experience and case knowledge, on workflow patterns. These studies could also quantify cognitive load, error rates, and workflow interruptions across institutions with varying IS configurations.

Conclusion

In summary, this study makes 3 significant contributions to the field of IS development in medical informatics. First, the identification of pain points in the interaction revealed the areas that require attention in the development of new applications. Furthermore, it demonstrated the importance of context exploration and development that is centered on context integration and user feedback. When designing and implementing new applications based on the insights gained, physicians can obtain a focused view of the necessary information. This can reduce the cognitive effort required to

XSL•FO
RenderX

interact with the program and possibly avoid time-consuming and potentially life-threatening workarounds.

Second, the insight into workflow, particularly the information in focus, provides transferable knowledge. While workflow steps may differ across hospitals, the underlying intentions associated with these steps remain consistent. Therefore, the combination of these intentions and the information in focus can serve as a basis for context-sensitive software development across institutions.

Third, this study addresses the discrepancy between established user experience principles and their application in health care IS development. The users' conceptual understanding of task execution, which is known in HCI as "mental models," is recognized as a foundational element in the field of user experience design [15,42,43]. However, their systematic exploration through workflow analysis in the medical domain remains limited. The findings of our study demonstrate the value of integrating the TA method with cognitive task analysis and SeqDs in revealing physicians' cognitive workflows and task conceptualizations. This approach captures not isolated cognitive structures but the observable manifestation of how physicians understand and execute clinical tasks within IS constraints. Integrating this model in point-of-care software design can enhance patient treatment by shared cognitive approaches across physicians, regardless of their work experience or previous case knowledge.

Acknowledgments

We thank the physicians of the Department of Dermatology (University Hospital Essen, Germany) for their dedicated participation in the usability studies. We thank Katharina Zeiner (Siemens, Munich) for the mentoring and fruitful discussions on user experience.

Funding

This work was funded by PhD grants from the DFG Research Training Group 2535 "Knowledge- and data-based personalization of medicine at the point of care (WisPerMed)."

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

EMH: Conceptualization, Methodology, Investigation, Data Curation, Visualization, Formal analysis, Writing- Original draft preparation
AK: Methodology, Data curation, Investigation, Formal analysis, Writing - Review & Editing
JS and CLB: Methodology, Data Curation, Writing - Review & Editing
GCL, EL, and DS: Conceptualization, Resources, Writing - Review & Editing
SS: Supervision, Conceptualization, Methodology, Writing - Review & Editing.

Conflicts of Interest

GCL reports advisory role or has received honoraria/travel costs from: Novartis, Pierre Fabre, Merck and Sunpharma.

Multimedia Appendix 1
Work steps, programs, code systems, and intercoder reliability.
[XLSX File (Microsoft Excel File), 67 KB - cancer_v12i1e79294_app1.xlsx]

References

1. RAND Europe, Open Evidence, BDI Research. In: Benchmarking deployment of eHealth among general practitioners. eHealth among general practitioners: Publications Office; 2018.
2. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
3. Johnson C, Johnston D, Crowley P, Culbertson H, Rippen H, Damico D, et al. EHR usability toolkit // EHR Usability Toolkit: A Background Report on Usability and Electronic Health Records: A background report on usability and electronic health records. Rockville, Maryland: AHRQ Publication; 2011.
4. van Offenbeek MAG, Vos JFJ, van den Hooff B, Boonstra A. When workarounds aggravate misfits in the use of electronic health record systems. Information Systems Journal 2023;34(2):293-326. [doi: 10.1111/isj.12478]
5. Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, Middleton B. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. J Biomed Inform 2012;45(6):1202-1216 [FREE Full text] [doi: 10.1016/j.jbi.2012.09.002] [Medline: 22995208]

XSL•FO
RenderX

6. Saitwal H, Feng X, Walji M, Patel V, Zhang J. Assessing performance of an Electronic Health Record (EHR) using Cognitive Task Analysis. Int J Med Inform 2010;79(7):501-506. [doi: 10.1016/j.ijmedinf.2010.04.001] [Medline: 20452274]
7. Akbar F, Mark G, Prausnitz S, Warton EM, East JA, Moeller MF, et al. Physician Stress During Electronic Health Record Inbox Work: In Situ Measurement With Wearable Sensors. JMIR Med Inform 2021;9(4):e24014 [FREE Full text] [doi: 10.2196/24014] [Medline: 33908888]
8. Shanafelt TD, Dyrbye LN, West CP. Addressing Physician Burnout: The Way Forward. JAMA 2017;317(9):901-902. [doi: 10.1001/jama.2017.0076] [Medline: 28196201]
9. Melnick ER, Dyrbye LN, Sinsky CA, Trockel M, West CP, Nedelec L, et al. The Association Between Perceived Electronic Health Record Usability and Professional Burnout Among US Physicians. Mayo Clin Proc 2020;95(3):476-487 [FREE Full text] [doi: 10.1016/j.mayocp.2019.09.024] [Medline: 31735343]
10. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
11. Osheroff J, Teich J, Levick D, Saldaña L, Velasco F, Sittig D, et al. Improving outcomes with clinical decision support: An implementer's guide. Second edition. Boca Raton, FL: HIMSS Publishing an imprint of Taylor and Francis; 2012.
12. Evans RS. Electronic Health Records: Then, Now, and in the Future. Yearb Med Inform 2016;Suppl 1(Suppl 1):S48-S61 [FREE Full text] [doi: 10.15265/IYS-2016-s006] [Medline: 27199197]
13. Middleton B, Bloomrosen M, Dente MA, Hashmat B, Koppel R, Overhage JM, American Medical Informatics Association. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. J Am Med Inform Assoc 2013;20(e1):e2-e8 [FREE Full text] [doi: 10.1136/amiajnl-2012-001458] [Medline: 23355463]
14. Kilsdonk E, Peute LW, Jaspers MWM. Factors influencing implementation success of guideline-based clinical decision support systems: A systematic review and gaps analysis. Int J Med Inform 2017;98:56-64. [doi: 10.1016/j.ijmedinf.2016.12.001] [Medline: 28034413]
15. Norman DA. Some Observations on Mental Models. In: Gentner D, Stevens AL. editors. Mental models. New York, London: Psychology Press Taylor & Francis Group; 1983:7-14.
16. Hartmann EM, Küper A, Swoboda J, Lodde GC, Livingstone E, Beckmann CL, et al. Requirements for a dashboard optimized for melanoma patient care through user-centered context exploration. Sci Rep 2024;14(1):1-10 [FREE Full text] [doi: 10.1038/s41598-024-67857-2] [Medline: 39080418]
17. Norouzi S, Galavi Z, Ahmadian L. Identifying the data elements and functionalities of clinical decision support systems to administer medication for neonates and pediatrics: a systematic literature review. BMC Med Inform Decis Mak 2023;23(1):263 [FREE Full text] [doi: 10.1186/s12911-023-02355-5] [Medline: 37974195]
18. Kushniruk AW, Patel VL. Cognitive evaluation of decision making processes and assessment of information technology in medicine. Int J Med Inform 1998;51(2-3):83-90. [doi: 10.1016/s1386-5056(98)00106-3] [Medline: 9794325]
19. Barber T, Toon L, Tandon P, Green LA. Exploring provider roles, continuity, and mental models in cirrhosis care: A qualitative study. Can Liver J 2023;6(1):14-23 [FREE Full text] [doi: 10.3138/canlivj-2022-0020] [Medline: 36908575]
20. Blijleven V, Hoxha F, Jaspers M. Workarounds in Electronic Health Record Systems and the Revised Sociotechnical Electronic Health Record Workaround Analysis Framework: Scoping Review. J Med Internet Res 2022;24(3):e33046 [FREE Full text] [doi: 10.2196/33046] [Medline: 35289752]
21. Rundo L, Pirrone R, Vitabile S, Sala E, Gambino O. Recent advances of HCI in decision-making tasks for optimized clinical workflows and precision medicine. J Biomed Inform 2020;108:103479 [FREE Full text] [doi: 10.1016/j.jbi.2020.103479] [Medline: 32561444]
22. 2024. American Cancer Society. URL: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/2024-cancer-facts-figures.html [accessed 2026-01-10]
23. 2024. NHS England Digital. 2024. Cancer Survival in England, cancers diagnosed. 2016. URL: https://digital.nhs.uk/data-and-information/publications/statistical/cancer-survival-in-england/cancers-diagnosed-2016-to-2020-followed-up-to-2021# [accessed 2023-02-16]
24. 2019. Office for National Statistics. 2019. Cancer survival in England - adults diagnosed. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancersurvivalratescancersurvivalinenglandadultsdiagnosed [accessed 2026-01-10]
25. Wolchok JD, Chiarion-Sileni V, Gonzalez R, Grob J, Rutkowski P, Lao CD, et al. Long-Term Outcomes With Nivolumab Plus Ipilimumab or Nivolumab Alone Versus Ipilimumab in Patients With Advanced Melanoma. J Clin Oncol 2022;40(2):127-137 [FREE Full text] [doi: 10.1200/JCO.21.02229] [Medline: 34818112]
26. Lewis C. Using the? thinking-aloud? method in cognitive interface design. method in cognitive interface design 1982.
27. Kip H, Keizer J, da Silva MC, Beerlage-de Jong N, Köhle N, Kelders SM. Methods for Human-Centered eHealth Development: Narrative Scoping Review. J Med Internet Res 2022;24(1):e31858 [FREE Full text] [doi: 10.2196/31858] [Medline: 35084359]
28. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: A scoping review. Int J Med Inform 2019;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]

XSL•FO
RenderX

29. Wronikowska MW, Malycha J, Morgan LJ, Westgate V, Petrinic T, Young JD, et al. Systematic review of applied usability metrics within usability evaluation methods for hospital electronic healthcare record systems: Metrics and Evaluation Methods for eHealth Systems. J Eval Clin Pract 2021;27(6):1403-1416 [FREE Full text] [doi: 10.1111/jep.13582] [Medline: 33982356]
30. Jacobsen J, Meyer L. Praxisbuch Usability und UX: Was alle wissen sollten, die Websites und Apps entwickeln. 3., aktualisierte und erweiterte Auflage. Bonn: Rheinwerk Computing; Rheinwerk Verlag; 2022.
31. Hwang W, Salvendy G. Number of people required for usability evaluation. Commun. ACM 2010;53(5):130-133. [doi: 10.1145/1735223.1735255]
32. Mayring P, Fenzl T. Qualitative Inhaltsanalyse. In: Baur N, Blasius J. editors. Handbuch Methoden der empirischen Sozialforschung. 2. Aufl. // 2. Auflage. Wiesbaden: Springer VS; Springer Fachmedien Wiesbaden; Springer; 2019:633-648.
33. Mayring P. Qualitative inhaltsanalyse? abgrenzungen, spielarten, weiterentwicklungen. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, Vol 20, No 3 (2019): Qualitative Content Analysis I 2019. [doi: 10.63485/599rn-wg167]
34. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
35. Mayring P. Qualitative Inhaltsanalyse: Grundlagen und Techniken. 12., vollständig überarbeitete und aktualisierte Aufl. Weinheim: Beltz; 2015.
36. Lindgren B, Lundman B, Graneheim UH. Abstraction and interpretation during the qualitative content analysis process. Int J Nurs Stud 2020;108:103632. [doi: 10.1016/j.ijnurstu.2020.103632] [Medline: 32505813]
37. Graneheim UH, Lindgren B, Lundman B. Methodological challenges in qualitative content analysis: A discussion paper. Nurse Educ Today 2017;56:29-34. [doi: 10.1016/j.nedt.2017.06.002] [Medline: 28651100]
38. Johnson-Laird P, Byrne R. Deduction: Lawrence Erlbaum Associates. Deduction: Lawrence Erlbaum Associates, Inc; 1991.
39. Nersessian NJ. In the Theoretician's Laboratory: Thought Experimenting as Mental Modeling. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1992;1992(2):291-301. [doi: 10.1086/psaprocbienmeetp.1992.2.192843]
40. Ford M, Johnson-Laird P. Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Cambridge, Massachusetts,: Harvard University Press; 1985.
41. Craik K. The nature of explanation. Cambridge: The Univesity Press; 1944.
42. Krug S. Don't make me think!: Web & Mobile Usability - das intuitive Web. Dritte Auflage. Cambridge: MIT Press; 2014.
43. Gentner D, Stevens A. editors. Mental models. New York, London: Psychology Press Taylor & Francis Group; 1983.
44. Holtzblatt K. Contextual design: Design for life. Second edition. Cambridge, MA: Morgan Kaufmann; 2017:9780128011362.
45. Walshe N, Drennan J, Hegarty J, O'brien S, Crowley C, Ryng S, et al. A cognitive task analysis of final year nursing students' situation awareness in simulated deteriorating patient events: a mixed methods study. Nurs Open 2024;11(4):e2154 [FREE Full text] [doi: 10.1002/nop2.2154] [Medline: 38606846]
46. Busse TS, Jux C, Laser J, Rasche P, Vollmar HC, Ehlers JP, et al. Involving health care professionals in the development of electronic health records: scoping review. JMIR Hum Factors 2023;10:e45598 [FREE Full text] [doi: 10.2196/45598] [Medline: 37428524]
47. Kilsdonk E, Peute LW, Riezebos RJ, Kremer LC, Jaspers MWM. Uncovering healthcare practitioners' information processing using the think-aloud method: from paper-based guideline to clinical decision support system. Int J Med Inform 2016;86:10-19 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.11.011] [Medline: 26725690]
48. Zayim N, Yıldız H, Yüce YK. Estimating cognitive load in a mobile personal health record application: a cognitive task analysis approach. Healthcare informatics research 2023;29(4):367-376 [FREE Full text] [doi: 10.4258/hir.2023.29.4.367] [Medline: 37964458]
49. Card S, Moran T, Newell A. The psychology of human-computer interaction. Reprinted. London: CRC Press Taylor and Francis Group; 2008.
50. Card SK, Moran TP, Newell A. The keystroke-level model for user performance time with interactive systems. Commun. ACM 1980;23(7):396-410. [doi: 10.1145/358886.358895]
51. Crandall B, Klein G, Hoffman R. Hoffman RR. Working Minds. Working Minds: The MIT Press; 2006.
52. Pluyter JR, Jacobs I, Langereis S, Cobben D, Williams S, Curfs J, et al. Looking through the eyes of the multidisciplinary team: the design and clinical evaluation of a decision support system for lung cancer care. Transl Lung Cancer Res 2020;9(4):1422-1432 [FREE Full text] [doi: 10.21037/tlcr-19-441] [Medline: 32953514]
53. Tory M, Möller T. Human factors in visualization research. IEEE Trans Vis Comput Graph 2004;10(1):72-84. [doi: 10.1109/TVCG.2004.1260759] [Medline: 15382699]
54. Padilla LM, Creem-Regehr SH, Hegarty M, Stefanucci JK. Decision making with visualizations: a cognitive framework across disciplines. Cogn Res Princ Implic 2018;3:29 [FREE Full text] [doi: 10.1186/s41235-018-0120-9] [Medline: 30238055]
55. Shneiderman B, Plaisant C, Hesse BW. Improving healthcare with interactive visualization. Computer 2013;46(5):58-66. [doi: 10.1109/MC.2013.38]

XSL•FO
RenderX

Abbreviations

AHLTA: Armed Forces Health Longitudinal Technology Application
CDSS: clinical decision support system
CTA: concurrent think-aloud
HCI: human-computer interaction
HIS: hospital information system
IS: information system
SeqD: sequence diagram
TA: think-aloud

Edited by N Cahill, M Balcarras; submitted 20.Jun.2025; peer-reviewed by D Ramasamy, H Anandan, L Wreyford; comments to author 14.Oct.2025; accepted 06.Jan.2026; published 02.Mar.2026.

Please cite as:
Hartmann EM, Küper A, Swoboda J, Beckmann CL, Lodde GC, Livingstone E, Schadendorf D, Sachweh S
Revealing the Physician's Task Conceptualization in Melanoma Treatment: Qualitative User Experience Study
JMIR Cancer 2026;12:e79294
URL: https://cancer.jmir.org/2026/1/e79294
doi:10.2196/79294
PMID:

©Eva Maria Hartmann, Alisa Küper, Jessica Swoboda, Catharina Lena Beckmann, Georg Christian Lodde, Elisabeth Livingstone, Dirk Schadendorf, Sabine Sachweh. Originally published in JMIR Cancer (https://cancer.jmir.org), 02.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementing UK Oncology Nursing Society–Informed Digital Symptom Triage With Episode-Based Review in Routine NHS Acute Oncology: Service Evaluation

Mohamed Sheeraz Mohamed Azhar[1], BSc, MBChB; Rishabh Ray[1], MBChB; Sagar Sharma[1], MBBS; Hiten Chauhan[1], MBChB; Margaret Ruth Robinson[1], MSc, BD, BSN; Yetunde Owoseni[1], MBChB; Anupama Vijay[1], MBBS; Sadaqat Hussain[1], MBBS, FRCP

University Hospitals of Leicester NHS Trust, Leicester, England, United Kingdom

Corresponding Author:
Mohamed Sheeraz Mohamed Azhar, BSc, MBChB
University Hospitals of Leicester NHS Trust
Infirmary Square
Leicester, England
United Kingdom
Phone: 44 0300 303 1573
Email: mohamedsheerazmohamed.azhar@nhs.net

Abstract

Background: Patients receiving systemic anticancer therapy (SACT) can deteriorate clinically between scheduled appointments; yet, acute oncology services often rely on reactive helplines with limited longitudinal symptom visibility.

Objective: The aim of this study was to evaluate the feasibility, safety, and workflow integration of OncsCare, a digital symptom triage platform mapping patient-reported symptoms to UK Oncology Nursing Society (UKONS) acuity tiers with episode-based clinician review.

Methods: This 10-week service evaluation (July to September 2025) implemented OncsCare within a UK tertiary acute oncology service. Patients completed daily symptom check-ins mapped to UKONS-informed green/amber/red tiers. Alerts were grouped into episode-level triage events using prespecified rules (48-hour window, symptom-domain continuity) to represent the operational workload. Outcomes included engagement, alert distribution, escalation pathways, timeliness of clinical response, and safety signals via structured case-finding.

Results: Thirty-two patients participated (none withdrew) in this study. Daily check-in completion rate was 91.7% (1444/1574 expected patient-days). From 362 amber/red alerts, 62 episodes were generated; 38.7% (24/62) were clinically actionable, resulting in telephone management (n=12, 50%), acute care assessment (n=9, 37.5%), emergency referral (n=2, 8.3%), or admission (n=1, 4.2%). Median time to first clinical response for in-hours red alerts was 47 minutes. Predefined safety case-finding identified no intervention-attributable safety signals. Patients reported increased home reassurance (n=17, 85%), and clinicians reported improved situational awareness without increased workload.

Conclusions: UKONS-informed digital triage with episode-based review demonstrated feasibility, with no intervention-attributable safety signals identified in this small single-center evaluation. This operational model addresses alert fragmentation and supports prospective multicenter evaluation.

(JMIR Cancer 2026;12:e92586) doi:10.2196/92586

KEYWORDS

acute oncology; digital health; patient-reported outcomes; symptom monitoring; triage; UKONS; quality improvement; implementation science

Introduction

Systemic anticancer therapy (SACT) is associated with acute toxicities that can evolve rapidly between scheduled clinical encounters. Symptoms such as diarrhea, vomiting, mucositis, dyspnea, and pain may precede time-critical complications, including infection, dehydration, and organ dysfunction. Acute oncology (AO) services therefore function as a safety net,

XSL•FO
RenderX

aiming to identify deterioration early and deliver appropriate escalation through structured triage. In the United Kingdom, the United Kingdom Oncology Nursing Society (UKONS) triage tool provides a nationally adopted framework for grading symptom severity and specifying escalation timeframes [1].

Despite widespread use of UKONS triage during telephone contacts, many AO services remain predominantly reliant on reactive, patient-initiated helplines. This model can be delayed and variably accessible, depends on patients recognizing deterioration and initiating contact, and provides limited longitudinal visibility of symptom trajectories between contacts. A recent UK-wide review led by the UK Acute Oncology Society (UKAOS) identified marked variation in the organization, governance, digital maturity, and integration of SACT helplines, alongside inconsistent data capture and limited alignment with wider urgent and emergency care pathways [2].

Electronic patient-reported outcome (ePRO) systems have demonstrated that symptom monitoring with clinician response improves symptom control and survival [3,4]. More recent trials have extended this evidence: the PRO-TECT trial demonstrated improved patient-reported outcomes across metastatic cancer types using electronic monitoring [5], and the SYMPRO-Lung study showed improved 2-year survival for patients with lung cancer monitored with web-based symptom tools [6]. Phase III evidence from the eRAPID program has further established the value of ePRO integration in routine chemotherapy care [7]. However, many platforms are designed primarily for longitudinal tracking and research workflows rather than operational integration into acute care pathways, and few map patient-reported symptoms to nationally adopted triage frameworks.

To our knowledge, few implementations have operationalized patient-reported symptom monitoring within a nationally standardized AO triage framework. To address these gaps, we developed OncsCare, a clinician-designed digital symptom monitoring and safety-netting platform that maps daily patient-reported symptoms to UKONS-informed red/amber/green acuity tiers with episode-based grouping of temporally related alerts to reduce alert fragmentation. Evaluation end points were prospectively informed by the RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance) framework [8], which underpins the 5 objectives below; implementation determinants were mapped retrospectively using the Consolidated Framework for Implementation Research (CFIR) [9], as detailed in the Methods and Discussion. We conducted a 10-week service evaluation with five objectives: (1) engagement and retention, (2) alert distribution and episode-level actionability, (3) escalation pathways and timeliness of clinical response, (4) safety signals using structured case-finding, and (5) patient and clinician experience.

Methods

Study Design

We conducted a 10-week quality improvement service evaluation (July 11, 2025 to September 19, 2025) within the AO service at University Hospitals of Leicester NHS Trust, a UK tertiary cancer center serving approximately 1.1 million people. The intervention was implemented as an adjunct to existing AO pathways and did not replace standard care. Participants retained access to usual services, including the 24-hour AO helpline, Oncology–Hematology Assessment Unit (OHAU) pathway, and emergency department (ED) assessment.

Two complementary implementation science frameworks informed the evaluation design. RE-AIM [8] was applied prospectively to define evaluation end points and structure outcome selection across engagement, adoption, and implementation domains. The CFIR [9] was applied retrospectively to map determinants of adoption across 5 domains from qualitative clinician debrief data. Both frameworks are elaborated in the Discussion.

Prior to implementation, informal focus groups with 3 patients receiving SACT and 2 caregivers informed onboarding materials and safety messaging. Key themes included the importance of clear instructions to bypass the application and use urgent/emergency pathways when needed, transparency about monitoring processes, and minimizing the burden of daily check-ins. These insights were incorporated into onboarding scripts and in-application safety-netting content.

To contextualize the evaluation, AO helpline activity was described using digitized call logs from January 2023 to June 2024. Metrics included weekly call volume and call outcomes categorized as managed without admission, escalated to OHAU, referred to ED/admitted, or administrative/other. Baseline data were descriptive and not intended for formal statistical comparison with the pilot period.

Participants

Eligible participants were adults (≥18 years) receiving active SACT with anticipated treatment duration ≥6 weeks, ECOG performance status 0-2, access to a smartphone or tablet (independently or with caregiver support), and ability to provide consent in English. Exclusion criteria included end-of-life care with estimated life expectancy <3 months, lack of capacity, enrolment in other intensive daily remote monitoring programs, or anticipated prolonged admission or travel during the pilot period. Patients were approached pragmatically during SACT visits, OHAU reviews, or outpatient appointments. Participants provided informed verbal consent documented in the clinical record using a standardized template.

Setting and Intervention

OncsCare comprises a patient-facing mobile app (iOS/Android) and a secure clinician dashboard. Participants completed a brief daily symptom check-in covering temperature, dyspnea, nausea/vomiting, diarrhea, constipation, pain, bleeding/bruising, fatigue, skin rash/reaction, peripheral neuropathy, and oral mucositis. Symptom severity was captured using a 0-3 ordinal scale with patient-facing anchors co-designed with AO nurses, mapped to UKONS-informed acuity tiers. On submission, tier-specific end-screen guidance was displayed: green alerts provided reassurance, amber alerts advised monitoring while awaiting review, and red alerts instructed patients to contact the AO helpline urgently or attend ED if unwell. Red alerts

XSL•FO
RenderX

generated real-time dashboard notifications; amber alerts were reviewed in scheduled twice-daily batches. Green submissions were available for the longitudinal context but did not require active clinician review. Dashboard monitoring occurred during weekday in-hours periods (Monday to Friday, 9 AM to 5 PM). This was distinct from the 24/7 AO helpline, which continued to operate across two 12-hour nursing shifts throughout the evaluation.

Alerts were grouped into episode-level triage events using a prespecified rule-based approach. An episode was initiated by any amber or red alert. Subsequent amber/red alerts were assigned to the same episode if they occurred within 48 hours of the most recent alert and shared symptom-domain continuity (gastrointestinal, respiratory, fever/temperature, pain, mucositis, neuropathy, fatigue/wellbeing, bleeding/bruising, skin). A new episode was created when the 48-hour window elapsed without qualifying alerts or when alerts reflected a distinct symptom pattern. Grouping consistency was reviewed periodically by the project lead.

Implementation followed Plan–Do–Study–Act (PDSA) cycles across 4 phases (Multimedia Appendix 1), iteratively refining onboarding processes, alert threshold calibration, dashboard efficiency (including episode-based alert grouping introduced in week 3), and safety-netting messaging in response to early user feedback.

Measures and Outcomes

Post-pilot qualitative feedback was obtained from 20 of 32 participants (62.5%) via a structured survey with open-ended questions. The patient survey comprised both closed and open-ended items across four domains: (1) usability and engagement (ease of use, frequency of check-ins, usefulness of reminders, clarity of symptom questions); (2) perceived reassurance and psychological impact (including reassurance, anxiety related to alerts, and perceived safety); (3) impact on care (including symptom self-management, perceived responsiveness to alerts, and overall satisfaction); and (4) acceptability and future use (including recommendation to others and interest in future participation). Free-text responses explored perceived benefits, challenges, and suggested improvements. The complete survey instrument is provided in Multimedia Appendix 2. Survey responses were collected using Google Forms and analyzed descriptively and thematically. Four structured clinician debriefs were conducted at weeks 3, 5, 8, and 10 using a semistructured format focusing on (1) impact on situational awareness and clinical prioritization, (2) workflow integration within existing AO pathways, (3) alert appropriateness and specificity, (4) usability of the dashboard interface, and (5) perceived impact on workload and service delivery. Thematic analysis was conducted using an inductive approach, with themes subsequently mapped to CFIR domains to support the structured interpretation of implementation determinants.

Primary feasibility outcomes were recruitment and retention, engagement (completed check-ins divided by expected patient-days), and alert distribution by acuity tier. Secondary outcomes included the proportion of actionable episodes, escalation pathways (telephone management, OHAU assessment, ED referral, direct admission), timeliness of clinical response for in-hours red alerts (defined as alerts submitted during dashboard-monitored hours, Monday to Friday 9 AM to 5 PM, distinct from the 24/7 AO helpline), out-of-hours escalation behavior, and safety signals (missed deterioration, delayed escalation, inappropriate reassurance). User experience outcomes were captured via patient surveys and clinician structured debriefs.

Potential safety signals were assessed using predefined case-finding and structured review. All participants were cross-referenced against OHAU attendances, ED presentations, and unplanned hospital admissions. For each acute care encounter, case notes were reviewed alongside dashboard audit logs, including symptom check-ins and alerts in the preceding 48 hours. Safety signals were predefined as (1) missed deterioration, (2) delayed escalation attributable to the workflow, (3) inappropriate reassurance leading to delayed help-seeking, or (4) technical failure contributing to harm.

Data Analysis

All analyses were descriptive (counts, proportions, medians, and ranges). Temporal trends in daily check-in volume were visualized using run charts with median lines. Qualitative feedback was reviewed thematically by 2 team members and reconciled through discussion. Given the quality improvement design and absence of a contemporaneous comparator, inferential statistical analyses were not performed.

Ethical Considerations

The work was delivered as a Trust-approved quality improvement initiative (reference: QI-AO-2025-03) and did not require research ethics committee review, in line with UK Health Research Authority guidance. A Data Protection Impact Assessment was completed and approved through University Hospitals of Leicester NHS Trust information governance processes (dated July 28, 2025). This manuscript follows SQUIRE (Standards for Quality Improvement Reporting Excellence) 2.0 reporting guidance [10].

Results

Baseline AO Service Context

Table 1 summarizes the baseline AO helpline activity. Baseline service data are reported to contextualize operational workload rather than for formal comparison with the pilot period, providing insight into underlying AO demand and the operational context within which digital triage would be deployed. In the 18 months preceding implementation (January 2023 to June 2024), the AO helpline received a median of 107 calls per week (IQR 95-118). Approximately 70% of the calls were from patients receiving active SACT. Call outcomes were as follows: 52% managed with advice without admission, 14% escalated to OHAU review, 24% referred to ED or admitted, and 10% administrative/other.

XSL•FO
RenderX

Table 1. Baseline acute oncology helpline activity (January 2023 to June 2024).

Metric	Values
Median (IQR) calls per week	107 (95-118)
Caller type (%)[a]	
Oncology patients receiving active SACT[b]	~70
Hematology patients receiving active SACT[c]	~30
Managed with telephone advice (%, no admission)[a]	52
Escalated to OHAU (%)[a,d]	14
Referred to ED[e] or admitted (%)[a]	24
Administrative/other (%)[a]	10

[a]These percentages are derived from aggregate service call logs over an 18-month period. Absolute call counts are not available from the call log system and are therefore not reported.

[b]SACT: systemic anticancer therapy.

[c]Hematology patients included in helpline activity were those receiving active systemic anticancer therapy (eg, chemotherapy for hematological malignancy) and managed within the joint Oncology–Hematology Assessment Unit pathway. This reflects the mixed oncology/hematology case-mix of the Acute Oncology helpline service and does not imply hematology patients were enrolled in the OncsCare pilot evaluation.

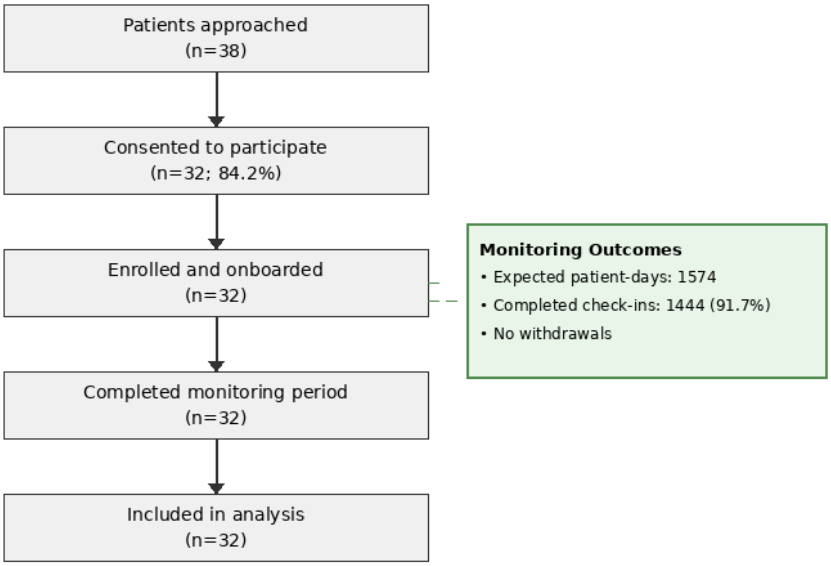
[d]OHAU: Oncology–Hematology Assessment Unit.

[e]ED: emergency department.

Recruitment and Retention

Figure 1 shows the recruitment and retention flowchart. Thirty-eight patients were approached; 32 (84.2%) consented, and none withdrew during the 10-week evaluation. Reasons for declining were concern about daily reporting burden (n=3), lack of smartphone ownership (n=2), and planned travel (n=1).

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of participant recruitment, retention, and inclusion in analysis during the 10-week service evaluation. Of the 38 patients assessed for eligibility, 6 were excluded (concern about daily reporting burden n=3, lack of smartphone ownership n=2, anticipated prolonged travel n=1). Thirty-two patients were enrolled; zero withdrew during the 10-week pilot. Overall check-in completion: 91.7% (1444/1574 expected patient-days).



The full participant demographics and clinical characteristics are presented in Table 2. The median age was 55 (range 33-77) years; 21 (65.6%) participants were females. Tumor groups were predominantly breast (n=18, 56.3%), with representation from colorectal (n=4), lungs (n=3), and other sites. Most participants managed the application independently (n=29, 90.6%); 3 used caregiver support. Treatment intent was curative in 17 (53.1%) participants and palliative in 15 (46.9%) participants. By treatment setting, 9 (28.1%) were receiving neoadjuvant therapy, 6 (18.8%) adjuvant, 2 (6.3%) other curative, and 15 (46.9%) metastatic. Treatment modality comprised chemotherapy alone in 16 (50%), chemo-immunotherapy in 10 (31.3%), chemotherapy with targeted therapy in 1 (3.1%), and targeted or biologic therapy alone in 5 (15.6%). ECOG performance status was 0 in 15 (46.9%), 1 in 16 (50%), and 2 in 1 (3.1%). Participants represented a heterogeneous AO population: 53.1% (n=17) were receiving treatment with curative intent (including neoadjuvant

XSL•FO
RenderX

and adjuvant settings) and 46.9% (n=15) palliative, with substantial representation of chemoimmunotherapy regimens (n=10, 31.3%).

Table 2. Participant demographics and clinical characteristics (N=32). Treatment modality categories are mutually exclusive as recorded at enrollment.

Characteristic	Values
Age (y), median (range)	55 (33-77)
Female sex, n (%)	21 (65.6)
Tumor site, n (%)	
Breast	18 (56.3)
Colorectal	4 (12.5)
Lung	3 (9.4)
Testicular	2 (6.3)
Other (gastric, ovarian, esophageal, pancreatic, sarcoma)	5 (15.6)
Treatment intent, n (%)	
Curative	17 (53.1)
Palliative	15 (46.9)
Treatment setting, n (%)	
Neoadjuvant	9 (28.1)
Adjuvant	6 (18.8)
Other curative	2 (6.3)
Metastatic	15 (46.9)
Treatment modality, n (%)	
Chemotherapy alone	16 (50)
Chemo-immunotherapy	10 (31.3)
Chemotherapy + targeted therapy	1 (3.1)
Targeted/biologic therapy alone	5 (15.6)
Route of administration, n (%)	
Oral SACT[a]	14 (43.8)
Intravenous SACT	16 (50)
Subcutaneous immunotherapy	2 (6.3)
ECOG[b] performance status, n (%)	
ECOG 0	15 (46.9)
ECOG 1	16 (50)
ECOG 2	1 (3.1)
Digital self-sufficiency, n (%)	29 (90.6)
Caregiver-assisted, n (%)	3 (9.4)

[a]SACT: systemic anticancer therapy.

[b]ECOG: Eastern Cooperative Oncology Group.

Objective 1: Engagement and Retention

Across the evaluation, 1444 symptom check-ins were completed (median 44 per participant; range 12-70). Of 1574 expected patient-days, the overall completion rate was 91.7% (1444/1574). Engagement was sustained without marked attrition (Figure S1C in Multimedia Appendix 3). High recruitment rate (32/38, 84.2%) and zero attrition were consistent with sustained engagement throughout the pilot period, supporting the feasibility of this monitoring model in routine outpatient oncology.

Objective 2: Alert Distribution and Episode-Level Actionability

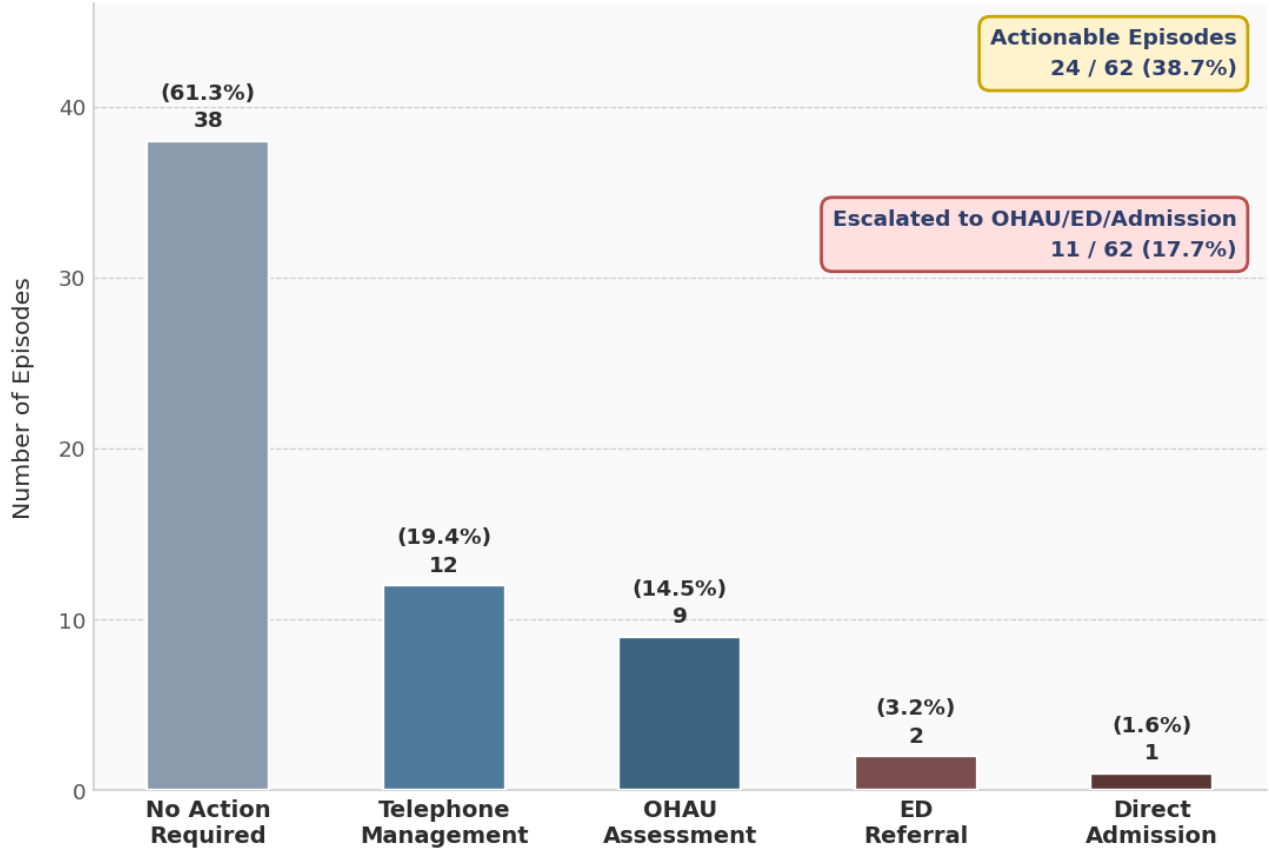
Across the 10-week evaluation, 1444 check-ins generated 1082 green (74.9%), 228 amber (15.8%), and 134 red (9.3%) alerts. The median number of amber/red alerts per participant was 11 (range 2-31). Red alert distribution was highly skewed: 11 of

XSL•FO
RenderX

the participants (34.4%) experienced no red alerts, while the majority of the remainder (14/32, 43.8%) had only 1-3 alerts. A small subset (3/32, 9.4%) accounted for a disproportionate number of alerts, each experiencing ≥10 high-acuity events, suggesting that alert burden is concentrated in a minority of patients. Of 134 red alerts, 69 (51.5%) occurred during staffed monitoring hours and were managed via the dashboard workflow; the remaining 65 (48.5%) occurred out-of-hours and were managed through standard AO helpline and emergency pathways consistent with onboarding safety-netting. Alert activity peaked in mid-August (weeks 5-7), consistent with mid-cycle toxicity periods (Figures S1A-S1C in Multimedia Appendix 3).

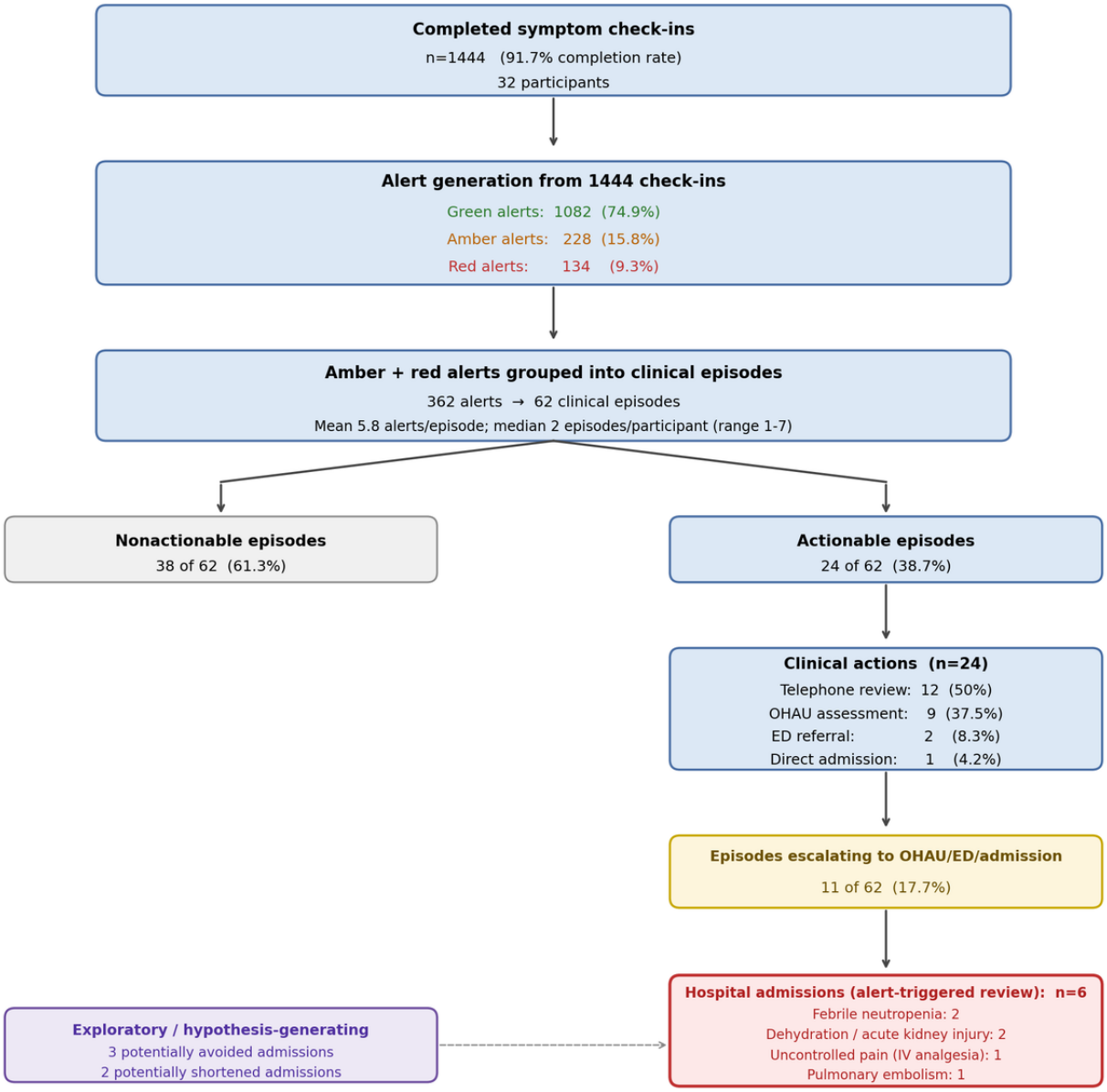
Figure 2 shows the distribution of clinical actions, while Figure 3 illustrates the operational flow from symptom check-ins through episode grouping to clinical outcomes. Table 3 summarizes the episode classifications. Using the prespecified episode grouping approach, 362 amber/red alerts were consolidated into 62 clinical episodes (mean 5.8 alerts per episode; range 1-18)—a 5.8-fold reduction in discrete review events. The median number of episodes per participant was 2 (range 1-7). Twenty-four episodes (38.7%) were clinically actionable; 38 (61.3%) were reviewed without intervention as symptoms were mild, self-limiting, or already under management.

Figure 2. Bar chart showing the distribution of clinical actions across all 62 episodes. Of 62 episodes, 38 (61.3%) required no intervention; among 24 (38.7%) actionable episodes, telephone management (n=12, 50%), OHAU assessment (n=9, 37.5%), ED referral (n=2, 8.3%), and direct admission (n=1, 4.2%) are shown. ED: emergency department; OHAU: Oncology–Hematology Assessment Unit.



XSL•FO
RenderX

Figure 3. Operational flow from symptom check-ins through episode grouping to clinical outcomes and downstream care pathways. AKI: acute kidney injury; ED: emergency department; FN: febrile neutropenia; IV: intravenous; OHAU: Oncology–Hematology Assessment Unit; PE: pulmonary embolism.



XSL•FO
RenderX

Table 3. Clinical episode classification and outcomes (n=62 episodes).

Episode category	Values, n (%)
Total clinical episodes	62 (100)
Episode actionability	
Clinically actionable	24 (38.7)
Non-actionable (reviewed, no intervention required)	38 (61.3)
Clinical actions (actionable episodes, n=24)[a]	
Telephone clinical review with safety-netting advice	12 (50)
OHAU[b] assessment	9 (37.5)
ED[c] referral	2 (8.3)
Direct hospital admission	1 (4.2)
Escalation outcomes (all episodes, n=62)[d]	
Episodes escalating to OHAU/ED/admission	11 (17.7)
Episodes managed with telephone advice only	12 (19.4)
Episodes reviewed, no intervention required	38 (61.3)
Hospital admissions following alert-triggered review	6 (9.7)

[a]Percentages for clinical actions are of actionable episodes (n=24).

[b]OHAU: Oncology–Hematology Assessment Unit.

[c]ED: emergency department.

[d]Percentages for escalation outcomes are of all episodes (n=62).

Among the 24 actionable episodes, telephone clinical review with safety-netting advice was the most frequent (n=12, 50%), followed by OHAU assessment (n=9, 37.5%), ED referral (n=2, 8.3%), and direct admission (n=1, 4.2%). Overall, 11 of 62 episodes (17.7%) resulted in escalation to OHAU, ED, or admission. Six hospital admissions occurred following alert-triggered review: febrile neutropenia (n=2), dehydration with acute kidney injury (n=2), uncontrolled pain requiring intravenous analgesia (n=1), and pulmonary embolism (n=1).

As a hypothesis-generating, exploratory analysis only, 5 cases were identified prospectively by the clinical team as potential examples where earlier digital detection may have supported admission avoidance or shortened length of stay. These were subsequently independently adjudicated via structured case-note review: 2 potentially shortened admissions (febrile neutropenia managed with timely antibiotics); 2 potentially avoided admissions (severe diarrhea managed in OHAU with IV fluids); and 1 complex episode managed safely without admission.

Objectives 3 and 4: Escalation Pathways, Timeliness, and Safety Signals

Using dashboard audit logs and retrospective review of acute care encounters, 2 instances were identified in which participants submitted a green check-in earlier in the day but subsequently presented to hospital later the same day. In both cases, presenting complaints (acute abdominal pain and febrile neutropenia, respectively) developed after the morning check-in; both patients appropriately activated standard urgent care pathways. Neither case represented inappropriate reassurance attributable to the intervention. These cases illustrate a fundamental limitation of point-in-time symptom assessment: a single daily check-in cannot detect acute intraday deterioration, and the platform does not provide continuous monitoring.

No other intervention-related safety signals were identified. All 69 red alerts submitted during in-hours monitoring periods (Monday to Friday 9 AM to 5 PM) were reviewed: 66 (95.7%) of the 69 alerts on the same day. Median time to first clinical response was 47 minutes (range 12 minutes to 6 hours). Three reviews exceeded 2 hours due to concurrent OHAU activity; in each instance, patients were contacted and managed via an established helpline pathway with no patient harm resulting.

Objective 5: Patient and Clinician Experience

Post-pilot qualitative feedback was obtained from 20 of the 32 participants (62.5%). Four primary themes were identified: reassurance through active monitoring (17/20, 85%), ease of use and low burden (20/20, 100%), clarity of symptom reporting and guidance (15/20, 75%), and confidence in timely clinical response (14/20, 70%). Illustrative quotes are provided below, with additional quotations in Multimedia Appendix 4. Regarding reassurance, one participant described, “Having a daily reminder to check in and feeling I was being supported.” Another noted, “The thought of someone monitoring your symptoms.” Regarding ease of use and low burden, “Easy to use. Clear instructions… and reassuring knowing there was someone checking on any symptoms that were worrying.” Some participants highlighted areas for refinement, including the need for clearer symptom grading and additional contextual input options.

These qualitative findings align with the high engagement and completion rates observed, suggesting that perceived clinical oversight and low interaction burden were key drivers of

XSL•FO
RenderX

sustained use. Within the CFIR framework, these themes map primarily to inner setting (workflow compatibility, perceived value) and intervention characteristics (usability, clarity of outputs), supporting the feasibility of the integration into routine AO practice. One clinician highlighted, "Ability to see at a glance the symptoms and concerns of patients and easily identify those needing help."

Clinician debriefs (n=3 clinicians; 4 sessions at weeks 3, 5, 8, and 10) identified five themes: (1) improved situational awareness across the outpatient cohort (3/3, 100%), (2) real-time alerts with clear audit trail (3/3, 100%), (3) ease of use with minimal onboarding (2/3, 67%), (4) early escalation potentially preventing or shortening admissions (3/3, 100%), and (5) need for alert specificity refinement and contextual data integration (3/3, 100%). All clinicians reported no increase in the perceived workload.

Discussion

Principal Findings

In this 10-week service evaluation, implementation of a UKONS-informed digital symptom triage platform within routine NHS AO pathways was feasible and acceptable, with sustained engagement (1444/1574, 91.7% completion), timely clinician response to in-hours high-acuity alerts (median 47 minutes), and no intervention-attributable safety signals identified via structured case-finding. The limited sample size and single-center design constrain the conclusions that can be drawn, and larger multicenter evaluation is required. This evaluation provides an operational signal that structured UKONS-aligned digital triage can be integrated into routine AO workflows without increasing clinician workload.

The findings of this evaluation are consistent with prior trials of ePRO monitoring, which have demonstrated improved symptom control, quality of life, and survival when structured symptom reporting is coupled with clinician response [3,4,7]. However, these systems were primarily designed for longitudinal monitoring and research workflows. In contrast, OncsCare was designed for direct integration into AO triage pathways, mapping symptoms to nationally adopted UKONS escalation thresholds and supporting real-time clinical decision-making. Unlike prior ePRO systems, which rely on the clinician interpretation of longitudinal symptom trends, OncsCare provides structured triage outputs aligned with existing clinical frameworks. This distinction highlights a potential evolution from monitoring-focused platforms to operational, workflow-integrated digital safety-netting interventions.

Alert Specificity and Algorithm Refinement

Only 38.7% (24/62) of the clinical episodes were actionable, meaning the majority of the episodes reviewed did not require clinical intervention. While this did not produce evidence of alert fatigue in this small cohort, a high nonactionable rate has implications for adoption readiness. Immediate next steps should prioritize practical implementation work: refinement of alert thresholds and composite alert rules informed by adjudicated episode outcomes, integration of treatment schedule data to contextualize expected versus unexpected toxicity, and incorporation of laboratory results (particularly neutrophil counts for temperature alerts). Adoption-readiness testing across two NHS sites, design and specification refinement, and early health economic modelling are the near-term priorities before broader scale-up is considered. The skewed distribution of the red alerts—with a small minority of patients accounting for a disproportionate alert burden—suggests that threshold calibration and targeted monitoring strategies should be priorities for future implementation, potentially enabling more efficient resource allocation.

This proportion of actionable episodes suggests that triage outputs retained clinical specificity while avoiding indiscriminate escalation—a balance that will require ongoing calibration as the platform scales. The patient perspective on nonactionable alerts also warrants attention. This evaluation did not systematically capture how participants perceived episodes in which a red-tier alert was raised but no clinical action was taken. Patients who receive urgent red-tier guidance but are subsequently told no action is needed may experience confusion or eroded trust in the alert system. Future studies should include patient-reported experience of nonactionable alerts as a specific outcome measure. Patient feedback also identified specific usability refinements for future iteration: clearer differentiation of repeat notifications, the addition of free-text fields for contextual symptom detail, and improved anchoring of symptom severity scales. These insights directly informed planned algorithm and interface refinement and are consistent with the iterative PDSA approach adopted throughout the evaluation.

Safety in Digital Urgent-Care Workflows

Safety in digital urgent-care workflows requires layered safeguards. OncsCare was implemented as an adjunct to existing AO pathways with bounded monitoring hours, explicit in-app and onboarding safety-netting, UKONS-informed escalation thresholds, audit logging, and retrospective safety case-finding linked to objective acute care encounters. The 2 same-day presentations reinforce that the platform complements rather than replaces clinical vigilance and patient self-escalation. Importantly, no cases of harm attributable to delayed or inappropriate triage were identified. Larger multicenter evaluation with prespecified safety end points and independent adjudication is required before safety conclusions can be drawn. The 5 potentially avoided or shortened admissions identified in exploratory analysis should be interpreted cautiously: these classifications are counterfactual and subject to ascertainment bias and are presented solely to inform hypothesis generation for future controlled evaluation.

Operational Implementation and Practical Implications

This evaluation highlights a transferable design pattern: mapping patient-reported symptoms to established clinical triage frameworks, coupled with episode-level grouping to reflect operational workload rather than raw alert volume. This reflects a shift from reactive, contact-driven triage to proactive, structured situational awareness at the service level. Episode-based grouping represents a further shift from alert-level monitoring to clinically meaningful workload representation, reducing discrete review events by 5.8-fold in

XSL•FO
RenderX

this evaluation. Near-term practical implications include augmenting SACT helplines with structured longitudinal symptom visibility, embedding digital monitoring within routine clinical governance using PDSA cycles, and developing audit infrastructure aligned with UKAOS recommendations for standardized data capture. Health economic evaluation is a priority for any future multicenter evaluation. A key perceived benefit, reported by all 3 clinicians in debrief, was enhanced cohort-level oversight, enabling risk-stratified prioritization rather than reliance on episodic, patient-initiated contact. This positions OncsCare within a broader landscape of digital triage approaches. In contrast, reactive models, including patient-initiated helplines and e-consult pathways, do not provide longitudinal visibility between contacts. OncsCare offers a clinically structured intermediate: structured daily reporting mapped to nationally adopted escalation thresholds, with episode-level workload consolidation designed to support sustainable clinical review. Future comparative evaluations should examine whether proactive symptom monitoring confers measurable benefits in safe escalation versus reactive-only alternatives.

Digital Equity and Accessibility

This evaluation demonstrates high engagement among participants with smartphone access and digital confidence. However, eligibility criteria requiring smartphone ownership and English-language consent likely excluded digitally marginalized populations. Digital health interventions risk widening health inequities if deployment favors populations with existing digital access and literacy [11,12]. Future implementations should adopt deliberate equity-focused strategies, including formal digital literacy assessment [13], alternative access modalities (SMS text message–based reporting, telephone check-ins), multilingual interfaces, and device loan programs. These considerations are addressed further in the Limitations section.

Implementation Science Framework Mapping

Retrospective CFIR mapping identified key adoption facilitators in the inner setting domain (leadership engagement, workflow compatibility) and outer setting domain (alignment with UKAOS policy framework) [9]. Barriers to wider adoption include intervention characteristics (alert specificity), inner setting (limited electronic health record integration), and equity of access (individual characteristics and outer setting). Within the RE-AIM framework [8], this evaluation addressed Reach (32 patients enrolled; 32/38, 84.2% consent rate), Adoption (clinician engagement with dashboard workflow and participation in structured debriefs), and Implementation fidelity (1444/1574, 91.7% check-in completion; 66/69, 95.7% same-day red alert review). Effectiveness (clinical outcomes) and Maintenance (sustained adoption beyond the pilot) will be formally evaluated in the prospective multicenter study.

Comparison With Other ePRO Platforms

While eRAPID [7] and PRO-CTCAE [14] have demonstrated efficacy in longitudinal symptom tracking and research contexts, OncsCare was designed specifically for operational integration into acute care triage workflows, with UKONS framework mapping, episode-based alert consolidation, and PDSA-driven implementation via routine clinical governance.

Limitations

This evaluation is limited by its 10-week duration, single-center design, modest sample size (n=32), and absence of a contemporaneous comparator.

Selection Bias and Generalizability

Pragmatic recruitment, smartphone ownership requirements, and English-language consent criteria introduce systematic selection bias. Findings should not be generalized to digitally excluded populations, older adults with lower digital literacy, those with limited English, or patients managed at nontertiary centers.

Clinical Lead Involvement

The lead author is the developer of the OncsCare platform, creating potential for confirmation and ascertainment bias in episode classification, actionability judgements, and thematic analysis. Independent adjudication of a random 21% sample (92.3% concordance for episode assignment, 84.6% for actionability) provides partial mitigation. Future evaluations should involve fully independent adjudication and should separate the platform developer from primary data collection roles.

Point-In-Time Assessment and Intraday Deterioration

A single daily check-in cannot detect acute clinical deterioration between submissions. The 2 same-day presentations illustrate this directly. The platform does not provide continuous monitoring, and safety-netting messaging must consistently emphasize this constraint.

Out-Of-Hours Coverage

Dashboard monitoring was restricted to weekday in-hours periods (Monday to Friday, 9 AM to 5 PM). This is distinct from the AO helpline, which operates 24/7 across two 12-hour nursing shifts and remained fully available throughout the evaluation. Nearly half of the red alerts (65/134, 48.5%) were submitted out-of-hours and were not captured by the dashboard workflow. While safety-netting directed patients to the 24/7 helpline and emergency services, we cannot assess whether digital monitoring influenced out-of-hours escalation behavior.

Alert Specificity

Only 38.7% (24/62) of the episodes were actionable. Episode grouping relied on rule-based logic rather than validated clinical algorithms, and threshold calibrations were not validated against an independent dataset.

Subtle Harms

Safety case-finding was linked to objective acute care encounters. Subtle harms not resulting in OHAU/ED/admission—such as patient anxiety from nonactionable red alerts or delayed self-escalation—may not have been detected.

XSL•FO
RenderX

Digital Equity in Future Implementations

Deliberate equity-focused strategies will be required for wider deployment, including formal digital literacy assessment using validated tools such as eHEALS [13]; alternative access modalities for non–smartphone users; multilingual interfaces co-designed with ethnically diverse patient groups; device loan programs; and explicit subgroup analysis by age, ethnicity, socioeconomic status, and digital literacy in the planned multicenter study (2027-2028).

Health Economics

Cost per avoided admission, clinician dashboard review time, and wider health system impact were not formally evaluated and are priorities for the multicenter evaluation.

Conclusions

This single-center service evaluation demonstrates the feasibility of implementing UKONS-informed digital symptom triage with episode-based review within routine NHS AO workflows. High patient engagement, episode-level actionability, timely clinical response, and absence of intervention-attributable safety signals in this small single-center evaluation support further multicenter evaluation. Planned next steps include adoption-readiness testing across 2 NHS sites, alert algorithm refinement, and early health economic modelling, with a view to informing the design of a comparative multicenter study.

Acknowledgments

We thank Jaya Jaitly for her contribution to patient involvement activities during the evaluation. We also thank the Acute Oncology nursing team and oncology registrars at University Hospitals of Leicester NHS Trust for operational support and the patients and caregivers who provided feedback to improve onboarding and safety-netting materials. The authors used Claude (Anthropic, claude.ai) to assist with manuscript drafting, editing, and language refinement during the preparation of this work. All content, data, analysis, and conclusions were produced, verified, and are the full responsibility of the authors. The use of artificial intelligence assistance was limited to writing support and did not involve data analysis, interpretation of results, or generation of clinical conclusions.

Funding

This work was delivered as an internally supported quality improvement initiative within the Department of Oncology, University Hospitals of Leicester NHS Trust. No external funding was received.

Data Availability

Deidentified aggregate data supporting the findings are available from the corresponding author on reasonable request, subject to University Hospitals of Leicester NHS Trust information governance approvals. Individual-level datasets are not publicly available due to privacy and consent constraints.

Authors' Contributions

Conceptualization: MSMA (lead)
Methodology: MSMA (lead), MR
Investigation: MSMA (lead), MR, SS, HC, YO, AV
Formal Analysis: MSMA (lead), RR
Data Curation: MSMA (lead), RR
Project Administration: MSMA (lead), RR, SS
Software: MSMA (lead)
Supervision: SH
Validation: MSMA (lead), MR, SH
Writing – original draft: MSMA (lead)
Writing – review & editing: All authors

Conflicts of Interest

MSMA is the clinical lead for the OncsCare platform evaluated in this service evaluation. At the time of this evaluation, there was no commercial deployment and no external commercial funding. SH has received honoraria from Eisai, Pfizer, Merck Serono, and Novartis; travel grants from Roche, Novartis, Merck, and Ipsen; and has been involved in clinical trials with Apopharma and Lilly. All other authors declare no competing interests.

Multimedia Appendix 1
Plan–Do–Study–Act (PDSA) cycles undertaken during the 10-week service evaluation, detailing iterative refinements to the OncsCare platform, including alert logic, workflow integration, and clinician review processes within acute oncology pathways.
[DOCX File , 16 KB - cancer_v12i1e92586_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Patient feedback survey instrument administered post-pilot to OncsCare participants (20/32, 62.5%). Survey sections cover demographics, app usability, symptom reporting experience, impact on care, and recommendations for future development.
[DOCX File , 11 KB - cancer_v12i1e92586_app2.docx]

Multimedia Appendix 3
Additional visualizations of alert patterns during the 10-week evaluation period. (A) Distribution of alert acuity over time showing green, amber, and red alerts. (B) Heat map of symptom acuity illustrating temporal patterns across participants. (C) Run chart of red alerts over time showing variation in high-acuity symptom reporting.
[PNG File , 281 KB - cancer_v12i1e92586_app3.png]

Multimedia Appendix 4
Full thematic analysis of patient and clinician feedback from the OncsCare service evaluation. Part A presents four patient themes derived from post-pilot survey free-text responses (n=20). Part B presents five clinician themes derived from structured debrief sessions and post-pilot survey (n=3). Representative quotes and supporting quantitative data are provided for each theme.
[DOCX File , 15 KB - cancer_v12i1e92586_app4.docx]

Multimedia Appendix 5
SQUIRE 2.0 checklist.
[PDF File (Adobe PDF File), 165 KB - cancer_v12i1e92586_app5.pdf]

References

1. UKONS 24-hour triage tool (version 3). UK Oncology Nursing Society (UKONS). URL: https://www.ukacuteoncology.co.uk/application/files/4217/3799/3595/UKONS_Triage_tool_poster_A3_V10.pdf [accessed 2026-05-10]
2. A UK review of systemic anti-cancer therapy helplines: getting it right first time for cancer patients. UK Acute Oncology Society. URL: https://www.ukacuteoncology.co.uk/application/files/7117/5260/7575/UKAOS_4_Nations_Report_v7_003_PDF_2.pdf [accessed 2026-05-10]
3. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
4. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
5. Basch E, Schrag D, Henson S, Jansen J, Ginos B, Stover AM, et al. Effect of electronic symptom monitoring on patient-reported outcomes among patients with metastatic cancer: a randomized clinical trial. JAMA 2022 Jun 28;327(24):2413-2422 [FREE Full text] [doi: 10.1001/jama.2022.9265] [Medline: 35661856]
6. Denis F, Basch E, Septans A, Bennouna J, Urban T, Dueck AC, et al. Two-year survival comparing web-based symptom monitoring vs routine surveillance following treatment for lung cancer. JAMA 2019 Jan 22;321(3):306-307 [FREE Full text] [doi: 10.1001/jama.2018.18085] [Medline: 30667494]
7. Absolom K, Warrington L, Hudson E, Hewison J, Morris C, Holch P, et al. Phase III randomized controlled trial of eRAPID: eHealth intervention during chemotherapy. J Clin Oncol 2021 Mar 01;39(7):734-747 [FREE Full text] [doi: 10.1200/JCO.20.02015] [Medline: 33417506]
8. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
9. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
10. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. BMJ Qual Saf 2016 Dec;25(12):986-992 [FREE Full text] [doi: 10.1136/bmjqs-2015-004411] [Medline: 26369893]
11. Palacholla RS, Fischer N, Coleman A, Agboola S, Kirley K, Felsted J, et al. Provider- and patient-related barriers to and facilitators of digital health technology adoption for hypertension management: scoping review. JMIR Cardio 2019 Mar 26;3(1):e11951 [FREE Full text] [doi: 10.2196/11951] [Medline: 31758771]
12. Evidence standards framework for digital health technologies. National Institute for Health and Care Excellence. 2019. URL: https://www.nice.org.uk/what-nice-does/digital-health/evidence-standards-framework-esf-for-digital-health-technologies [accessed 2026-05-10]
13. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]

XSL•FO
RenderX

14. Dueck AC, Mendoza TR, Mitchell SA, Reeve BB, Castro KM, Rogak LJ, National Cancer Institute PRO-CTCAE Study Group. Validity and reliability of the US National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). JAMA Oncol 2015 Nov;1(8):1051-1059 [FREE Full text] [doi: 10.1001/jamaoncol.2015.2639] [Medline: 26270597]

Abbreviations

AO: Acute Oncology
CFIR: Consolidated Framework for Implementation Research
ED: emergency department
ePRO: electronic patient-reported outcome
OHAU: Oncology–Hematology Assessment Unit
PDSA: Plan–Do–Study–Act
RE-AIM: Reach, Effectiveness, Adoption, Implementation, Maintenance
SACT: systemic anticancer therapy
SQUIRE: Standards for Quality Improvement Reporting Excellence
UKAOS: UK Acute Oncology Society
UKONS: UK Oncology Nursing Society

Edited by M Balcarras; submitted 31.Jan.2026; peer-reviewed by E Marshall, C Drossaert; comments to author 02.Mar.2026; revised version received 04.May.2026; accepted 04.May.2026; published 21.May.2026.

Please cite as:
Mohamed Azhar MS, Ray R, Sharma S, Chauhan H, Robinson MR, Owoseni Y, Vijay A, Hussain S
Implementing UK Oncology Nursing Society–Informed Digital Symptom Triage With Episode-Based Review in Routine NHS Acute Oncology: Service Evaluation
JMIR Cancer 2026;12:e92586
URL: https://cancer.jmir.org/2026/1/e92586
doi:10.2196/92586
PMID:

©Mohamed Sheeraz Mohamed Azhar, Rishabh Ray, Sagar Sharma, Hiten Chauhan, Margaret Ruth Robinson, Yetunde Owoseni, Anupama Vijay, Sadaqat Hussain. Originally published in JMIR Cancer (https://cancer.jmir.org), 21.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Cost-Effectiveness of AI-Powered Decentralized Production for Advanced Therapies in the Hospital–Generated Chimeric Antigen Receptor T-Cell Therapy for Multiple Myeloma From a German Hospital Perspective: Early Health Technology Assessment Study

Iris W A Boot[1], MSc; Anna Louisa Weltin[2], MSc; Ysanne de Graaf[1,3], MSc; Hubertus J M Vrijhoef[1], PhD

[1]Panaxea, 's-Hertogenbosch, The Netherlands

[2]Department Bio-adaptive Production, Fraunhofer Institute for Production Technology IPT, Aachen, North Rhine-Westphalia, Germany

[3]Faculty of Health, Medicine and Life Sciences, Maastricht University, Maatricht, Limburg, The Netherlands

Corresponding Author:
Anna Louisa Weltin, MSc
Department Bio-adaptive Production
Fraunhofer Institute for Production Technology IPT
Steinbachstraße 17
Aachen, North Rhine-Westphalia, 52074
Germany
Phone: 49 241 8904 236
Email: louisa.weltin@ipt.fraunhofer.de

Abstract

Background: Although effective, current chimeric antigen receptor T-cell (CAR-T) production methods—centralized, manual, and complex—are cost-intensive, time-consuming, and prone to variability. AIDPATH (AI-Powered Decentralized Production for Advanced Therapies in the Hospital) proposes a decentralized, automated alternative that integrates patient-specific data; optimizes resource use; and potentially improves cell viability, manufacturing efficiency, and patient outcomes.

Objective: The aim of this study was to compare CAR-T therapy produced by the AIDPATH platform to both ciltacabtagene autoleucel (cilta-cel) and standard of care (SoC) for patients with triple-class refractory multiple myeloma (MM) over a 40-year time horizon from the hospital perspective in Germany.

Methods: A partitioned survival model reflecting 3 health states (progression-free disease, progressed disease, and death) was used. The analysis used clinical trial data for cilta-cel, real-world data for SoC, and estimated parameters for AIDPATH due to the developmental status of the platform. The primary outcome was the incremental cost-effectiveness ratio; secondary outcomes included sensitivity and scenario analyses.

Results: On the basis of the model used in this analysis, AIDPATH was dominant compared to both cilta-cel and SoC. Most costs for CAR-T therapies were driven by acquisition and adverse events. Sensitivity analyses showed that the results were most influenced by discount rates and assumptions about progression-free survival. Scenario analyses, including reduced adverse events and shorter vein-to-vein times for AIDPATH, further supported its cost-effectiveness.

Conclusions: This is the first study to assess the cost-effectiveness of a CAR-T product generated with AI support in Germany from the hospital perspective. AIDPATH was found to be a cost-effective alternative to both cilta-cel and SoC, making it a promising option for future implementation. Although further data are needed, this study provides valuable guidance for health care stakeholders, reimbursement discussions, and future research.

(JMIR Cancer 2026;12:e92253) doi:10.2196/92253

KEYWORDS

cost-effectiveness analysis; chimeric antigen receptor T-cell therapy; CAR-T therapy; early health technology assessment; early HTA; artificial intelligence manufacturing; AI manufacturing; health economic modeling; automated decentralized manufacturing

XSL•FO
RenderX

Introduction

Chimeric antigen receptor (CAR) T-cells are a new potent treatment option for patients with cancer, with long-term data demonstrating robust efficacy and overall low levels of toxicity. CAR T-cell (CAR-T) therapy works by collecting a patient's T-cells, genetically modifying them to express tumor-targeting CARs, and reinfusing them to produce a targeted immune response [1]. However, as CAR-T therapy advances, a critical bottleneck is emerging in the manufacturing and distribution of these specialized cells [2].

Currently, CAR-T cells are predominantly manufactured in centralized facilities. This process introduces variability in the quality of the produced cells and limits scalability. Moreover, there is often a significant time gap between T-cell collection and CAR-T product infusion, referred to as the vein-to-vein time, which can lead to patient deterioration [2]. This can compromise product effectiveness or even render patients unsuitable for infusion. Although the standard production methods are groundbreaking, they are operationally complex and costly.

AIDPATH (AI-Powered Decentralized Production for Advanced Therapies in the Hospital) is a European Union–funded project (grant 101016909) addressing these challenges by automating and decentralizing CAR-T cell manufacturing. AIDPATH's AI technology enables a deeper understanding of the production process and facilitates the identification of correlations between production and patient data, thereby supporting its patient-centric approach. The platform for producing this therapy is being built and is expected to optimize scheduling and resource planning, thereby reducing costs and hospital resource use. By overcoming the limitations of centralized manufacturing, AIDPATH has the potential to enhance efficiency, cost-effectiveness, and patient access to and outcomes of cancer treatment [3].

The objective of this study was to assess the cost-effectiveness of CAR-T cell therapy produced by the AIDPATH platform for patients with multiple myeloma (MM) in Germany, from the hospital perspective, over a 40-year time horizon, compared to manually produced and centrally distributed CAR-T cell therapy (ciltacabtagene autoleucel [cilta-cel]) and standard of care (SoC).

The results of this analysis will provide insights into whether transitioning to automated AI-supported CAR-T cell production is expected to be a cost-effective alternative. These results may support evidence-based decision-making for policymakers, health care administrators, and clinicians considering adoption of the AIDPATH platform.

Methods

Study Design and Population

To assess the cost-effectiveness of CAR-T cell therapy generated by the AIDPATH platform compared to SoC and manually produced, centrally distributed CAR-T cell therapy (cilta-cel; Johnson & Johnson), a partitioned survival model was developed as part of this early health technology assessment (HTA). The partitioned survival model consists of health states that a patient can occupy at a given time point, which reflect the clinical trajectory of patients [4].

The study population comprised adult patients with MM in Germany who had not responded to or had experienced a relapse/refractory (R/R) after at least 3 prior lines of cancer treatment. These treatments include an immunomodulator, a proteasome inhibitor, and a cluster of differentiation 38 antibody. In cases where the disease no longer responds to agents in each of the 3 classes, it is referred to as "triple-class refractory" MM [5]. This patient group is currently eligible for CAR-T cell therapy within the German health care system [6].

Given that social, environmental, and other contextual factors influence disease progression [7], this study focuses on a single country. As the first AIDPATH platform is being implemented at the University Hospital Würzburg, a leading center for CAR-T cell research and treatment, Germany was selected as the setting for this analysis.

Model Overview

The partitioned survival model used in this study consists of 3 health states: progression-free disease (PFD), progressive disease (PD), and death (Figure 1). For the cohort receiving cilta-cel, the starting point was "apheresis" on the decision tree. Hereafter, patients were either not infused and streamed into the partitioned survival model or started a bridging period while they waited for their CAR-T therapy to be produced. After the bridging period, the infused cohort streamed into the partitioned survival model. All cohorts start in the PFD health state. Patients enter the model in the PFD health state and receive either CAR-T cell therapy (derived from AIDPATH or cilta-cel) or SoC. Patients enter the model at 61 years old, based on the selected patient group's average age of 61.7 years in the cilta-cel clinical trial [8]. A cycle length of 1 week was used, and the time horizon for the analysis was 40 years (ie, 2080 cycles). This is in line with the Dutch Health Care Institute (*Zorginstituut Nederland* [*ZIN*]) cilta-cel study and represents a lifelong time horizon [5].

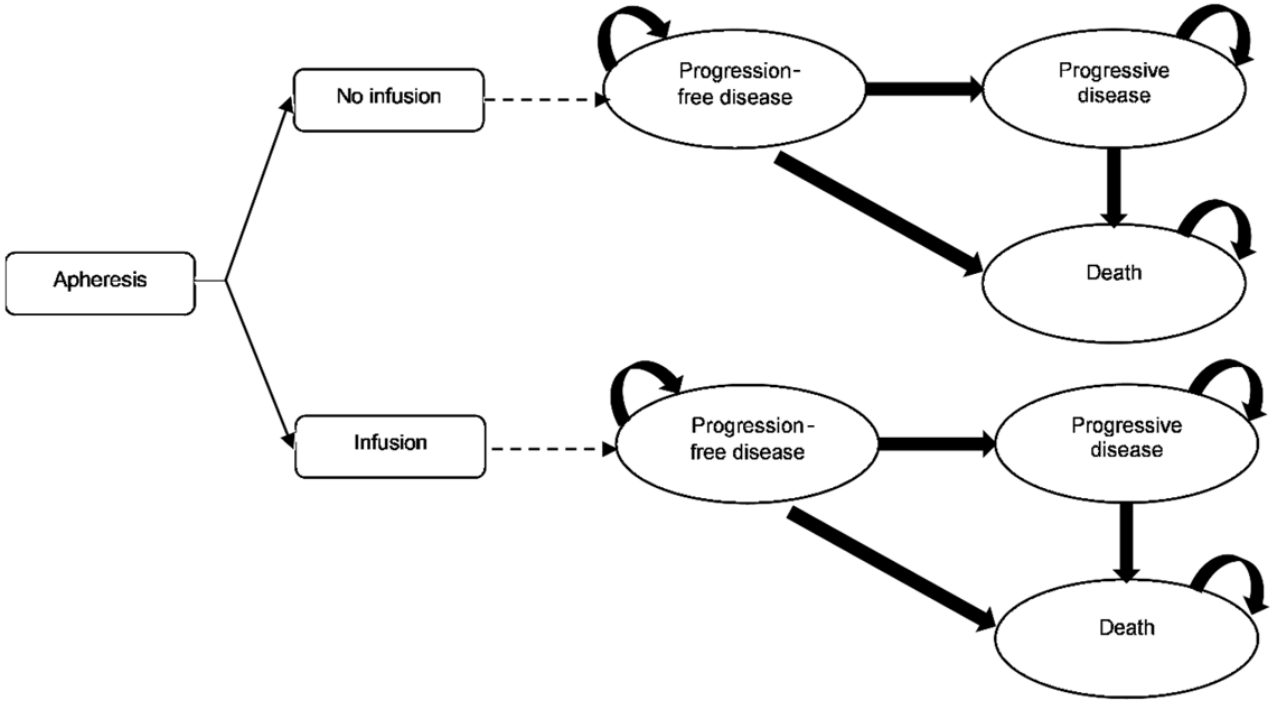
The number of patients in the PD health state at a certain time equals the overall survival (OS) proportion minus the progression-free survival (PFS) proportion. The number of patients in the death state equals 1–OS. OS and PFS rates for SoC were based on German real-world clinical population data. Here, time to next treatment was used as a proxy for OS [9]. The OS and PFS rates for the CAR-T therapy arms were based on Kaplan-Meier curves (median OS was not reached at the data cutoff; 27.7-month PFS rate: 54.9%) digitized from published clinical trial data [8-10] and reconstructed into pseudoindividual patient data to enable parametric survival modeling. Parametric survival models (including exponential, Weibull, Gompertz, gamma, log-logistic, log-normal, and generalized gamma distributions) were fitted to the reconstructed data to enable extrapolation beyond the observed trial period. Model selection was based on statistical goodness of fit (Akaike information criterion [AIC]), visual inspection of fitted vs observed Kaplan-Meier curves, and assessment of hazard functions. The analyses were then used to extrapolate the survival data beyond the time frame covered by clinical trials. Detailed methodology for extrapolation, data fitting, and model

XSL•FO
RenderX

selection can be found in Section A in Multimedia Appendix 1 [8,11-41].

The mortality rate for the PFD health state equals background mortality, retrieved from German life tables, obtained from the World Health Organization [42]. The mortality rate for the PD health state is cancer-specific mortality. The life tables used were from the most recent year available (2019) and for ages 61 to 100 years.

Figure 1. Partitioned survival model pathways and health states.



Data on health state utilities were derived from the CARTITUDE study [8] and the Dutch Health Care Institute (*ZIN*) cilta-cel study [5]. In these studies, the EQ-5D-5L questionnaire was used to assess utility scores, and mixed-effects models were used to estimate utilities for the PD state when data were missing [5]. Utilities were not age-adjusted. Data on efficacy and safety from the SoC therapies and cilta-cel CAR-T therapy were collected from publicly available trial results. Data from the AIDPATH platform will be collected once it is set up. When data are not available from the AIDPATH platform, estimates are based on cilta-cel.

Assuming similar clinical outcomes (PFS and OS), adverse event rates, and manufacturing success rates for AIDPATH and cilta-cel was an intentionally conservative modeling choice. In the base case, the manufacturing failure rate for AIDPATH was assumed to be equal to that of cilta-cel (3%) [43]. Manufacturing failures were not modeled as separate clinical events; instead, their impact was incorporated through increased production costs, reflecting the need for repeated manufacturing attempts to achieve a successful infusion. This assumption was made despite the AIDPATH platform incorporating a highly automated and standardized manufacturing process designed to reduce operator-dependent variability and production errors.

Consequently, assuming equivalent performance likely underestimates the potential clinical and operational advantages of the AIDPATH platform.

The costs for cilta-cel are based on the list price of €420,000 (€1=US $1.13; 2025 annual average) [5]. These costs are applied once. As AIDPATH does not have marketing authorization (yet), costs were based on production costs plus an estimation of marketing costs, totaling €120,000. The total cost of AIDPATH treatment was estimated at €120,000 based on a bottom-up aggregation of key cost components informed by published literature. Manufacturing (production) costs were assumed to be approximately €60,000 per treatment [11,44]. A 30% markup (€20,000) was applied to account for sales and marketing expenses [45]. Hospital-related costs were estimated at €27,000 [46], and quality control costs were approximately €13,000 [47]. Preinfusion period costs consisted of apheresis, bridging therapy, and conditioning therapy; these applied to both types of CAR-T therapies. Refer to Table S1 in Section B of Multimedia Appendix 1 for the cost breakdown.

Hospital resource use was based on data from the CARTITUDE-1 study [8], *ZIN* Carvykti economic analysis [5], and the AIDPATH consortium. Hospital costs and medication costs were based on the German fee tariff schedule (2023) [12], the *Kassenärtzliche Bundesvereinigung*, and publicly available economic evaluations [5,6] (Table 1). All costs were indexed to 2025 euros based on German inflation [48].

XSL•FO
RenderX

Table 1. Baseline parameters.

Parameters	Base case	Lower limit[a]	Upper limit	References
Utilities				
Pre-infusion period	0.763	0.675	0.841	[5]
PFS[b] (on treatment)	0.804	0.721	0.875	[5]
PFS (off treatment)	0.763	0.675	0.841	[5]
Progressed disease	0.653	0.558	0.743	[5]
Death	0	0	0	—[c]
Disutility CAR-T[d,e]	1.286	—	—	—
Disutility SoC[f]	0.225	0.149	0.311	—
General information				
Time horizon (years)	40	10	50	[5]
Discount rate outcomes (%)	3	0	5	[49]
Discount rate costs (%)	3	0	5	[49]
Average age (years)	61.70	37.51	85.89	[8]
Average body weight (kg)	80.7	49.07	112.33	[8]
Average body surface area (m^2)	1.91	1.16	2.66	[8]
AIDPATH[g] vein-to-vein time (days)	26	16	36	—
Cilta-cel vein-to-vein time (days)	51	31	71	[8]
Costs (€, 2025)[h]				
Leukapheresis	2612.00	—	—	[11]
AIDPATH treatment	120,000.00	—	—	—
Bridging chemotherapy (12 days)	3058.64	—	—	[8,13]
Conditioning lymphodepleting chemotherapy acquisition	1178.22	—	—	[6]
Cilta-cel treatment	420,000.00	—	—	[5]
Bridging chemotherapy (32 days)	—	—		[8,13]
Conditioning lymphodepleting chemotherapy acquisition	—	—		[6]
Administration lymphodepleting therapy	961.76	—	—	[14]
NUB[i]-fee for CAR-T infusion	2035.00	—	—	—
SoC drugs per week[j]	3506.10	—	—	—
Fifth line treatment (one time)	9087.80	—	—	[9,50,51]
Pre-infusion costs AIDPATH[k]	5198.62	—	—	—
Pre-infusion costs Cilta cel[k]	10,243.94	—	—	—
Hospital stay				
Inpatient care per day	186.76	—	—	[12]
Intensive care unit	1831.53	—	—	[12]
day care unit hematology and internal oncology	406.98	—	—	[12]
Palliative care/end of life	8163.92	—	—	[52]

[a]Lower and upper limits represent the values used in sensitivity analyses to characterize uncertainty around the base-case parameter estimate.
[b]PFS: progression-free survival.
[c]Not applicable.
[d]CAR-T: chimeric antigen receptor T-cell.

XSL•FO
RenderX

[e]Total average of treatment, spread over multiple cycles.

[f]SoC: standard of care.

[g]AIDPATH: AI-Powered Decentralized Production for Advanced Therapies in the Hospital.

[h]€1=US $1.13 (2025 annual average).

[i]NUB: Neue Untersuchungs- und Behandlungsmethoden (new diagnostic and treatment methods).

[j]Detailed calculation is provided in Section B in Multimedia Appendix 1.

[k]Detailed calculation is provided in Section C in Multimedia Appendix 1.

SoC encompasses the physician's choice of treatment for patients with triple-class refractory MM. These data were obtained from the German Therapy Monitor, a registry of real-world oncology data. It contains anonymized data on patients with approximately 4000 R/R MM, obtained from a representative sample of 108 German centers. Refer to Table S2 in Section C in Multimedia Appendix 1 for an overview of included drugs and Table S3 in Section C in Multimedia Appendix 1 for detailed calculations.

When patients enter the PD health state, costs for follow-up treatment were applied once for all treatments. The treatment given as fifth-line R/R MM treatment is one of the regimens given in the fourth line that the patient has not received before [9]. Therefore, the average fifth-line treatment price for SoC and CAR-T is the same as that for the fourth line. The average treatment duration in the fifth line is assumed to be 4 months [50]. Monthly costs of fifth-line treatment are €10,349.98. The percentage of patients with fourth-line R/R MM that receive fifth-line treatment is 0.22, according to attrition rates in a study by Dhanasiri [51]. This results in the applied 1-time follow-up treatment costs being €9108.00 (4×€10,349.98×0.22).

All baseline parameters are presented in Table 1.

The primary outcome of the analyses is the incremental cost-effectiveness ratio (ICER), which is calculated by dividing the difference in total costs (incremental cost) by the difference in quality-adjusted life years (QALYs; incremental effect) [53].

Secondary outcomes include the avoidance of adverse events. Adverse events were only included if they incurred health care costs (required physician input and/or hospitalization and/or medication) or had a large impact on patient quality of life (refer to Table S4 in Section D in Multimedia Appendix 1 for a list of included adverse events). As the occurrence rates of adverse events for AIDPATH are not yet available, they are assumed to be equal to those of cilta-cel.

Following the methodological guidelines of the German National Institute for Quality and Efficiency in Health Care (*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen* [*IQWiG*]), the German institute responsible for evaluating health care interventions, outcomes, and costs was discounted at 3% [49].

The model was constructed using Excel (version 2608; Microsoft Corp).

Sensitivity Analyses

A 1-way sensitivity analysis was performed to assess the model results for robustness and to identify the main cost drivers. Each parameter was varied by 20% [54]. Refer to Table 1 for the lower- and upper-limit values.

Scenario Analyses

Various scenarios were performed to assess the economic implications of the competing treatment strategies and to test the robustness of the different assumptions made in the model.

In the first scenario analysis, the risk of adverse events following AIDPATH-produced CAR-T cell therapy was decreased by 10% and 20%. CAR-T–specific side effects after AIDPATH-produced CAR-T cell infusion are expected to decrease. Furthermore, the CARs are equipped with an epidermal growth factor receptor marker, which may prevent fatal outcomes of CAR-T cell–related toxicity [55].

In the second scenario analysis, the shipping and manufacturing time of CAR-T cells produced by the AIDPATH platform was lowered. In this scenario, we assumed the manufacturing time (14 days) plus shipping time (0 days) as the bridging period for AIDPATH (the average bridging therapy duration for cilta-cel is 32 days). A shorter bridging period is expected to reduce the incidence of bridging-related adverse events. It also affects the number of patients who remain alive and eligible to receive lymphodepleting therapy and the subsequent CAR-T infusion. Because bridging therapy is shorter, the direct costs of bridging treatment decrease, along with the costs of managing adverse events associated with bridging chemotherapy.

To explore uncertainty in important assumptions related to AIDPATH, additional scenario analyses were performed. In the third scenario, the PFS rate of AIDPATH was varied by ±10% to assess the impact of potential differences in treatment efficacy. In the fourth scenario, the production costs of AIDPATH were increased by 10% and 20% to reflect a higher manufacturing failure rate. An increase in manufacturing failure rate implies that more production attempts are required per successfully treated patient, thereby increasing the cost per treatment. This approach allows an indirect assessment of the impact of manufacturing failure on cost-effectiveness outcomes.

To address uncertainty in the AIDPATH cost estimate, a fifth scenario analysis was added to explore the impact of the total costs of an AIDPATH treatment. To do so, costs were varied to €200,000 and €300,000.

Ethical Considerations

This study presents an early HTA based exclusively on secondary data obtained from previously published literature, publicly available sources, and aggregated data. No primary data were collected, and no individual-level patient data were accessed.

According to German and Dutch regulations governing research involving human participants, ethics approval from an ethics committee (*Ethikkommission*) is not required for studies that

XSL•FO
RenderX

rely solely on anonymized, aggregated, or publicly available data and do not involve interventions with human participants [56,57]. Therefore, formal ethics approval and informed consent were not required for this study.

Results

The mixture-cure log-normal distribution was selected for the extrapolation of PFS (AIC=324.56) and OS (AIC=431.23) data (refer to Figures S1 and S2 and Tables S1-S3 in Section E in Multimedia Appendix 1 for the results of the selection process).

Primary Result

The results indicate that AIDPATH-produced CAR-T cell therapy is projected to be both more effective and less costly than current SoC and cilta-cel treatment. Specifically, patients receiving AIDPATH therapy are expected to gain more QALYs while incurring lower overall treatment costs. This improvement in cost-effectiveness is reflected in the ICERs, which are negative for both comparisons: –€3075.08 when comparing AIDPATH to SoC and –€1,986,526.56 when comparing AIDPATH to cilta-cel. A negative ICER indicates that AIDPATH therapy dominates the alternatives, meaning it provides better health outcomes at a lower cost. Detailed results are presented in Table 2.

Most costs of the CAR-T cell therapies can be attributed to the acquisition, followed by adverse events and monitoring (Table 3).

Table 2. Primary outcomes in the base case analysis.

Treatments	Costs (€)[a]	Quality-adjusted life years	Incremental cost-effectiveness ratio (AIDPATH[b] compared to alternative)	Adverse event costs (€)[a]
AIDPATH	167,895.77	5.10	Reference	24,131.66
Cilta-cel	436,967.32	4.96	–1,986,526.56	24,131.66
Standard of care	180,114.69	1.13	–3075.08	4722.48

[a]€1=US $1.13 (2025 annual average).

[b]AIDPATH: AI-Powered Decentralized Production for Advanced Therapies in the Hospital.

Table 3. Cost breakdown of primary results.

Treatments	Acquisition of treatment (€)[a]	Administration of treatment (€)[a]	Monitoring of patient (€)[a]	Adverse event (€)[a]	Follow-up treatment (€)[a]	End of life (€)[a]
AIDPATH[b]	113,216.30	1927.01	12,864.01	24,131.66	9087.80	6668.99
Cilta-cel	382,779.98	1863.47	12,388.68	24,131.66	9087.80	6715.74
Standard of care	150,019.83	6125.16	2214.26	4772.48	9087.80	7898.01

[a]€1=US $1.13 (2025 annual average).

[b]AIDPATH: AI-Powered Decentralized Production for Advanced Therapies in the Hospital.

Sensitivity Analyses

An overview of the model input parameters with the largest impact on the ICER when comparing AIDPATH-produced CAR-T cell therapy to SoC is shown in Figure S3 and Table S8 in Section F in Multimedia Appendix 1. The results show that the discount rate of outcomes and PFS rate have the largest impact on the ICER. This indicates that assumptions about the valuation of future health benefits and long-term patient outcomes play a critical role in determining the overall cost-effectiveness of AIDPATH therapy.

An overview of the model input parameters with the largest impact on the ICER, when comparing AIDPATH therapy to cilta-cel CAR-T cell therapy, is shown in Figure S4 and Table S9 in Section F in Multimedia Appendix 1. The results show that the discount rate of outcomes and the vein-to-vein time have the largest impact on the ICER. This suggests that both the assumptions regarding how future health benefits are valued over time and the time required for manufacturing and delivering the CAR-T cell product play a crucial role in determining the relative cost-effectiveness of AIDPATH therapy compared to cilta-cel.

Scenario Analyses

The scenario analyses demonstrated that the results were generally robust to variations in key assumptions. Reducing adverse event risks (scenario 1) slightly lowered costs, with QALYs unchanged, and AIDPATH remained dominant vs both comparators. Shortening manufacturing and shipping times (scenario 2) increased QALYs with modest cost increases; AIDPATH remained dominant compared to cilta-cel but was less cost-saving compared to SoC.

Varying PFS rates (scenario 3) affected both costs and QALYs; however, AIDPATH remained dominant across all variations. Increasing production costs to reflect lower manufacturing success rates (scenario 4) raised total costs without affecting health outcomes. Despite these increases, AIDPATH remained cost-effective compared with cilta-cel and SoC. Increasing treatment prices (scenario 5) substantially increased costs (Table 4).

XSL•FO
RenderX

Table 4. Results of scenario analyses.

Scenarios and treatments	Costs (€)[a]	Quality-adjusted life years	Incremental cost-effectiveness ratio (compared to SoC[b])	AE[c] costs (€)
Scenario 1: lower risk of AEs with AIDPATH[d] CAR-T[e] therapy				
AIDPATH –10%	165,518.17	5.10	Reference	21,754.06
Cilta-cel	436,967.32	4.96	–1,938,922.50	24,131.66
SoC	180,114.69	1.13	–3676.71	4722.48
AIDPATH –20%	163,132.67	5.10	Reference	19,368.56
Cilta-cel	436,967.32	4.96	–1,955,961.79	24,131.66
SoC	180,114.69	1.13	–4277.59	4722.48
Scenario 2: shorter shipping and manufacturing time of AIDPATH CAR-T				
AIDPATH	172,296.18	5.25	Reference	24,131.66
Cilta-cel	436,967.32	4.96	–912,659.10	24,131.66
SoC	180,114.69	1.13	1897.70	4722.48
Scenario 3: variation in progression-free survival rates for AIDPATH treatment				
AIDPATH –10%	154,277.14	5.01	Reference	24,131.66
Cilta-cel	436,967.32	4.96	–5,653,803.60	24,131.66
SoC	180,114.69	1.13	–6659.16	4722.48
AIDPATH +10%	168,610.22	5.15	Reference	24,131.66
Cilta-cel	436,967.32	4.96	–1,412,405.79	24,131.66
SoC	180,114.69	1.13	–2861.81	4722.48
Scenario 4: variation in manufacturing success rate (ie, variation in production costs)				
AIDPATH +10%	173,556.58	5.10	Reference	24,131.66
Cilta-cel	436,967.32	4.96	–1,881,505.29	24,131.66
SoC	180,114.69	1.13	–1651.92	4722.48
AIDPATH +20%	179,217.40	5.10	Reference	24,131.66
Cilta-cel	436,967.32	4.96	–1,841,070.86	24,131.66
SoC	180,114.69	1.13	–226.02	4722.48
Scenario 5: increased costs of AIDPATH CAR-T treatment				
AIDPATH (€200,000)	243,373.30	5.10	Reference	24,131.66
Cilta-cel	436,967.32	4.96	–1,382.814.43	24,131.66
SoC	180,114.69	1.13	15,934.16	4722.48
AIDPATH (€300,000)	337,720.22	5.10	Reference	24,131.66
Cilta-cel	436,967.32	4.96	–708,907.86	24,131.66
SoC	180,114.69	1.13	36,699.13	4722.48

[a]€1=US $1.13 (2025 annual average).

[b]SoC: standard of care.

[c]AE: adverse event.

[d]AIDPATH: AI-Powered Decentralized Production for Advanced Therapies in the Hospital.

[e]CAR-T: chimeric antigen receptor T-cell.

Discussion

Principal Findings

The objective of this study was to assess the cost-effectiveness of CAR-T cell therapy produced by the AIDPATH platform for patients with triple-class refractory MM in Germany, compared to manually produced and centrally distributed CAR-T cell therapy (cilta-cel) and SoC. In this model, given its assumptions and from a hospital perspective, AIDPATH CAR-T therapy was shown to be dominant over a period of 40 years, offering improved outcomes and lower overall costs than both

XSL•FO
RenderX

comparators. Despite the inferiority, only a small difference was observed in QALY outcomes between the CAR-T treatments in the base case analysis. This is despite the expectation that shorter vein-to-vein time would increase the proportion of patients remaining eligible for infusion by limiting progression during the manufacturing period [2]. Due to limited evidence, equal eligibility was conservatively assumed, and thus, the observed QALY difference primarily reflects earlier treatment administration rather than differences in infusion eligibility.

The findings suggest that the AIDPATH platform could be considered for implementation within the German health care setting. The results may provide a basis for estimating the return on investment for relevant stakeholders. Importantly, the base case assumptions were intentionally conservative. Equivalent clinical outcomes (PFS and OS), adverse event rates, and manufacturing failure rates were assumed for AIDPATH and cilta-cel, despite the expectation that AIDPATH's automated and standardized manufacturing process may reduce variability and production errors in practice. The scenario analyses, which included further conservative adjustments, such as increased costs and limited efficacy gains, demonstrated that AIDPATH remained cost-effective, supporting the robustness of the findings.

The sensitivity analyses illustrate the sensitivity of the results to changes in methodological choices. In both sensitivity analyses, only a few parameters had a large impact on the outcomes. These parameters are important to consider when modifying or scaling up the platform to safeguard its cost-effectiveness compared to alternative treatment strategies.

The results of this study can be used to guide discussions with German stakeholders, including investors and policymakers, regarding the incremental cost-effectiveness of CAR-T cell therapy and reimbursement decisions. Furthermore, the findings can be considered exemplary for countries with comparable health care systems. Finally, AIDPATH uses a patient-centric approach to CAR-T cell production, and based on the results, this approach shows potential.

Strengths and Limitations

To our knowledge, this is the first early HTA study to assess the cost-effectiveness of AI-supported CAR-T cell therapy production in Germany from the hospital perspective. This study is tailored to the German health care setting and provides a model that can be updated when new data become available.

Although this cost-effectiveness study provides valuable insights to support decision-making, it has several important limitations that must be acknowledged. First, by focusing narrowly on cost-effectiveness, the study may overlook the broader potential benefits of automated CAR-T therapy production. These may include system-level advantages associated with decentralized manufacturing, such as reduced logistical complexity, lower risks related to cold chain management and cross-border shipment, and improved access to treatment through reduced patient travel or expanded treatment capacity. In addition, automated production may improve treatment convenience and patient satisfaction and streamline clinical workflows. These qualitative outcomes can significantly impact health care delivery and patient experience but are often difficult to adequately quantify in economic evaluations [58]. Moreover, as with many applications of medical AI, there are pressing concerns about appropriate regulation and evaluation [59]. HTA has been recognized as a potentially key method for evaluating AI-driven innovations. However, this model does not fully address AI-specific challenges, such as the assessment of continuously learning systems, the explainability of AI decision-making, and interoperability with existing health care infrastructure. To address these gaps, Boverhof et al [59] proposed a comprehensive framework of 29 evaluation elements across 4 domains, developed through a systematic and targeted review, which could serve as a more nuanced approach for assessing the value of AI in medicine. This broader perspective is crucial for ensuring that automated CAR-T cell production and similar technologies are appraised not only for their economic impact but also for their systemic and societal implications.

Second, no specific data regarding the safety and effectiveness of CAR-T therapy generated by the AIDPATH platform are available yet. The platform is currently being tested, and it is important to update the model as soon as clinical trial data become available. In addition, survival extrapolation is an important source of uncertainty in this analysis. A mixture-cure log-normal model was selected based on statistical goodness-of-fit, visual inspection, and clinical plausibility. Alternative parametric models were explored during model selection; however, these showed inferior fit and/or less plausible long-term extrapolations.

Third, a limitation in achieving the objective of this study is the reliance on Dutch CARTITUDE trial data for key comparator inputs, specifically for the centrally manufactured CAR-T therapy cilta-cel [5]. Due to the lack of German-specific data, key cost and resource use parameters for cilta-cel were derived from a Dutch study, and these values were also used as a proxy for the AIDPATH-generated CAR-T cell therapy. However, this approach introduces uncertainty, as health care delivery, cost structures, and reimbursement mechanisms differ between the Dutch and German hospital systems. For example, hospital tariffs, staff costs, and care pathways in Germany may be higher or more complex, potentially leading to an underrepresentation of true costs from the German hospital perspective. This could result in a more favorable cost-effectiveness profile for CAR-T cell therapies than would be observed with German-specific data. Future studies should aim to validate these inputs with empirical data from the German context to strengthen the reliability of the cost-effectiveness conclusions. Because this approach was used for both CAR-T cell therapy arms, it is unlikely that an updated model would change the conclusions of the current study.

Despite these limitations, an HTA at this early stage is considered highly valuable because it provides critical insights into the potential clinical and economic impact of this new medical product. It supports informed decision-making during early development stages, helps prioritize investments, guides research efforts, and identifies key data gaps to address before implementation [60].

XSL•FO
RenderX

Conclusions

This study is the first to assess the cost-effectiveness of AI-supported, hospital-based CAR-T cell therapy produced by the AIDPATH platform for patients with MM in Germany from the hospital perspective. On the basis of the assumptions inherent in this model, AIDPATH therapy was found to be cost-effective and dominant compared with both cilta-cel and SoC from the hospital perspective over a 40-year time horizon. This was found in the base case analysis and was confirmed by all scenario analyses performed.

However, key inputs were partly based on Dutch data due to the lack of German-specific figures, introducing uncertainty, particularly in cost estimates. Additionally, no clinical data are yet available for AIDPATH-generated CAR-T cell therapy, and broader benefits, such as patient satisfaction and system efficiency, were not captured in the cost-effectiveness analysis.

Despite these challenges, this early HTA provides valuable insights to guide further research, inform reimbursement discussions, and support planning for future implementation in Germany and similar health care systems.

Acknowledgments

The authors acknowledge all partners who collaborated in the European Union Horizon 2020 AIDPATH (AI-Powered Decentralized Production for Advanced Therapies in the Hospital) project. During the preparation of this manuscript, the authors used ChatGPT (version 5-5; OpenAI) to polish the wording and formulation of sentences for English grammar and clarity. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

The AIDPATH project was funded by the European Union (grant 101016909). The authors declare that no additional financial support was received for the preparation of this manuscript.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

Conceptualization: IWAB, YdG, HJMV
Data curation: IWAB, YdG
Formal analysis: IWAB, YdG
Supervision: HJMV
Writing—original draft: IWAB
Writing—review and editing: IWAB, YdG, ALW, HJMV

Conflicts of Interest

None declared.

Multimedia Appendix 1
A: Methodology for extrapolation, data fitting, and model selection of overall survival (OS) and progression-free survival (PFS) data; B: preinfusion costs; C: standard of care drugs; D: adverse events; E: results of extrapolation, data fitting, and model selection for OS and PFS data; and F: scenario analysis results.
[DOCX File , 18477 KB - cancer_v12i1e92253_app1.docx]

References

1. Cappell KM, Kochenderfer JN. Long-term outcomes following CAR T cell therapy: what we know so far. Nat Rev Clin Oncol 2023 Jun;20(6):359-371 [FREE Full text] [doi: 10.1038/s41571-023-00754-1] [Medline: 37055515]
2. Zhang X, Zhang H, Lan H, Wu J, Xiao Y. CAR-T cell therapy in multiple myeloma: current limitations and potential strategies. Front Immunol 2023 Feb 20;14:1101495 [FREE Full text] [doi: 10.3389/fimmu.2023.1101495] [Medline: 36891310]
3. Hort S, Herbst L, Bäckel N, Erkens F, Niessing B, Frye M, et al. Toward rapid, widely available autologous CAR-T cell therapy - artificial intelligence and automation enabling the smart manufacturing hospital. Front Med (Lausanne) 2022 Jun 6;9:913287 [FREE Full text] [doi: 10.3389/fmed.2022.913287] [Medline: 35733863]
4. Minacori R, Bonastre J, Lueza B, Marguet S, Levy P. How to model survival in cost-effectiveness analysis? Differences between Markov and partitioned survival analysis models. Value Health 2015 Nov;18(7):A704. [doi: 10.1016/j.jval.2015.09.2639]

XSL•FO
RenderX

5. Diemer FS. Farmacotherapeutisch rapport ciltacabtagene autoleucel (Carvykti®) bij recidiverend en refractair multipel myeloom na ten minste drie eerdere behandelingen [Article in Dutch]. Zorginstituut Nederland. 2022. URL: https://tinyurl.com/mv33y36s [accessed 2026-04-23]
6. Ciltacabtagene Autoleucel (CARVYKTI®) [Article in German]. Gemeinsamer Bundesausschuss. 2023. URL: https://www.g-ba.de/downloads/92-975-6400/2023_02_14_Janssen_Dossier_Ciltacel_rrMM_Modul3.pdf [accessed 2026-04-23]
7. Herbolsheimer F, Schaap LA, Edwards MH, Maggi S, Otero Á, Timmermans EJ, et al. Physical activity patterns among older adults with and without knee osteoarthritis in six European countries. Arthritis Care Res (Hoboken) 2016 Feb;68(2):228-236 [FREE Full text] [doi: 10.1002/acr.22669] [Medline: 26212673]
8. Berdeja JG, Madduri D, Usmani SZ, Jakubowiak A, Agha M, Cohen AD, et al. Ciltacabtagene autoleucel, a B-cell maturation antigen-directed chimeric antigen receptor T-cell therapy in patients with relapsed or refractory multiple myeloma (CARTITUDE-1): a phase 1b/2 open-label study. Lancet 2021 Jul 24;398(10297):314-324. [doi: 10.1016/S0140-6736(21)00933-8] [Medline: 34175021]
9. Merz M, Goldschmidt H, Hari P, Agha M, Diels J, Ghilotti F, et al. Adjusted comparison of outcomes between patients from CARTITUDE-1 versus multiple myeloma patients with prior exposure to PI, Imid and anti-CD-38 from a German registry. Cancers (Basel) 2021 Nov 29;13(23):5996 [FREE Full text] [doi: 10.3390/cancers13235996] [Medline: 34885106]
10. Martin T, Usmani SZ, Berdeja JG, Agha M, Cohen AD, Hari P, et al. Ciltacabtagene autoleucel, an anti-B-cell maturation antigen chimeric antigen receptor T-cell therapy, for relapsed/refractory multiple myeloma: CARTITUDE-1 2-year follow-up. J Clin Oncol 2023 Feb 20;41(6):1265-1274 [FREE Full text] [doi: 10.1200/JCO.22.00842] [Medline: 35658469]
11. Ran T, Eichmüller SB, Schmidt P, Schlander M. Cost of decentralized CAR T-cell production in an academic nonprofit setting. Int J Cancer 2020 Dec 15;147(12):3438-3445. [doi: 10.1002/ijc.33156] [Medline: 32535920]
12. DRG-Entgelttarif 2022 für die Klinikum Würzburg Mitte gGmbH [Article in German]. Klinikum Würzburg Mitte. 2022. URL: https://www.kwm-klinikum.de/assets/downloads/KWM%20DRG%20Entgelt%2001.8.2022.pdf [accessed 2026-04-23]
13. Jagannath S, Joseph N, Crivera C, Kharat A, Jackson CC, Valluri S, et al. Component costs of CAR-T therapy in addition to treatment acquisition costs in patients with multiple myeloma. Oncol Ther 2023 Jun;11(2):263-275 [FREE Full text] [doi: 10.1007/s40487-023-00228-5] [Medline: 37014590]
14. Tisagenlecleucel (Kymriah®). Gemeinsamer Bundesausschuss. 2022. URL: https://www.g-ba.de/downloads/92-975-5819/2022_05_25_Modul3A_Tisagenlecleucel.pdf [accessed 2026-04-23]
15. Latimer NR. Survival Analysis For Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data. London, UK: National Institute for Health and Care Excellence (NICE); 2013.
16. Rutherford MJ, Dickman PW, Coviello E, Lambert PC. Estimation of age-standardized net survival, even when age-specific data are sparse. Cancer Epidemiol 2020 Aug;67:101745. [doi: 10.1016/j.canep.2020.101745] [Medline: 32554300]
17. Jackson C, Thom H, V G. Different models for different purposes: focused model comparison in R. MRC Biostatistics Unit. URL: https://r-mirror.zim.uni-due.de/web/packages/fic/vignettes/fic.pdf [accessed 2026-09-23]
18. Medicinal Products Act in the version published on 12 December 2005 (Federal Law Gazette [BGBl.]) Part I p. 3394, last amended by Article 11 of the Act of 6 May 2019 (Federal Law Gazette I p. 646). LawEuro. 2021. URL: https://laweuro.com/?p=14930 [accessed 2026-09-23]
19. Hecken. Justification to the resolution of the Federal Joint Committee (G-BA) on an amendment of the pharmaceuticals directive (AM-RL): Annex XII – benefit assessment of medicinal products with new active ingredients according to section 35a SGB V polatuzumab vedotin (diffuse large B-Cell lymphoma, combination with bendamustine and rituximab). Gemeinsamer Bundesausschuss. 2020. URL: https://www.g-ba.de/downloads/40-1465-6796/2020-08-20_AM-RL-XII_Polatuzumab-Vedotin_D-507_TrG_EN.pdf [accessed 2026-09-23]
20. Justification to the resolution of the Federal Joint Committee (G-BA) on an amendment of the pharmaceuticals directive (AM-RL): Annex XII – benefit assessment of medicinal products with new active ingredients according to section 35a SGB V daratumumab (new therapeutic indication: multiple myeloma, newly diagnosed, patients eligible for autologous stem cell transplant, combination with bortezomib, thalidomide, and dexamethasone). Gemeinsamer Bundesausschuss. 2020. URL: https://www.g-ba.de/downloads/40-1465-6794/2020-08-20_AM-RL-XII_Daratumumab_D-522_TrG_EN.pdf [accessed 2026-09-23]
21. -. GmbH J-C. Daratumumab (Darzalex®) {German] 2018.
22. Palumbo A, Bringhen S, Larocca A, Rossi D, Di Raimondo F, Magarotto V, et al. Bortezomib-melphalan-prednisone-thalidomide followed by maintenance with bortezomib-thalidomide compared with bortezomib-melphalan-prednisone for initial treatment of multiple myeloma: updated follow-up and improved survival. J Clin Oncol 2014 Mar 01;32(7):634-640 [FREE Full text] [doi: 10.1200/JCO.2013.52.0023] [Medline: 24449241]
23. Durie BG, Hoering A, Abidi MH, Rajkumar SV, Epstein J, Kahanic SP, et al. Bortezomib with lenalidomide and dexamethasone versus lenalidomide and dexamethasone alone in patients with newly diagnosed myeloma without intent for immediate autologous stem-cell transplant (SWOG S0777): a randomised, open-label, phase 3 trial. Lancet 2017 Feb 04;389(10068):519-527 [FREE Full text] [doi: 10.1016/S0140-6736(16)31594-X] [Medline: 28017406]
24. Usmani SZ, Weiss BM, Plesner T, Bahlis NJ, Belch A, Lonial S, et al. Clinical efficacy of daratumumab monotherapy in patients with heavily pretreated relapsed or refractory multiple myeloma. Blood 2016 Jul 07;128(1):37-44 [FREE Full text] [doi: 10.1182/blood-2016-03-705210] [Medline: 27216216]

XSL•FO
RenderX

25. Chang HJ, Lee JH, Do YR, Bae SH, Lee JL, Nam SH, et al. A combination of melphalan, prednisone, and 50 mg thalidomide treatment in non-transplant-candidate patients with newly diagnosed multiple myeloma. Korean J Intern Med 2011 Dec;26(4):403-409 [FREE Full text] [doi: 10.3904/kjim.2011.26.4.403] [Medline: 22205840]
26. Moreau P, Masszi T, Grzasko N, Bahlis NJ, Hansson M, Pour L, et al. Oral ixazomib, lenalidomide, and dexamethasone for multiple myeloma. N Engl J Med 2016 Apr 28;374(17):1621-1634. [doi: 10.1056/NEJMoa1516282] [Medline: 27119237]
27. Multiple myeloma [Article in Dutch]. Nederlandse Vereniging voor Hematologie. 2024. URL: https://publicatie.hematologienederland.nl/richtlijnen/multipel-myeloom/ [accessed 2026-09-23]
28. Mateos MV, Masszi T, Grzasko N, Hansson M, Sandhu I, Pour L, et al. Impact of prior therapy on the efficacy and safety of oral ixazomib-lenalidomide-dexamethasone vs. placebo-lenalidomide-dexamethasone in patients with relapsed/refractory multiple myeloma in TOURMALINE-MM1. Haematologica 2017 Oct;102(10):1767-1775 [FREE Full text] [doi: 10.3324/haematol.2017.170118] [Medline: 28751562]
29. Latif A, Kapoor V, Lateef N, Ahsan MJ, Usman RM, Malik SU, et al. Incidence and management of carfilzomib-induced cardiovascular toxicity; a systematic review and meta-analysis. Cardiovasc Hematol Disord Drug Targets 2021;21(1):30-45. [doi: 10.2174/1871529X21666210412113017] [Medline: 33845729]
30. Stewart AK, Rajkumar SV, Dimopoulos MA, Masszi T, Špička I, Oriol A, et al. Carfilzomib, lenalidomide, and dexamethasone for relapsed multiple myeloma. N Engl J Med 2015 Jan 08;372(2):142-152. [doi: 10.1056/NEJMoa1411321] [Medline: 25482145]
31. Dimopoulos MA, Lonial S, White D, Moreau P, Weisel K, San-Miguel J, et al. Elotuzumab, lenalidomide, and dexamethasone in RRMM: final overall survival results from the phase 3 randomized ELOQUENT-2 study. Blood Cancer J 2020 Sep 04;10(9):91 [FREE Full text] [doi: 10.1038/s41408-020-00357-4] [Medline: 32887873]
32. San-Miguel JF, Hungria VT, Yoon SS, Beksac M, Dimopoulos MA, Elghandour A, et al. Panobinostat plus bortezomib and dexamethasone: impact of dose intensity and administration frequency on safety in the PANORAMA 1 trial. Br J Haematol 2017 Oct;179(1):66-74. [doi: 10.1111/bjh.14821] [Medline: 28653400]
33. Richardson PG, Attal M, Campana F, Le-Guennec S, Hui AM, Risse ML, et al. Isatuximab plus pomalidomide/dexamethasone versus pomalidomide/dexamethasone in relapsed/refractory multiple myeloma: ICARIA phase III study design. Future Oncol 2018 May;14(11):1035-1047 [FREE Full text] [doi: 10.2217/fon-2017-0616] [Medline: 29268619]
34. Palumbo A, Chanan-Khan A, Weisel K, Nooka AK, Masszi T, Beksac M, et al. Daratumumab, bortezomib, and dexamethasone for multiple myeloma. N Engl J Med 2016 Aug 25;375(8):754-766. [doi: 10.1056/NEJMoa1606038] [Medline: 27557302]
35. Kim SJ, Bang SM, Choi YS, Jo DY, Kim JS, Lee H, et al. Bendamustine in heavily pre-treated multiple myeloma patients: results of a retrospective analysis from the Korean Multiple Myeloma Working Party. Blood Res 2016 Sep;51(3):193-199 [FREE Full text] [doi: 10.5045/br.2016.51.3.193] [Medline: 27722131]
36. Mickisch G, Gore M, Escudier B, Procopio G, Walzer S, Nuijten M. Costs of managing adverse events in the treatment of first-line metastatic renal cell carcinoma: bevacizumab in combination with interferon-alpha2a compared with sunitinib. Br J Cancer 2010 Jan 05;102(1):80-86 [FREE Full text] [doi: 10.1038/sj.bjc.6605417] [Medline: 19920817]
37. Campone M, Yang H, Faust E, Kageleiry A, Signorovitch JE, Zhang J, et al. Cost of adverse events during treatment with everolimus plus exemestane or single-agent chemotherapy in patients with advanced breast cancer in Western Europe. J Med Econ 2014 Dec;17(12):837-845 [FREE Full text] [doi: 10.3111/13696998.2014.959589] [Medline: 25164472]
38. Wehler E, Zhao Z, Pinar Bilir S, Munakata J, Barber B. Economic burden of toxicities associated with treating metastatic melanoma in eight countries. Eur J Health Econ 2017 Jan;18(1):49-58 [FREE Full text] [doi: 10.1007/s10198-015-0757-y] [Medline: 26721505]
39. Ihbe-Heffinger A, Paessens B, Jacobs VR, Ehlken B, Shlaen M, Berger K, et al. Resource use and costs associated with routine management of febrile neutropenia in German hospitals. Blood 2007;110(11):3338. [doi: 10.1182/blood.V110.11.3338.3338]
40. Deb A, Podmore B, Barnett R, Beier D, Galetzka W, Qizilbash N, et al. Clinical and economic burden of pneumococcal disease among individuals aged 16 years and older in Germany. Epidemiol Infect 2022 Nov 08;150:e204 [FREE Full text] [doi: 10.1017/S0950268822001182] [Medline: 36345842]
41. Fleischmann C, Hartmann M, Hartog CS, Welte T, Heublein S, Thomas-Rueddel D, et al. Epidemiology of sepsis in Germany: incidence, mortality and associated costs of care 2007-2013. Intensive Care Med Exp 2015 Oct 1;3(S1):A50. [doi: 10.1186/2197-425X-3-S1-A50]
42. The old GHO Minerva interface and GHO Athena API are retired. World Health Organization. URL: https://www.who.int/data/gho/legacy?lang=en [accessed 2026-04-23]
43. Lopes AG, Noel R, Sinclair A. Cost analysis of vein-to-vein CAR T-cell therapy: automated manufacturing and supply chain. Cell Gene Ther Insights 2020 Jun 02;6(3):487-510. [doi: 10.18609/cgti.2020.058]
44. Weltin AL, De Graaf Y, Goudarzi A, Müller M, Herbst L, Nießing B, et al. Industrializing CAR-T cell therapy: impact of automation on cost and space efficiency of manufacturing facilities. Front Bioeng Biotechnol 2026 Jan 6;13:1612248 [FREE Full text] [doi: 10.3389/fbioe.2025.1612248] [Medline: 41568231]

XSL•FO
RenderX

45. Thielen FW, Heine RJ, Berg SV, Ham RM, Groot CA. Towards sustainability and affordability of expensive cell and gene therapies? Applying a cost-based pricing model to estimate prices for Libmeldy and Zolgensma. Cytotherapy 2022 Dec;24(12):1245-1258 [FREE Full text] [doi: 10.1016/j.jcyt.2022.09.002] [Medline: 36216697]
46. Roedl K, Ahmadi P, Essmann S, Aamir S, Haar M, Ayuk F, et al. Economic evaluation of critically ill adult CAR-T cell recipients-analysis from a healthcare payer perspective. Med Klin Intensivmed Notfmed 2025 Dec;120(Suppl 1):1-7. [doi: 10.1007/s00063-024-01230-z] [Medline: 39739037]
47. Nikoniuk A, Lilova K, Thomas M, Szita N. The potential of microfluidic cell analysis in CAR T cell therapy manufacturing. Front Bioeng Biotechnol 2025 Oct 29;13:1613836 [FREE Full text] [doi: 10.3389/fbioe.2025.1613836] [Medline: 41234554]
48. GENESIS Online. Statistisches Bundesamt (Destatis). URL: https://www-genesis.destatis.de/datenbank/online/ [accessed 2026-04-23]
49. General methods. IQWiG. URL: https://www.iqwig.de/en/methods/methods-paper/ [accessed 2026-04-23]
50. Yong K, Delforge M, Driessen C, Fink L, Flinois A, Gonzalez-McQuire S, et al. Multiple myeloma: patient outcomes in real-world practice. Br J Haematol 2016 Oct;175(2):252-264 [FREE Full text] [doi: 10.1111/bjh.14213] [Medline: 27411022]
51. Dhanasiri S, Hollier-Hann G, Stothard C, Dhanda DS, Davies FE, Rodriguez-Otero P. Treatment patterns and outcomes in triple-class exposed patients with relapsed and refractory multiple myeloma: findings from the multinational ITEMISE study. Clin Ther 2021 Nov;43(11):1983-96.e3 [FREE Full text] [doi: 10.1016/j.clinthera.2021.09.013] [Medline: 34736769]
52. Becker C, Leidl R, Schildmann E, Hodiamont F, Bausewein C. A pilot study on patient-related costs and factors associated with the cost of specialist palliative care in the hospital: first steps towards a patient classification system in Germany. Cost Eff Resour Alloc 2018 Oct 16;16:35 [FREE Full text] [doi: 10.1186/s12962-018-0154-3] [Medline: 30349423]
53. Incremental cost-effectiveness ratio (ICER). York Health Economics Consortium. 2016 Oct. URL: https://yhec.co.uk/glossary/incremental-cost-effectiveness-ratio-icer/ [accessed 2025-04-23]
54. Boot IW, Planken RN, den Hartog AW, Vrijhoef HJ. Early-stage health technology assessment of fractional flow reserve coronary computed tomography versus standard diagnostics in patients with stable chest pain in the Netherlands. PLoS One 2024 Jun 13;19(6):e0305189 [FREE Full text] [doi: 10.1371/journal.pone.0305189] [Medline: 38870138]
55. Shabaneh TB, Moffett HF, Stull SM, Derezes T, Tait LJ, Park S, et al. Safety switch optimization enhances antibody-mediated elimination of CAR T cells. Front Mol Med 2022 Oct 11;2:1026474 [FREE Full text] [doi: 10.3389/fmmed.2022.1026474] [Medline: 39086975]
56. Your research: subject to the WMO or not? Central Committee on Research Involving Human Subjects. URL: https://english.ccmo.nl/investigators/legal-framework-for-medical-scientific-research/your-research-is-it-subject-to-the-wmo-or-not [accessed 2026-04-23]
57. Obtaining ethical approval. KonsortSWD. URL: https://www.konsortswd.de/en/topics/best-practices-research-ethics/researchers/ethical-approval/ [accessed 2026-04-23]
58. Neumann PJ, Garrison LP, Willke RJ. The history and future of the "ISPOR value flower": addressing limitations of conventional cost-effectiveness analysis. Value Health 2022 Apr;25(4):558-565 [FREE Full text] [doi: 10.1016/j.jval.2022.01.010] [Medline: 35279370]
59. Boverhof BJ, Redekop WK, Visser JJ, Uyl-de Groot CA, Rutten-van Mölken MP. Broadening the HTA of medical AI: a review of the literature to inform a tailored approach. Health Policy Technol 2024 Jun;13(2):100868. [doi: 10.1016/j.hlpt.2024.100868]
60. Gavan SP. Supporting translation with early health technology assessment: from definition to action. Int J Technol Assess Health Care 2025 Jun 02;41(1):e39. [doi: 10.1017/S0266462325100147] [Medline: 40452623]

Abbreviations

AIC: Akaike information criterion
AIDPATH: AI-Powered Decentralized Production for Advanced Therapies in the Hospital
CAR: chimeric antigen receptor
CAR-T: chimeric antigen receptor T-cell
Cilta-cel: ciltacabtagene autoleucel
HTA: health technology assessment
ICER: incremental cost-effectiveness ratio
IQWiG: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
MM: multiple myeloma
OS: overall survival
PD: progressive disease
PFD: progression-free disease
PFS: progression-free survival
QALY: quality-adjusted life year
R/R: relapse/refractory
SoC: standard of care

XSL•FO
RenderX

ZIN: Zorginstituut Nederland

Edited by M Balcarras; submitted 27.Jan.2026; peer-reviewed by J Guo, C Br; comments to author 18.Mar.2026; revised version received 07.May.2026; accepted 13.May.2026; published 13.Oct.2026.

Please cite as:
Boot IWA, Weltin AL, de Graaf Y, Vrijhoef HJM
Cost-Effectiveness of AI-Powered Decentralized Production for Advanced Therapies in the Hospital–Generated Chimeric Antigen Receptor T-Cell Therapy for Multiple Myeloma From a German Hospital Perspective: Early Health Technology Assessment Study
JMIR Cancer 2026;12:e92253
URL: https://cancer.jmir.org/2026/1/e92253
doi:10.2196/92253
PMID:

©Iris W A Boot, Anna Louisa Weltin, Ysanne de Graaf, Hubertus J M Vrijhoef. Originally published in JMIR Cancer (https://cancer.jmir.org), 13.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Detection of Cervical Nodal Metastasis in Oral Squamous Cell Carcinoma Using Elastic Scattering Spectroscopy: Ex Vivo Diagnostic Modeling Study

Waseem Jerjes[1], PhD; Zaid Hamdoon[2], PhD; Dara Rashed[3], PhD; Colin Hopper[3], MD

[1]Department of Primary Care and Public Health, Imperial College London, 90 Wood Lane, London, United Kingdom

[2]College of Dental Medicine, University of Sharjah, Sharjah, United Arab Emirates

[3]Unit of Oral and Maxillofacial Surgery, UCL Eastman Dental Institute, London, United Kingdom

Corresponding Author:
Waseem Jerjes, PhD
Department of Primary Care and Public Health, Imperial College London, 90 Wood Lane, London, United Kingdom

Abstract

Background: Cervical nodal metastasis remains one of the strongest determinants of treatment planning and prognosis in oral squamous cell carcinoma. Histopathology remains the reference standard, but it does not provide immediate information at the point of surgical specimen assessment. Elastic scattering spectroscopy (ESS) offers rapid, label-free optical interrogation of tissue microstructure, and may support ex vivo nodal assessment.

Objective: This study aimed to evaluate the diagnostic accuracy of ESS combined with machine learning classification for ex vivo detection of cervical lymph node metastasis in patients with oral squamous cell carcinoma.

Methods: We conducted a single-center observational ex vivo diagnostic modeling study in patients with oral squamous cell carcinoma undergoing primary tumor resection and neck dissection. A total of 3 ESS spectra were acquired from each bivalved lymph node before routine histopathological processing. Internal validation used repeated nested 5-fold cross-validation grouped by patient, so every spectrum and lymph node from a patient remained in the same fold. All preprocessing, scaling, dimensionality reduction, within-model-family configuration, and hyperparameter tuning, probability calibration where required, and decision-threshold selection were restricted to training data within the nested validation process; comparisons between candidate model families were exploratory and were based on pooled patient-grouped out-of-fold predictions. Spectrum-level probabilities were aggregated within each lymph node, and histopathology was the reference standard.

Results: A total of 52 patients were assessed for eligibility, 45 were included, and 280 fresh lymph nodes were assessed by ESS; 96 were pathology positive, and 184 were pathology negative. Of 840 acquired spectra, 760 passed quality control. Of the 280 nodes, 274 retained at least 2 usable spectra and comprised the analytical cohort; 6 pathology-negative nodes yielded no analyzable node-level ESS result. The hybrid model was selected because it had the highest pooled patient-grouped out-of-fold node-level area under the receiver operating characteristic curve (AUROC), with an internally validated, exploratory estimate of 0.966 (95% CI 0.929 - 0.988). The same out-of-fold predictions informed model-family selection and performance reporting; this estimate does not represent independent validation of a prospectively fixed final model. At the cross-validation–derived summary threshold, secondary descriptive operating characteristics were sensitivity 93.8% (95% CI 86.4% - 97.6%), specificity 95.5% (95% CI 90.1% - 98.3%), accuracy 94.9% (95% CI 90.4% - 97.3%), positive predictive value 91.8% (95% CI 82.9% - 96.3%), and negative predictive value 96.6% (95% CI 92.4% - 98.8%). The Brier score was 0.063, with calibration-in-the-large of –0.03 and a calibration slope of 0.94. Patient-level sensitivity was 96% and specificity was 90%. Pathology-positive and pathology-negative nodes also showed differences across amplitude-related and shape-related spectral features.

Conclusions: ESS showed promising ex vivo discrimination of metastatic cervical lymph nodes in this single-center model-development and internal-validation study. The findings support further investigation of ESS as a potential pathology-triage adjunct rather than a replacement for histopathology.

(JMIR Cancer 2026;12:e103960) doi:10.2196/103960

KEYWORDS

oral squamous cell carcinoma; cervical lymph node metastasis; elastic scattering spectroscopy; optical biopsy; diagnostic accuracy; decision support systems; optical diagnostics; head and neck cancer; ex vivo tissue assessment

XSL•FO
RenderX

Introduction

Head and neck squamous cell carcinoma remains a major global cancer burden, and oral squamous cell carcinoma is one of its most clinically demanding subtypes because small differences in regional disease can substantially alter treatment. Contemporary global estimates confirm the scale of cancer incidence and mortality, while epidemiological reviews continue to identify tobacco, alcohol, areca nut, human papillomavirus infection, and social determinants as important contributors to disease distribution and late presentation [1,2]. For the surgeon and oncologist, however, the decisive issue at the time of treatment planning is often not the primary tumor alone, but whether malignant disease has reached the cervical lymphatic system. Once cervical nodal metastasis is identified, staging, operative planning, adjuvant radiotherapy or chemoradiotherapy, surveillance intensity, and prognosis all change [3].

The clinical importance of the neck in oral cancer is reflected in the depth of work examining predictors and patterns of nodal metastasis. Recent systematic reviews show that histopathological features of the primary tumor, including depth of invasion and adverse growth patterns, are associated with cervical lymph node metastasis, and that nodal involvement varies across neck levels in ways that influence decisions about the extent of neck dissection [4,5]. Yet the fundamental problem remains unresolved—occult or small-volume nodal metastases may be difficult to identify before treatment, while enlarged or reactive nodes may be mistaken for malignant disease. The consequence is a familiar clinical tension between undertreatment of metastatic disease and overtreatment of patients whose necks are ultimately histologically negative.

Preoperative imaging has improved this balance but has not eliminated uncertainty. Computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography-computed tomography (PET-CT), ultrasound, and radiomics-based approaches provide valuable information, and PET-CT may offer high specificity and negative predictive value (NPV) in selected clinically node-negative settings [6,7]. However, meta-analyses and radiomics reviews also highlight variable sensitivity, dependence on tumor site and study design, and a continuing need for prospective validation before advanced computational imaging can be relied upon for occult nodal disease [6-8]. Imaging therefore remains essential, but it cannot replace detailed histopathological examination of dissected lymph nodes. Equally, histopathology is accurate but not immediate. Routine processing, sectioning, and reporting take time, and the workload generated by multiple lymph nodes can be considerable.

A second challenge is that nodal assessment is intrinsically heterogeneous. A neck dissection specimen may contain many nodes of different sizes and anatomical levels, and metastatic deposits may be focal, central, peripheral, or microscopic. Standard histopathology is designed to manage this complexity, but the process depends on careful grossing, representative sectioning, and prioritization of tissue that appears suspicious. A rapid adjunctive technology would therefore be valuable if it could identify nodes with a high probability of metastasis before or during routine processing. Such a test would not need to provide the final diagnosis to be useful; it would need to improve where attention is directed, how quickly suspicious nodes are reviewed, and how confidently low-risk nodes are handled within established pathology protocols.

Optical diagnostic technologies have been developed to address exactly this gap; the need for rapid, biologically meaningful information from tissue without waiting for full laboratory processing. Recent reviews of oral cancer diagnostics and head and neck photodiagnosis describe several label-free or minimally invasive modalities, including Raman spectroscopy, fluorescence, hyperspectral imaging, and infrared spectroscopy combined with machine learning [9-13]. These techniques exploit changes in tissue architecture, scattering, absorption, and molecular composition that accompany dysplasia, invasion, and metastatic transformation. Their appeal in head and neck oncology is not that they replace histopathology, but that they may provide a near-immediate adjunct to help guide sampling, prioritize suspicious tissue, and improve the efficiency of diagnostic pathways. In lymph nodes, such optical signals may reflect not only the presence of malignant cells themselves but also the architectural consequences of metastatic replacement, including altered cellular density, stromal organization, vascularity, and loss of normal nodal microstructure.

Elastic scattering spectroscopy (ESS) is particularly suited to this role because it is rapid, contact-based, and sensitive to cellular and subcellular structure. When light enters tissue, it may be scattered or absorbed depending on wavelength, tissue architecture, nuclear size, stromal organization, and the distribution of chromophores. Earlier work in oral cancer showed that ESS could detect cervical intranodal metastasis in bivalved lymph nodes using intensity differences at selected wavelengths, although specificity was limited when a simple threshold was used [14]. Broader optical biopsy literature has also shown that elastic scattering can provide a strong optical signal in several epithelial and nodal settings, supporting the concept that tissue microstructure can be converted into clinically useful diagnostic information [15,16].

This study builds on that foundation by treating nodal assessment as a node-level diagnostic problem rather than a single-contact spectral problem. We hypothesized that metastatic lymph nodes would show reproducible amplitude-related and shape-related spectral differences and that combining full-spectrum information with selected optical features might improve discrimination compared with a simple single-wavelength rule. Because modern diagnostic algorithms require more than high apparent accuracy, we also designed the analysis to reflect principles of transparent prediction model reporting, diagnostic accuracy reporting, bias assessment, and early clinical evaluation of decision support systems [17-20]. The aim was to evaluate whether ESS could provide accurate ex vivo node-level discrimination of metastatic cervical nodal disease in oral squamous cell carcinoma and to define a clinically interpretable workflow for future validation.

XSL•FO
RenderX

Methods

Study Design, Setting, and Timeline

This observational ex vivo diagnostic modeling study was conducted at University College London Hospitals. The optical-imaging study received a favorable ethical opinion from The Joint UCL/UCLH Committees on the Ethics of Human Research (Committee A), confirmed in a letter dated July 31, 2006 (REC [Research Ethics Committee] reference 06/Q0505/39). Patients underwent primary tumor resection and neck dissection between September 2016 and March 2021, with fresh lymph-node preparation, ESS acquisition, and routine histopathological assessment undertaken during the same surgical episode. The archived spectra and pathology records were curated between January and December 2025, and the modeling analyses were completed between January and May 2026.

Participants and Specimens

Potentially eligible patients were assessed during the recruitment period. Adults with histologically confirmed oral squamous cell carcinoma were eligible when neck dissection formed part of their planned surgery, fresh cervical lymph nodes were available before fixation, and written informed consent had been obtained. Patients were excluded when consent was not obtained, suitable fresh lymph nodes were unavailable before fixation, or the ESS system was unavailable at the time of surgery. The patient was the enrollment and validation unit; lymph nodes were clustered within patients, and repeated spectra were clustered within lymph nodes. Within each enrolled neck dissection, ESS was performed on fresh lymph nodes that could be isolated and made available before fixation without delaying or compromising routine pathological processing. The ESS-assessed nodes therefore represented a pragmatic availability-based sample and did not necessarily constitute the complete pathological nodal yield from each neck dissection. Selection for ESS assessment occurred before the final histopathological diagnosis was known.

Ethical Considerations

The application for the optical-imaging study, “The Role of Optical Imaging (Elastic Scattering Spectroscopy, Raman Spectroscopy and Second-harmonic Imaging Microscopy) as a Diagnostic Tool in Head and Neck Cancer,” was submitted to The Joint UCL/UCLH Committees on the Ethics of Human Research (Committee A) on April 26, 2006. A favorable ethical opinion was confirmed in a letter dated July 31, 2006 (REC reference 06/Q0505/39), following consideration of further information by the Chair on July 28, 2006. The approved documents included protocol version 1.4, dated March 13, 2006; participant information sheet version 2, dated June 12, 2006; and participant consent form version 1.

The study protocol described comparison of optical imaging techniques with routine histopathology using clinically obtained oral or head and neck tissue and eligible archived material. The protocol described in vivo assessment before routine biopsy and ex vivo assessment of clinically removed oral or head and neck tissue, without additional research-only biopsy or surgery or interference with routine histopathological processing. The study continuation was subject to annual local organizational governance checks. The protocol assigned responsibility for study conduct, protocol compliance, data custody, and reporting to the Chief Investigator, who is a coauthor of this report.

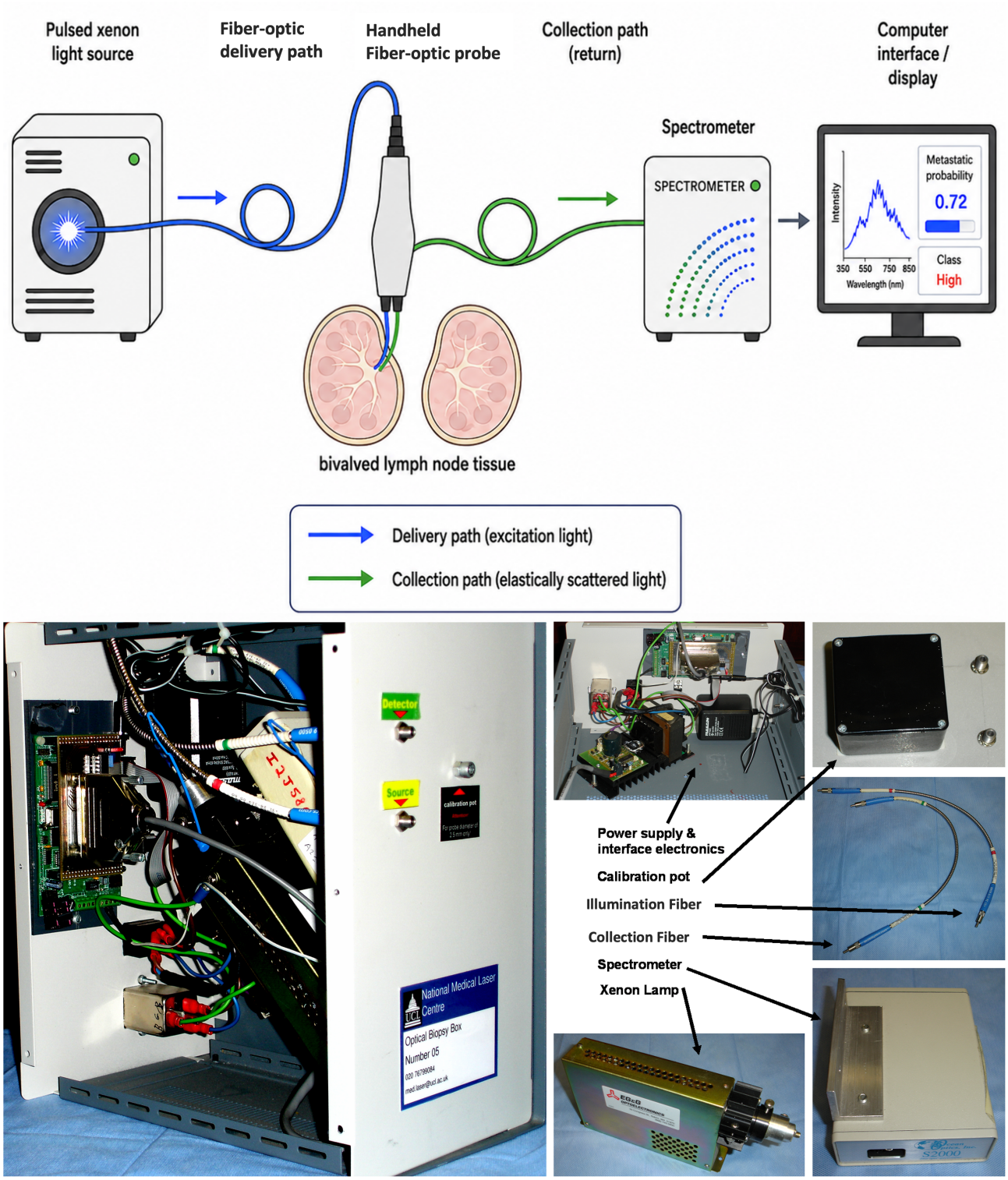
This report concerns fresh cervical lymph-node ESS measurements acquired between September 2016 and March 2021. The archived spectra and pathology records were curated between January and December 2025, and modeling was undertaken between January and May 2026. All participants provided written informed consent for study participation. The analytical dataset was deidentified before modeling, direct identifiers were removed, and access was restricted to the study team. Participants received no payment or other compensation.

Procedure and ESS System

After surgical removal, the neck dissection specimen was orientated with pathological guidance to preserve anatomical context and nodal identity. Individual lymph nodes were isolated and bivalved to expose the internal cut surface. Two trained operators performed ESS acquisition according to a written protocol, with each lymph node assessed by 1 operator. The freshly cut surface was gently blotted, visibly blood-contaminated areas were avoided, and the probe was held approximately perpendicular to the surface in gentle contact without visible tissue compression. In total, 3 spectra were acquired from spatially separated points across the exposed cut surface of each bivalved lymph node. Operator, acquisition batch, and the times of excision, bivalving, and first ESS acquisition were recorded. The median interval from excision to the first acquisition was 18 (IQR 13 - 26) minutes. The overall diagnostic pathway is illustrated in Figure 1, and the probe-tissue interrogation geometry is illustrated in Figure 2.

XSL•FO
RenderX

Figure 1. Elastic scattering spectroscopy (ESS) diagnostic system and hardware components. Upper panel: schematic representation of the ESS system used for ex vivo interrogation of bivalved lymph node tissue. Excitation light from a pulsed xenon source is delivered to the tissue through the illumination fiber, shown in blue. The handheld fiber-optic probe is placed in gentle contact with the cut surface of one half of the bivalved lymph node. Elastically scattered light is collected through the adjacent collection fiber and transmitted to the spectrometer, shown in green. Spectral data are then transferred to the computer interface for display and downstream classification. The computer readout is a conceptual illustration of how spectral information and a future model-derived probability could be displayed; it does not represent a prospectively deployed clinical classifier used during specimen acquisition in this study. Lower panel: representative photographs of the ESS hardware, including the optical biopsy box, detector and source ports, calibration port, power supply and interface electronics, illumination and collection fibers, spectrometer, and xenon lamp. Blue arrows indicate the excitation/delivery pathway; green arrows indicate the collected elastically scattered return signal; the gray arrow indicates transfer of spectral information from the spectrometer to the computer display. The system architecture is consistent with the described ESS platform comprising a pulsed xenon arc lamp, fiber-optic probe, spectrometer, and computer-controlled acquisition interface.



The ESS system comprised a pulsed xenon arc lamp light source, a PC-compatible spectrometer based on a modified Ocean Optics platform with linear charge-coupled device array detection, a fiber-optic probe, and a laptop computer for system control and

XSL•FO
RenderX

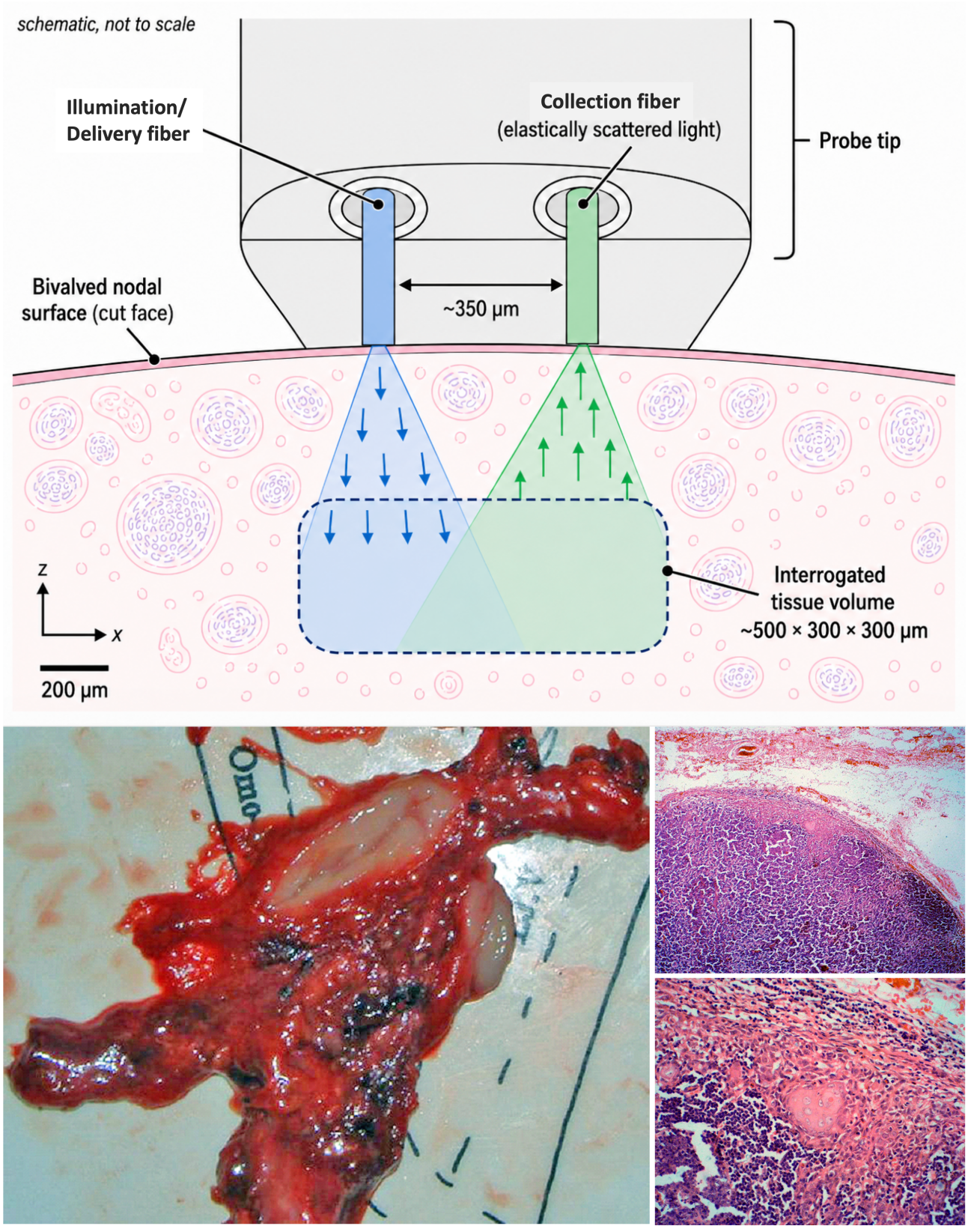
data display. Light was delivered through a 200 micrometer core illumination fiber and scattered light was collected through a neighboring fiber connected to the spectrometer. The probe had a fixed center-to-center separation of approximately 350 micrometers between the delivery and collection fibers. This geometry interrogated a superficial tissue volume of approximately 500 micrometers in length, 300 micrometers in width, and 300 micrometers in depth. The probe was placed in gentle optical contact with the nodal cut surface to minimize reflected light from the surface and to ensure that collected photons had undergone scattering within the tissue.

Before tissue measurements, the system was calibrated using a reflectance standard to normalize the spectral response and reduce variation related to probe transmission, lamp output, and thermal drift. Each acquisition was computer-controlled and included a background measurement followed immediately by pulsed spectral capture with automatic background subtraction. The wavelength range available from the system was broader than the analytical range, but all spectra were subsequently restricted to a common 350‐850 nm window to ensure consistency across measurements. Optical measurement sites were matched to the corresponding lymph node, and the final nodal histopathological diagnosis was used as the class label.

XSL•FO
RenderX

Figure 2. Elastic scattering spectroscopy (ESS) fiber-optic probe geometry, bivalved lymph node sampling, and histopathological correlation. Upper panel: schematic representation of the elastic scattering spectroscopy fiber-optic probe in gentle contact with the exposed cut surface of a bivalved lymph node. The illumination or delivery fiber transmits excitation light into the superficial nodal tissue, shown by the blue cone and downward blue arrows. The adjacent collection fiber captures elastically scattered photons returning from the interrogated tissue volume, shown by the green collection field and upward green arrows. The delivery and collection fibers are separated by approximately 350 µm at the probe tip. With this probe geometry, the sampled optical field interrogates a superficial tissue volume of approximately 500×300×300 µm. The schematic is not drawn to scale and is intended to illustrate probe-tissue geometry rather than exact photon trajectories. Lower left: gross appearance of a surgically excised lymph node specimen after bivalving, exposing the internal nodal surface used for ex vivo optical interrogation. Lower right: representative hematoxylin and eosin-stained sections used for pathological correlation. The upper histological image shows nodal tissue architecture, while the lower histological image demonstrates metastatic keratinizing squamous cell carcinoma within nodal tissue, with malignant squamous nests, keratinization, and associated stromal or inflammatory response. Histopathology remained the reference standard against which the ESS spectral measurements were compared. The histological panels are representative images from the present study material and were not used as quantitative inputs to the ESS classification models.

XSL•FO
RenderX



Histopathological Reference Standard

After ESS acquisition, each bivalved lymph node was fixed in 10% neutral-buffered formalin for 24 - 48 hours, embedded in paraffin, and represented by 1 block from each half. One 4-µm hematoxylin and eosin section was examined from each block. Two deeper levels at 200-µm intervals and pancytokeratin AE1/AE3 immunohistochemistry were used when the initial sections were equivocal or a small metastatic deposit was suspected. The final diagnosis was assigned at lymph-node level.

For this analysis, metastatic deposits were categorized as macrometastases when greater than 2 mm, micrometastases when greater than 0.2 mm and no greater than 2 mm, and

XSL•FO
RenderX

isolated tumor cells when no greater than 0.2 mm. Nodes containing macrometastases or micrometastases were classified as pathology positive. No isolated-tumor-cell-only nodes were identified in the cohort. Pathological T and N categories were retained as documented in the original pathology reports. All tumors were retrospectively restaged according to the UICC TNM eighth edition using the recorded depth-of-invasion and extranodal-extension data. Extranodal extension was recorded when metastatic tumor extended beyond the lymph-node capsule.

Pathologists were blinded to the ESS spectra, probability estimates, and model outputs. Routine histopathological assessment was undertaken by consultant pathologists experienced in head and neck pathology using the contemporaneous clinical reporting pathway. Where a requested clinicopathological characteristic was absent from the archived gross or histopathological record, it was recorded as unavailable rather than inferred.

The study cohort comprised 45 patients and 280 ESS-assessed lymph nodes. Of these nodes, 96 were pathology positive and 184 were pathology negative. Three spectra were acquired per node, giving 840 spectra before quality control. Nodes were sampled from both upper and lower neck levels, with 168 nodes from levels I-III and 112 from levels IV-V. The median node size was 14 mm, and acquisition was performed across 9 batches.

Spectra were imported as wavelength-intensity arrays and reviewed using sequential quality-control rules before modeling. Spectra were excluded if they contained zero or failed channels, excessive noise at the spectral edges, clipping or saturation, duplicate entries, or incomplete records. These prespecified quality-control rules were applied before model fitting and without access to the histopathological class label. Of 840 raw spectra, 80 were excluded, leaving 760 retained spectra. At node level, 274 of 280 ESS-assessed nodes retained at least 2 usable spectra and 212 retained all 3 spectra. In each of the remaining 6 pathology-negative nodes, none of the 3 acquired spectra passed the requirement for an analyzable node-level ESS result. These 6 nodes were therefore reported as node-level index-test no-results and were not included in the 274-node modeling and performance denominator. The node-level no-result rate was 2.1% (6/280).

Two complementary preprocessing approaches were used. For feature-based analyses intended to preserve optical amplitude, spectra were lightly smoothed and scaled after restriction to the 350 - 850 nm window. For full-spectrum analyses intended to emphasize spectral shape, standard normal variate (SNV) normalization was applied after smoothing. Principal component analysis was used where dimensionality reduction was required, particularly for support vector machine modeling. These preprocessing choices were applied consistently so that amplitude-dominant and shape-dominant analytical strategies could be compared.

A total of 8 candidate optical features were specified before the final comparative modeling analysis on the basis of clinical interpretability and previous ESS work. I360, I450, I630, and I700 were defined as the spectral intensity at the wavelength channel nearest to 360, 450, 630, and 700 nm, respectively. The I630/I360 ratio was calculated from these 2 intensity measurements. Total spectral area was calculated by trapezoidal integration from 350 to 850 nm. Blue trough depth was defined as I450 minus the minimum spectral intensity between 390 and 420 nm. Red shoulder height was defined as the maximum spectral intensity between 650 and 700 nm minus I630. Retained spectra were first summarized by their median within each lymph node so that every node contributed 1 observation per feature. Effect sizes were expressed as rank-biserial correlations. 95% CIs were estimated by resampling patients as clusters while retaining all lymph nodes from each selected patient. Moreover, 2-sided *P* values were obtained by permuting node-level pathology labels within patient while preserving each patient's number of pathology-positive and pathology-negative nodes. Each permutation-based *P* value was estimated using 10,000 patient-preserving permutations.

Model Development

Model development compared a baseline I630 rule, in which I630 was defined as the median intensity at the wavelength channel nearest 630 nm across the retained spectra from each lymph node and nodes were classified using a decision threshold selected exclusively from patient-grouped inner-validation training data by maximizing the Youden index, feature-based logistic regression, full-spectrum elastic-net logistic regression, a linear support vector machine, and a hybrid elastic-net model. The amplitude branch comprised the 8 prespecified optical features extracted from Savitzky-Golay–smoothed raw-amplitude spectra, with robust scaling fitted within the inner training folds. The shape branch comprised principal-component scores derived from standard-normal-variate-normalized spectra, with principal component analysis fitted within the inner training folds. Candidate component counts of 10, 20, 30, and 40 were evaluated, with 20 components selected most frequently across the outer fits. The hybrid elastic-net model combined the selected principal-component scores with the 8 scaled optical features. Regularization strength, l1 ratio, and class weighting were selected within the inner folds; C=0.3 and an l1 ratio of 0.5 were the most frequently selected values. Model-family comparisons were exploratory. Each candidate model was tuned independently within the patient-grouped inner validation process, and the hybrid model was carried forward for detailed calibration, error, and robustness analyses because it had the highest pooled patient-grouped out-of-fold area under the receiver operating characteristic curve (AUROC) point estimate. No formal pairwise test of superiority between model families was prespecified.

Internal Validation and Leakage Control

Primary internal validation used 10 repetitions of nested 5-fold cross-validation implemented using stratified group k-fold splitting, with patient ID as the grouping variable. Within each repetition, pathology class was stratified as closely as possible across folds while all spectra and lymph nodes from an individual patient remained within a single fold. Fold assignment was shuffled before each repetition using random seeds 202601 to 202610, generating 50 outer training-test splits in total. Every

XSL•FO
RenderX

spectrum and lymph node from a given patient remained in the same outer fold, so lymph nodes from the same patient could not occur simultaneously in the corresponding outer training and test sets. Within each outer training set, model development and hyperparameter tuning used 4-fold stratified group cross-validation, again grouped by patient and with shuffling enabled. The repetition-specific random seed was retained for the corresponding inner-validation procedure, and no patient represented in an outer test fold was available to any stage of inner model development. Smoothing, scaling, SNV transformation, principal component analysis, component selection, regularization parameters, class weighting, probability calibration where required, and decision threshold selection were fitted or selected using training data only. Outer test patients were used only once the corresponding analytical pipeline had been fixed, solely to generate held-out predictions.

Model Selection and Decision Threshold

Within each candidate model family, mean inner-fold AUROC was used to select hyperparameters and model configuration. Within each outer training set, the node-level decision threshold was selected from patient-grouped inner-fold out-of-fold predictions using the Youden index. Thus, threshold optimization itself was restricted to training data and did not use the corresponding outer test outcomes. Each node received 1 held-out probability in each of the 10 outer repetitions, and these probabilities were averaged to obtain a single patient-grouped out-of-fold probability for each node. The 50 training-derived thresholds had a median of 0.38 (IQR 0.35‐0.41). For the pooled descriptive confusion matrix and associated threshold-dependent operating characteristics, this median training-derived threshold was applied once to the averaged held-out node probabilities. No pooled outcome or outer-test outcome was used to select or adjust the threshold. Because this single summary threshold was applied after repeated internal cross-validation rather than prospectively fixed before validation, sensitivity, specificity, accuracy, positive predictive value (PPV), and NPV were interpreted as secondary descriptive operating characteristics. AUROC and calibration calculated from the pooled patient-grouped out-of-fold probabilities were the principal performance measures in this exploratory internal-validation analysis. Because these predictions also informed the choice of the hybrid model family, its reported performance estimates may be affected by selection optimism and do not constitute independent validation of a prospectively fixed final model.

Calibration and Statistical Uncertainty

Calibration was assessed using each node's averaged patient-grouped out-of-fold probability and was summarized using the Brier score, calibration-in-the-large, calibration slope, and a grouped calibration plot. For the plot, the observed proportion of pathology-positive nodes was compared with the mean predicted probability within each probability group; groups were not constrained to contain equal numbers of nodes. 95% CIs for node-level sensitivity, specificity, accuracy, PPV, NPV, AUROC, and calibration measures were estimated using 2000 patient-clustered bootstrap resamples, with all lymph nodes belonging to each resampled patient retained together. The reported bootstrap intervals were calculated from the completed patient-grouped out-of-fold prediction dataset without refitting the complete model-development procedure within each bootstrap replicate; they therefore quantify uncertainty in the held-out predictions conditional on the completed modeling process. Patient-level CIs for sensitivity, specificity, accuracy, PPV, and NPV were calculated using the Wilson method. The patient-level AUROC CI was estimated by bootstrap resampling of patients.

Node-Level and Patient-Level Classification

Because the primary diagnostic unit was the lymph node, spectrum-level probabilities were averaged to generate 1 probability for each node. Mean probability across retained spectra was the primary node-level aggregation method. Majority vote, median probability, maximum probability, and classification using a single representative spectrum were evaluated as secondary aggregation strategies. For the single-spectrum analysis, the first quality-control-passing acquisition from each lymph node according to the recorded acquisition order was selected, without reference to histopathological status or model output.

A secondary patient-level analysis was undertaken using the cross-validation–derived summary node-level threshold. A patient was reference-standard positive when at least 1 ESS-assessed lymph node was metastatic on histopathology and model positive when at least 1 assessed node exceeded the summary threshold. The maximum averaged out-of-fold node probability within each patient was used for patient-level receiver operating characteristic curve analysis. The patient-level analysis therefore represented only the ESS-assessed node set and was not interpreted as complete pathological staging of the neck.

Performance was summarized using sensitivity, specificity, accuracy, PPV, NPV, and AUROC. Robustness analyses included leave-one-patient-out validation, leave-one-batch-out validation, exclusion of the lowest-quality 10% of retained spectra ranked by signal-to-noise ratio, use of the first 2 quality-control–passing spectra from each node according to recorded acquisition order, restriction to nodes retaining all 3 spectra, area-normalized spectra, and shape-focused preprocessing.

Analysis Plan and Reproducibility

The 3-acquisition sampling protocol and 8 candidate optical features were specified before the final modeling analysis. The comparison of machine learning algorithms, model-family selection, and associated hyperparameter grids was exploratory and was not prospectively registered. Accordingly, comparisons between candidate model families were interpreted descriptively, and no claim of formal superiority was made. Analyses were conducted using Python (version 3.12.8; Python Software Foundation), *NumPy* 2.2.1, *pandas* 2.2.3, *SciPy* 1.15.0, and *scikit-learn* 1.6.0. Random seeds 202601 to 202610 were used for the 10 outer repetitions. No formal sample-size calculation was undertaken; the analysis used the complete available cohort meeting the eligibility requirements and having linked ESS and histopathological data.

XSL•FO
RenderX

Reporting followed TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis)+AI, STARD (Standards for Reporting of Diagnostic Accuracy Studies)-AI, and STARD 2015 [17,18,21]. Completed reporting checklists are provided in Checklists 1-3 Table S1 and Figure S1 in Multimedia Appendix 1 provide the analytical workflow, software versions, preprocessing sequence, candidate component counts, hyperparameter grid, threshold-selection method, and leakage-control procedures.

Results

Overview

In total, 52 patients were assessed for eligibility. Of the total, 7 were excluded before enrollment; 3 because consent was not obtained, 2 because no suitable fresh lymph nodes were available before fixation, and 2 because the ESS system was unavailable at the time of surgery. The final cohort comprised 45 patients, with a median age of 61 (IQR 53 - 69) years; 29 were male and 16 were female. Moreover, 25 patients had at least 1 pathology-positive ESS-assessed node, and 20 had no positive assessed nodes. Primary tumor site, pathological T category, clinical and pathological N category, depth of invasion, histological grade, extranodal extension, and positive-node burden are summarized in Table 1.

XSL•FO
RenderX

Table . Participant and primary tumor characteristics. Percentages may not total 100 because of rounding.

Characteristic	Overall cohort
Patients, n	45
Age (y), median (IQR)	61 (53 - 69)
Age range (y)	34 - 82
Sex, n (%)	
Male	29 (64.4)
Female	16 (35.6)
Primary tumor site, n (%)	
Oral tongue	17 (37.8)
Floor of mouth	9 (20)
Mandibular alveolus	7 (15.6)
Buccal mucosa	5 (11.1)
Retromolar trigone	4 (8.9)
Maxillary alveolus or hard palate	3 (6.7)
Pathological T category, n (%)	
pT1	8 (17.8)
pT2	17 (37.8)
pT3	8 (17.8)
pT4a	12 (26.7)
Clinical N category, n (%)	
cN0	23 (51.1)
cN1	9 (20)
cN2	13 (28.9)
Pathological N category, n (%)	
pN0	20 (44.4)
pN1	8 (17.8)
pN2	17 (37.8)
Depth of invasion (mm), median (IQR)	8 (5-13)
Histological grade, n (%)	
Well differentiated	8 (17.8)
Moderately differentiated	29 (64.4)
Poorly differentiated	8 (17.8)
Patients with ≥1 pathology-positive ESS-assessed[a] node, n (%)	25 (55.6)
Extranodal extension among node-positive patients, n/N (%)	9/25 (36)
Positive nodes per node-positive patient, median (IQR); range	3 (2-5); 1 - 9

[a]ESS: elastic scattering spectroscopy.

The patients contributed 280 ESS-assessed cervical lymph nodes, of which 96 were pathology positive and 184 were pathology negative. The median node size was 14 (IQR 9 - 21) mm. Levels I-III accounted for 168 nodes and levels IV-V for 112 nodes. The median number of ESS-assessed nodes per patient was 6 (IQR 4 - 8). The study timeline, participant flow, nodal sampling, and acquisition characteristics are summarized in Table 2.

XSL•FO
RenderX

Table . Study timeline, participant flow, nodal sampling, and acquisition characteristics. The REC[a] dates refer to the optical-imaging study. The study scope, provisions for annual local organizational governance checks, and responsibilities of the Chief Investigator are described in the Ethical Considerations section.

Variable	Value
Study timeline	
Optical-imaging study: REC application submitted	April 26, 2006
Optical-imaging study: favorable ethical opinion confirmed, REC 06/Q0505/39	July 31, 2006
Surgery, tissue collection, and ESS[b] acquisition	September 2016-March 2021
Data curation	January-December 2025
Exploratory model development and internal validation	January-May 2026
Participant flow	
Patients assessed for eligibility, n	52
Patients excluded before enrollment, n (%)	7 (13.5)
Written consent not obtained, n	3
No suitable fresh lymph nodes available before fixation, n	2
ESS system unavailable at the time of surgery, n	2
Patients included, n	45
Nodal sampling	
Lymph nodes assessed by ESS, n	280
Pathology-positive nodes, n (%)	96 (34.3)
Pathology-negative nodes, n (%)	184 (65.7)
Nodes from neck levels I-III, n (%)	168 (60)
Nodes from neck levels IV-V, n (%)	112 (40)
Node size (mm), median (IQR)	14 (9 - 21)
ESS-assessed nodes per patient, median (IQR)	6 (4-8)
Acquisition characteristics	
ESS spectra acquired, n	840
Spectra acquired per node, n	3
Acquisition batches, n	9
Operator A	31 patients; 193 nodes
Operator B	14 patients; 87 nodes
Time from excision to first ESS acquisition (min), median (IQR)	18 (13 - 26)
Time from excision to first ESS acquisition (min), range	7 - 48

[a]REC:

[b]ESS: elastic scattering spectroscopy.

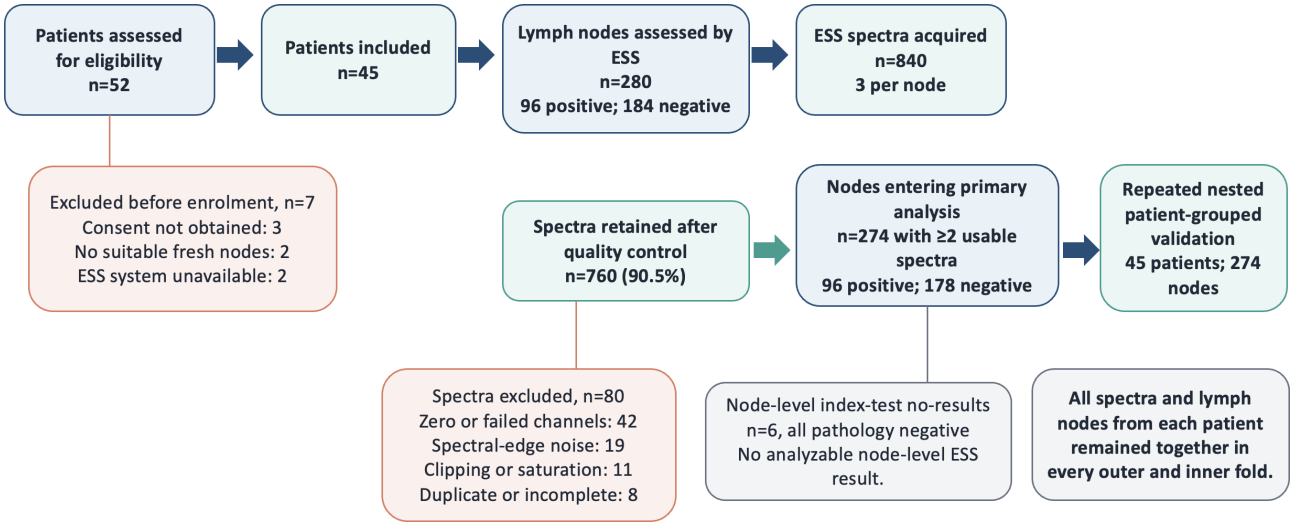
Acquisition and Quality Control

A total of 3 spectra were acquired from each ESS-assessed lymph node, giving 840 acquisitions. Quality control excluded 80 spectra—42 because of zero or failed channels, 19 because of excessive spectral-edge noise, 11 because of clipping or saturation, and 8 because of duplicate or incomplete records. The final spectrum set comprised 760 retained spectra, equivalent to 90.5% of acquisitions.

At node level, 274 lymph nodes retained at least 2 usable spectra and entered the analytical cohort. These comprised 96 pathology-positive and 178 pathology-negative nodes. In each of 6 additional pathology-negative nodes, none of the 3 acquired spectra met the requirement for an analyzable node-level ESS result. These 6 nodes were reported as node-level index-test no-results and were not included in the 274-node modeling or performance denominator. Moreover, 212 nodes retained all 3 spectra. Participant, specimen, acquisition, and analysis flow is illustrated in Figure 3, and quality-control outcomes are reported in Table 3.

XSL•FO
RenderX

Figure 3. Participant, specimen, acquisition, and analysis flow. Of 52 patients assessed for eligibility, 7 were excluded before enrollment: 3 because consent was not obtained, 2 because no suitable fresh lymph nodes were available before fixation, and 2 because the elastic scattering spectroscopy (ESS) system was unavailable at the time of surgery. In total, 45 patients were included and contributed 280 ESS-assessed lymph nodes, of which 96 were pathology positive and 184 were pathology negative. Moreover, 3 spectra were acquired from each node, giving 840 acquisitions. Quality control excluded 80 spectra, leaving 760 retained spectra. At node level, 274 nodes retained at least 2 usable spectra and comprised the analytical cohort; these included 96 pathology-positive and 178 pathology-negative nodes. Furthermore, 6 pathology-negative nodes produced no analyzable node-level ESS result and were reported as index-test no-results. All spectra and lymph nodes from each patient remained together during patient-grouped internal validation.



The 6 node-level index-test no-results were pathology negative and did not contribute to the 274-node modeling or performance denominator. Percentages for spectrum-level quality-control exclusions use 840 acquired spectra as the denominator; node percentages use 280 ESS-assessed nodes unless otherwise stated. Percentages shown for pathology-positive and pathology-negative analyzable nodes use 274 nodes as the denominator.

Table . Spectral quality control and node-level index-test no-results. The 6 node-level index-test no-results were pathology negative and did not contribute to the 274-node modeling or performance denominator. Percentages for spectrum-level quality-control exclusions use 840 acquired spectra as the denominator; node percentages use 280 elastic scattering spectroscopy–assessed nodes unless otherwise stated. Percentages shown for pathology-positive and pathology-negative analyzable nodes use 274 nodes as the denominator.

Quality-control step	Values, n (%)
Raw spectra acquired	840 (100)
Spectra with zero or failed channels	42 (5)
Spectra excluded for excessive spectral-edge noise	19 (2.3)
Spectra excluded for clipping or saturation	11 (1.3)
Spectra excluded as duplicate or incomplete records	8 (1)
Total excluded spectra	80 (9.5)
Spectra retained	760 (90.5)
Nodes retaining ≥2 usable spectra	274 (97.9)
Pathology positive	96 (35 of analyzed nodes)
Pathology negative	178 (65 of analyzed nodes)
Nodes retaining all 3 spectra	212 (75.7)
Node-level index-test no-results	6 (2.1)

Spectral Feature Differences

Pathology-positive lymph nodes showed higher intensities at 360, 450, and 630 nm and greater total spectral area. In contrast, pathology-negative nodes showed higher intensity at 700 nm, a higher I630/I360 ratio, a deeper local blue trough, and a larger red shoulder. The differences were therefore bidirectional rather than representing a uniform increase in spectral amplitude and supported the inclusion of both amplitude-related and shape-related information. The node-level median ESS profiles are illustrated in Figure 4, and numerical medians and IQRs for the prespecified node-level spectral features are reported in Table 4.

XSL•FO
RenderX

Figure 4. Median and representative elastic scattering spectroscopy profiles of pathology-negative and pathology-positive lymph nodes. The left panel shows node-level median elastic scattering spectroscopy (ESS) profiles across the 350-850 nm analytical range. Numerical medians and IQRs for the prespecified node-level spectral features are reported in Table 4. The pathology-positive profile showed higher short-wavelength intensity and a smoother, more gradually declining pattern, whereas the pathology-negative profile showed a deeper local blue trough and a later rise in the red and near-red region. The upper-right and lower-right panels show representative raw spectra generated by the ESS acquisition software for pathology-negative and pathology-positive nodes, respectively, illustrating within-class acquisition variability. The raw-spectrum panels retain the native acquisition-software wavelength and intensity scales and should not be used for direct between-class amplitude comparison. AU: arbitrary unit.

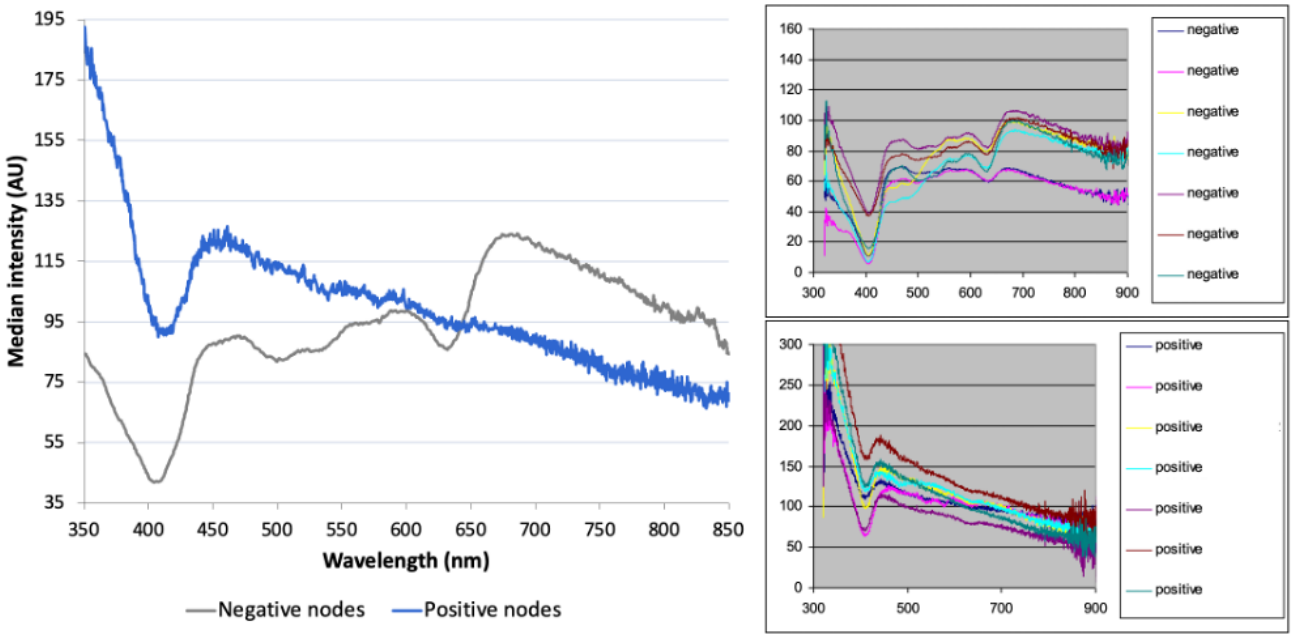


Table . Node-level spectral feature differences by pathological status. Spectra were first aggregated within each lymph node. Positive rank-biserial correlations indicate higher values in pathology-positive nodes, whereas negative correlations indicate higher values in pathology-negative nodes. CIs were estimated by resampling patients while retaining all lymph nodes from each selected patient. Blue trough depth was calculated as I450 minus the minimum intensity from 390 to 420 nm. Red shoulder height was calculated as the maximum intensity from 650 to 700 nm minus I630. The reported *P* values are exploratory, unadjusted for multiplicity, and should be interpreted together with the corresponding effect estimates and CIs rather than as confirmatory hypothesis tests.

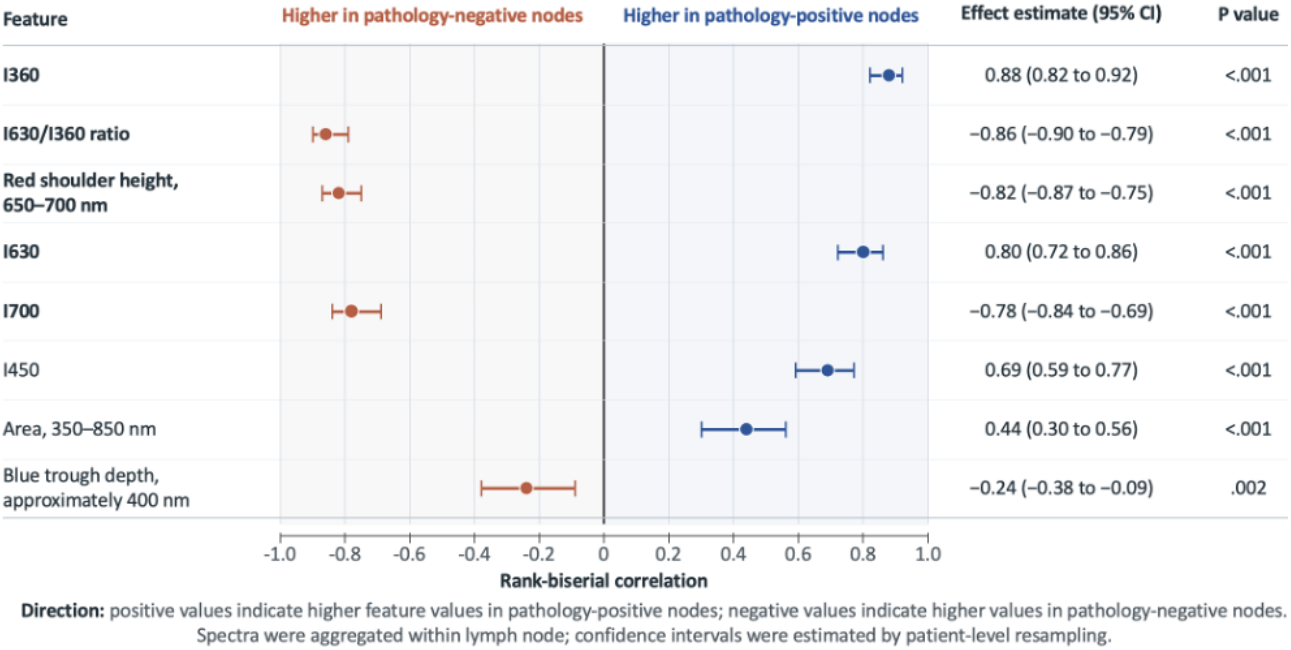
Feature	Pathology-negative nodes, median (IQR)	Pathology-positive nodes, median (IQR)	Rank-biserial correlation (95% CI)	*P* value
I360, AU[a]	80 (69 - 92)	168 (149 - 186)	0.88 (0.82 to 0.92)	<.001
I450, AU	85 (76 - 95)	121 (109 - 134)	0.69 (0.59 to 0.77)	<.001
I630, AU	88 (84 - 92)	96 (93 - 103)	0.80 (0.72 to 0.86)	<.001
I700, AU	120 (107 - 132)	88 (81 - 96)	–0.78 (–0.84 to –0.69)	<.001
I630/I360 ratio	1.10 (0.99 - 1.22)	0.58 (0.53 - 0.66)	–0.86 (–0.90 to –0.79)	<.001
Area, 350 - 850 nm, AU·nm	45,800 (41,600-50,700)	50,600 (46,900-55,400)	0.44 (0.30 to 0.56)	<.001
Blue trough depth, approximately 400 nm, AU	43.1 (35.8 - 50.5)	31.4 (25.6 - 38.2)	–0.24 (–0.38 to –0.09)	.002
Red shoulder height, 650 - 700 nm, AU	31.6 (24.8 - 38.9)	6.7 (3.9 - 10.2)	–0.82 (–0.87 to –0.75)	<.001

[a]AU: arbitrary unit.

Pathology-positive nodes had higher I360, I450, and I630 intensities and greater total spectral area, whereas pathology-negative nodes had higher I700 intensity, a higher I630/I360 ratio, a deeper blue trough, and a larger red shoulder. The largest effects were observed for I360, with a rank-biserial correlation of 0.88 (95% CI 0.82 - 0.92); the I630/I360 ratio, with a correlation of –0.86 (95% CI –0.90 to –0.79); red shoulder height, with a correlation of –0.82 (95% CI –0.87 to –0.75); I630, with a correlation of 0.80 (95% CI 0.72 - 0.86); and I700, with a correlation of –0.78 (95% CI –0.84 to –0.69). Moreover, 7 features had *P*<.001, while blue trough depth remained significant at *P*=.002. Full medians, IQRs, effect sizes, CIs, and *P* values are reported in Table 4 and Figure 5.

XSL•FO
RenderX

Figure 5. Effect sizes for prespecified node-level spectral features. Points show rank-biserial correlations, and horizontal lines show patient-clustered 95% CIs. Positive correlations indicate higher feature values in pathology-positive nodes, whereas negative correlations indicate higher values in pathology-negative nodes. Features are ordered by the absolute magnitude of the effect estimate. Spectra were first aggregated within each lymph node, and CIs were estimated by resampling patients while retaining all lymph nodes from each selected patient. The corresponding *P* values were derived from patient-preserving permutation analyses and are exploratory and unadjusted for multiplicity.



Patient-Grouped Internal Validation

Repeated nested patient-grouped cross-validation yielded an internally validated, exploratory AUROC estimate of 0.966 (95% CI 0.929 - 0.988) for the selected hybrid model. Because the hybrid model family was selected using the same pooled out-of-fold predictions, this result is not an independent validation estimate for a prospectively fixed final model. At the cross-validation–derived summary threshold, the pooled descriptive confusion matrix comprised 90 true-positive, 6 false-negative, 170 true-negative, and 8 false-positive lymph nodes. The corresponding secondary descriptive operating characteristics were as follows: sensitivity 93.8% (95% CI 86.4% - 97.6%), specificity 95.5% (95% CI 90.1% - 98.3%), accuracy 94.9% (95% CI 90.4% - 97.3%), PPV 91.8% (95% CI 82.9% - 96.3%), and NPV 96.6% (95% CI 92.4% - 98.8%). The 50 training-derived decision thresholds had a median of 0.38 (IQR 0.35 - 0.41); no pooled outcome or outer-test outcome was used to select or modify this threshold. The Brier score was 0.063, calibration-in-the-large was –0.03, and the calibration slope was 0.94. The calibration plot showed close agreement between observed and predicted probabilities across most of the probability range, with modest overprediction at the highest probabilities. Classification and calibration results are presented in Table 5 and Figure 6.

XSL•FO
RenderX

Table . Selected hybrid model discrimination, descriptive operating characteristics, calibration, aggregation, and robustness. Threshold-dependent measures are presented as secondary descriptive operating characteristics. For the primary patient-grouped node-level analysis, the cross-validation–derived summary threshold was applied to each node's averaged held-out probability to obtain the pooled confusion matrix. For the patient-grouped spectrum-level analysis and the analysis restricted to nodes retaining all 3 spectra, sensitivity, specificity, and accuracy were calculated separately within each of the 10 outer repetitions and then averaged; these secondary percentages therefore represent means of repetition-specific metrics rather than proportions derived from a single pooled confusion matrix. AUROC[a] estimates were derived from held-out probabilities within the reported internal-validation analyses. The primary patient-grouped AUROC of 0.966 is an internally validated, exploratory estimate for a model family selected using the same pooled out-of-fold predictions and does not represent independent validation of a prospectively fixed final model. Patient-level PPV[b] was 92.3% (95% CI 75.9%-97.9%) and NPV[c] was 94.7% (95% CI 75.4%-99.1%). Patient-level CIs for sensitivity, specificity, accuracy, PPV, and NPV were calculated using the Wilson method; the patient-level AUROC CI was estimated by bootstrap resampling of patients.

Domain or analysis	Analytical set	Sensitivity,%	Specificity,%	Accuracy,%	AUROC	Additional result
Validation level						
Primary patient-grouped node-level analysis	274 nodes	93.8	95.5	94.9	0.966	90 TP[d]; 6 FN[e]; 170 TN[f]; 8 FP[g]
Secondary node-grouped node-level analysis	274 nodes	96.9	96.6	96.7	0.989	93 TP; 3 FN; 172 TN; 6 FP
Patient-grouped spectrum-level analysis	760 spectra	89.6	92.0	91.2	0.944	Secondary analysis
Secondary patient-level analysis	45 patients	96.0	90.0	93.3	0.956	24 TP; 1 FN; 18 TN; 2 FP
Calibration of the primary hybrid model						
Brier score	274 nodes	—[h]	—	—	—	0.063 (95% CI 0.044 - 0.086)
Calibration-in-the-large	274 nodes	—	—	—	—	−0.03 (95% CI −0.21 to 0.15)
Calibration slope	274 nodes	—	—	—	—	0.94 (95% CI 0.74 - 1.17)
Training-derived threshold distribution	Outer splits	—	—	—	—	Median 0.38 (IQR 0.35 - 0.41) across 50 outer training sets
Node-level aggregation						
Single prespecified spectrum	274 nodes	88.5	89.9	89.4	0.930	—
Majority vote	274 nodes	92.7	92.7	92.7	0.953	—
Mean probability	274 nodes	93.8	95.5	94.9	0.966	Primary method
Median probability	274 nodes	92.7	94.9	94.2	0.963	—
Maximum probability	274 nodes	97.9	87.6	91.2	0.959	Higher false-positive rate
Robustness analyses						
Leave-one-patient-out validation	274 nodes	92.7	94.9	94.2	0.961	—
Leave-one-batch-out validation	274 nodes	91.7	93.8	93.1	0.952	—
Excluding the lowest-quality 10% of spectra	266 nodes	94.7	95.9	95.5	0.969	—
Using 2 spectra per node	274 nodes	91.7	93.8	93.1	0.951	—

XSL•FO
RenderX

Domain or analysis	Analytical set	Sensitivity,%	Specificity,%	Accuracy,%	AUROC	Additional result
Nodes retaining all 3 spectra	212 nodes	94.7	95.6	95.3	0.964	—
Area-normalized spectra	274 nodes	85.4	87.1	86.5	0.895	—
SNV[i] plus first derivative	274 nodes	87.5	89.3	88.7	0.912	—

[a]AUROC: area under the receiver operating characteristic curve.

[b]PPV: positive predictive value.

[c]NPV: negative predictive value.

[d]TP: true positive.

[e]FN: false negative.

[f]TN: true negative.

[g]FP: false positive.

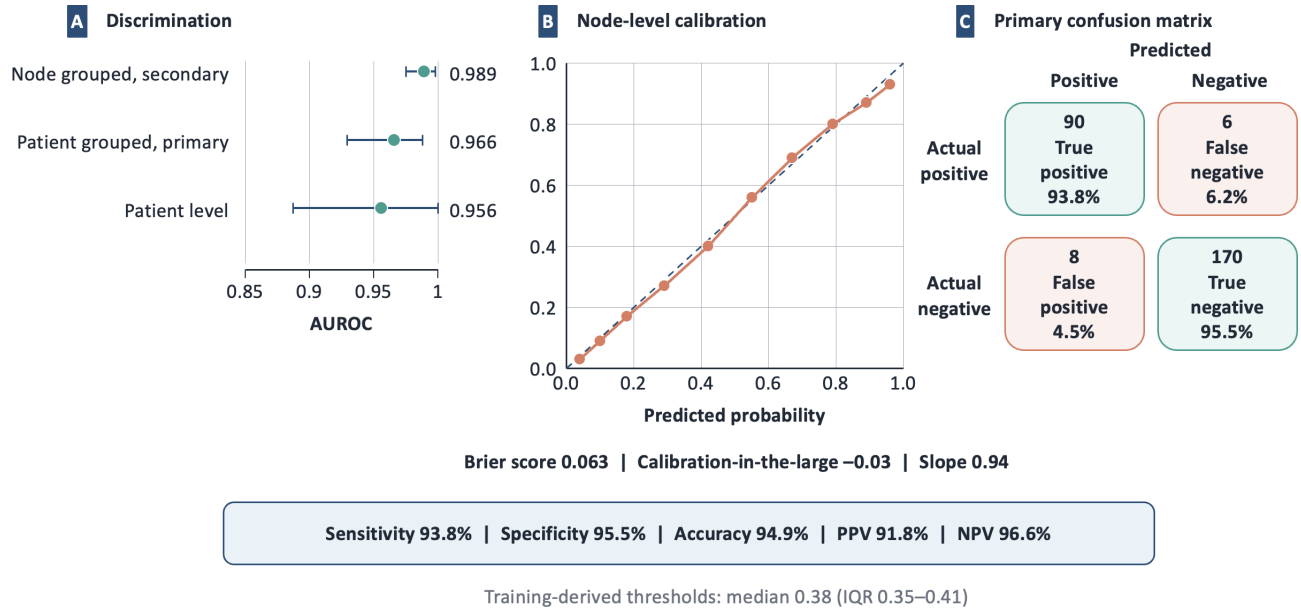
[h]Not applicable.

[i]SNV: standard normal variate.

Among the candidate model families, the hybrid model had the highest pooled patient-grouped out-of-fold AUROC point estimate. Comparative performance is reported in Table 6 and Figure 7. Because model-family comparisons were exploratory and no formal pairwise superiority test was prespecified, these differences should not be interpreted as establishing superiority of one model family over another.

In the secondary node-grouped analysis, the hybrid model produced 93 true-positive, 3 false-negative, 172 true-negative, and 6 false-positive nodes. These counts corresponded to sensitivity of 96.9%, specificity of 96.6%, accuracy of 96.7%, PPV of 93.9%, NPV of 98.3%, and an AUROC of 0.989. The higher point estimates obtained when grouping was restricted to lymph nodes illustrate the potential for optimistic performance estimates when patient-level clustering is not preserved.

Figure 6. Exploratory discrimination, calibration, and descriptive classification performance of the selected hybrid model under internal validation. Panel A shows area under the receiver operating characteristic curve (AUROC) estimates for the primary patient-grouped node-level analysis, the secondary node-grouped analysis, and the patient-level analysis. The primary patient-grouped AUROC of 0.966 is an internally validated, exploratory estimate for the selected model and does not represent independent validation of a prospectively fixed final model. Points show AUROC estimates, and horizontal bars show the corresponding 95% CIs for the primary patient-grouped node-level analysis, the secondary node-grouped analysis, and the patient-level analysis. Panel B compares observed pathology-positive proportions with mean predicted probabilities calculated from averaged patient-grouped out-of-fold node predictions. Each point represents a probability group, and the groups were not constrained to contain equal numbers of nodes. The diagonal line represents perfect calibration. The Brier score was 0.063, calibration-in-the-large was –0.03, and the calibration slope was 0.94. Panel C shows the descriptive confusion matrix obtained by applying the cross-validation-derived summary threshold to averaged patient-grouped held-out probabilities: 90 true-positive, 6 false-negative, 170 true-negative, and 8 false-positive nodes. The 50 training-derived thresholds had a median of 0.38 (IQR 0.35-0.41). No pooled outcome or outer-test outcome was used to select or adjust the summary threshold. NPV: negative predictive value; PPV: positive predictive value.



XSL•FO
RenderX

Table . Patient-grouped internal validation of candidate node-level models. AUROC[a] estimates were derived from pooled patient-grouped out-of-fold predictions and represent exploratory internal-validation performance. The hybrid model family was selected on the basis of these same predictions, its AUROC of 0.966 should therefore not be interpreted as independent validation of a prospectively fixed final model. For each candidate model family, sensitivity, specificity, accuracy, PPV[b], and NPV[c] were calculated using that model family's own summary threshold derived from patient-grouped inner-validation predictions and are presented as secondary descriptive operating characteristics. 95% CIs were estimated using 2000 patient-clustered bootstrap resamples. Model-family comparisons were exploratory and no formal pairwise superiority testing was prespecified.

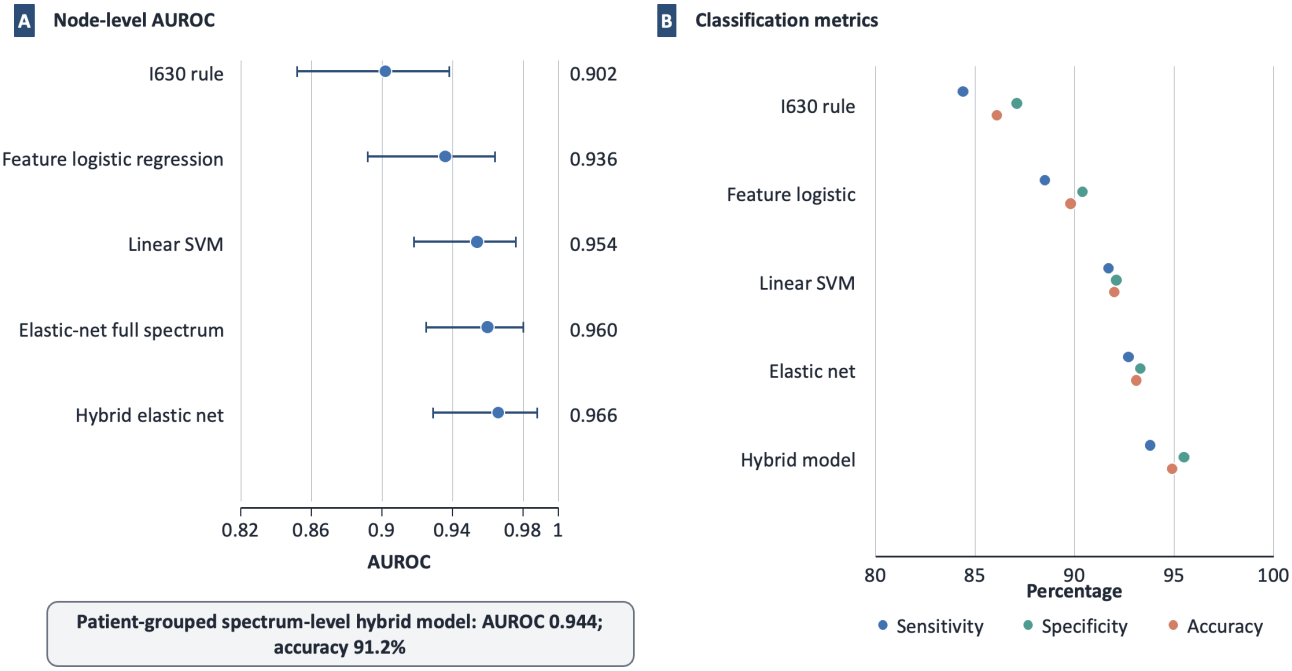
Model	Sensitivity,% (95% CI)	Specificity,% (95% CI)	Accuracy,% (95% CI)	PPV,% (95% CI)	NPV,% (95% CI)	AUROC (95% CI)
I630 rule	84.4 (75.0 - 91.0)	87.1 (79.8 - 92.4)	86.1 (80.4 - 90.2)	77.9 (65.8 - 86.4)	91.2 (84.9 - 95.0)	0.902 (0.852 - 0.938)
Feature logistic regression	88.5 (79.5 - 94.6)	90.4 (83.2 - 94.8)	89.8 (84.5 - 93.6)	83.3 (73.0 - 90.2)	93.6 (88.0 - 97.0)	0.936 (0.892 - 0.964)
Elastic-net full-spectrum model	92.7 (84.4 - 97.1)	93.3 (87.0 - 97.2)	93.1 (88.3 - 96.4)	88.1 (78.5 - 94.8)	96.0 (91.2 - 98.8)	0.960 (0.925 - 0.980)
Linear support vector machine	91.7 (83.2 - 96.5)	92.1 (85.7 - 96.5)	92.0 (86.9 - 95.8)	86.3 (76.2 - 93.5)	95.3 (90.4 - 98.4)	0.954 (0.918 - 0.976)
Hybrid elastic-net model	93.8 (86.4 - 97.6)	95.5 (90.1 - 98.3)	94.9 (90.4 - 97.3)	91.8 (82.9 - 96.3)	96.6 (92.4 - 98.8)	0.966 (0.929 - 0.988)

[a]AUROC: area under the receiver operating characteristic curve.

[b]PPV: positive predictive value.

[c]NPV: negative predictive value.

Figure 7. Exploratory comparison of candidate elastic scattering spectroscopy (ESS) models under patient-grouped internal validation. Panel A shows node-level area under the receiver operating characteristic curve (AUROC) and 95% CIs for the I630 rule, feature logistic regression, linear support vector machine, elastic-net full-spectrum model, and hybrid elastic-net model. Panel B shows the corresponding descriptive sensitivity, specificity, and accuracy estimates at each model family's own cross-validation-derived summary threshold. The hybrid elastic-net model family was selected because it had the highest pooled patient-grouped out-of-fold AUROC point estimate among the candidates. Its AUROC of 0.966 is an internally validated, exploratory estimate based on the same predictions used for model-family selection, rather than independent validation of a prospectively fixed final model. No formal pairwise superiority test was prespecified. The patient-grouped spectrum-level hybrid analysis had an AUROC of 0.944 and a mean accuracy of 91.2% across the 10 outer repetitions. SVM: support vector machine.



Patient-Level Performance

At patient level, 25 patients were reference-standard positive and 20 were negative within the ESS-assessed node set. Using the cross-validation–derived summary node-level threshold, 24 of 25 positive patients and 18 of 20 negative patients were classified correctly, giving 1 false-negative and 2 false-positive patient classifications. Patient-level sensitivity was 96% (95% CI 80.5% - 99.3%), specificity was 90% (95% CI 69.9% - 97.2%), accuracy was 93.3% (95% CI 82.1% - 97.7%), PPV was 92.3% (95% CI 75.9% - 97.9%), and NPV was 94.7% (95% CI 75.4% - 99.1%). The patient-level AUROC, based on the maximum node probability within each patient, was 0.956 (95% CI 0.887 - 1.000).

XSL•FO
RenderX

Patient-grouped spectrum-level analysis produced the same model ranking but lower performance than node-level analysis. The spectrum-level hybrid model had an AUROC of 0.944; mean sensitivity, specificity, and accuracy across the 10 outer repetitions were 89.6%, 92%, and 91.2%, respectively. This reduction supported the use of repeated acquisition and node-level aggregation rather than classification of individual spectra as independent diagnostic observations. Spectrum-level results are reported in Table 5 and Figure 7.

Aggregation and Robustness Analyses

Mean probability across retained spectra was the prespecified primary node-level aggregation method and provided a favorable balance of sensitivity and specificity in this analysis. Under patient-grouped validation, a single representative spectrum achieved accuracy of 89.4% and an AUROC of 0.930, while majority vote achieved accuracy of 92.7% and an AUROC of 0.953. Median probability performed similarly to mean probability. Maximum probability increased sensitivity to 97.9% but reduced specificity to 87.6%, resulting in more false-positive classifications.

The primary result remained stable in leave-one-patient-out validation, which produced an AUROC of 0.961, and leave-one-batch-out validation, which produced an AUROC of 0.952. Excluding the lowest-quality 10% of spectra produced an AUROC of 0.969. Restriction to 2 spectra per node reduced the AUROC to 0.951, whereas restriction to nodes retaining all 3 spectra gave an AUROC of 0.964. Area-normalized and shape-focused analyses produced lower AUROCs of 0.895 and 0.912, respectively, indicating that amplitude information contributed materially to discrimination. Full results are reported in Table 5.

Error Analysis

Using the cross-validation–derived summary threshold, the primary patient-grouped analysis classified 260 of 274 analyzable lymph nodes correctly, including 90 true-positive and 170 true-negative nodes. Moreover, 14 analyzable nodes were misclassified; 6 were false negative and 8 were false positive. The additional 6 pathology-negative ESS-assessed nodes that produced no analyzable node-level ESS result were reported separately as index-test no-results and were not included among the 14 model misclassifications.

Of the 6 false-negative nodes, 4 contained micrometastases measuring 0.3 - 1.6 mm, while 2 contained small peripheral macrometastases measuring 2.3 mm and 2.8 mm. The median false-negative node size was 8.5 mm, and none showed extranodal extension. The 8 false-positive nodes were clustered within 2 pathology-negative patients and showed marked reactive or inflammatory changes, including follicular hyperplasia, sinus histiocytosis, vascular congestion, surface hemorrhage, or suppurative inflammation. Summary findings are reported in Table 7, and detailed case-level findings are reported in Table S2 in Multimedia Appendix 1.

Table . Error analysis of the selected patient-grouped hybrid model. Misclassification status was determined using the cross-validation–derived summary node-level threshold of 0.38. The 6 pathology-negative node-level index-test no-results were reported separately and are not included in this table. Detailed case-level findings are reported in Table S2 in Multimedia Appendix 1.

Error category	n	Node size (mm), median (range)	Deposit size (mm), median (range)	Metastatic class	Neck level	Histopathological context	Out-of-fold probability, median (range)
False negative	6	8.5 (5-13)	1.2 (0.3 - 2.8)	4 micrometastases; 2 small peripheral macrometastases	4 at levels I-II; 2 at levels III-IV	Reactive change in 2; focal necrosis in 1; keratinization in 1; extranodal extension in 0	0.113 (0.028 - 0.239)
False positive	8	16.5 (9-22)	—[a]	All pathology negative	5 at levels I-II; 3 at levels III-IV	All arose in 2 patients with marked reactive or inflammatory change; focal necrosis in 1; keratinization in 0	0.915 (0.420 - 0.995)

[a]Not applicable.

Technical Repeatability and Variability

Within-node agreement across the 3 spatially separated acquisitions was highest for I630 intensity, with an intraclass correlation coefficient (ICC) of 0.91 (95% CI 0.88 - 0.93) and a median coefficient of variation of 7.4%. The corresponding ICCs were 0.89 for total spectral area, 0.84 for red shoulder height, 0.78 for blue trough depth, and 0.87 for spectrum-level hybrid probability. Intraclass correlation coefficients were calculated using a 1-way random-effects, absolute-agreement, single-measure model (ICC 1,1), with 95% CIs estimated from the corresponding F distribution.

Patient-grouped AUROC was 0.968 for acquisitions undertaken by operator A and 0.961 for acquisitions undertaken by operator B. The difference was 0.007 (95% CI –0.031 to 0.044). Acquisitions completed within 30 minutes of excision produced an AUROC of 0.967, compared with 0.958 for acquisitions completed after 30 minutes. Leave-one-batch-out validation produced an AUROC of 0.952. Because each lymph node was acquired by a single operator, the operator analysis compared

XSL•FO
RenderX

independent operator-specific case series rather than paired measurements of the same nodes and therefore did not estimate interoperator reproducibility directly. Operator, batch, and timing analyses were interpreted as sensitivity analyses because differences in patient and nodal case mix could not be fully separated from technical effects. Results are reported in Table S3 and Figure S2 in Multimedia Appendix 1.

Discussion

Principal Findings

This single-center model-development and internal-validation study found that ESS could discriminate metastatic from nonmetastatic cervical lymph nodes ex vivo. The selected hybrid model had an internally validated, exploratory AUROC estimate of 0.966 under patient-grouped cross-validation. This estimate may be affected by selection optimism because the same pooled out-of-fold predictions informed both model-family selection and performance reporting, and it should not be interpreted as independent validation of a prospectively fixed final model. At the cross-validation–derived summary threshold, secondary descriptive operating characteristics were sensitivity 93.8%, specificity 95.5%, and accuracy 94.9%. By comparison, validation grouped only at lymph-node level produced an AUROC of 0.989 and accuracy of 96.7%. The lower patient-grouped estimates illustrate the importance of preserving patient-level clustering when multiple lymph nodes originate from the same individual. The clinically relevant model output was a lymph-node probability derived from repeated spectra rather than classification of individual spectral acquisitions as independent observations.

Pathology-positive nodes showed markedly higher short-wavelength intensity and greater total spectral area, whereas pathology-negative nodes showed a higher I630/I360 ratio, higher intensity at 700 nm, a deeper local blue trough, and a larger red shoulder. This bidirectional spectral pattern may help explain why the hybrid model had a higher discrimination point estimate than the simpler candidate models. However, model-family comparisons were exploratory and were not designed to establish formal superiority. Possible contributors to the observed wavelength-dependent differences include changes in superficial scattering, cellular density, chromophore absorption, stromal architecture, and vascular content. This study did not include site-matched quantitative optical-histological measurements capable of establishing these mechanisms. These explanations are therefore hypothesis generating and should not be interpreted as demonstrated biological causes of the spectral differences.

This study extends, but is independent of, the earlier ESS study of cervical lymph nodes [14]. The previously published study examined 347 spectra obtained from 130 formalin-fixed lymph nodes in 13 patients and reported high sensitivity but lower specificity using selected wavelength intensities and a simple cutoff. The present cohort comprised fresh lymph nodes from 45 different patients. No participant, lymph node, raw spectrum, pathology outcome, or derived model output from the earlier dataset was included in this analytical dataset or used for model fitting, hyperparameter tuning, probability calibration, threshold selection, or performance estimation. All data-derived spectral and performance figures in this report were generated from the present cohort rather than from the previously published dataset. This analysis also differed by using 3 spatially separated acquisitions per node, explicit quality control, node-level probability aggregation, and internal validation grouped by patient.

The direction of some individual spectral features differed between the 2 studies. The earlier study found generally greater scattering intensity in pathology-positive measurements, including at longer wavelengths, whereas the present node-level analysis found higher I700 intensity, a higher I630/I360 ratio, a deeper blue trough, and a larger red shoulder in pathology-negative nodes. The differing wavelength-specific directions should not be interpreted as direct replication or contradiction because the studies differed materially in specimen state, acquisition context, analytical unit, preprocessing, aggregation, and modeling. The earlier specimens were formalin fixed and analyzed at individual-measurement level, whereas the present specimens were assessed fresh and spectral information was aggregated within lymph nodes. Taken together, the studies support the broader hypothesis that metastatic involvement is associated with differences in optical scattering profiles, while the wavelength-specific direction and magnitude of those differences require independent investigation under standardized acquisition conditions.

Repeated acquisition reduced dependence on a single contact site but did not remove technical variation. Within-node ICCs ranged from 0.78 for blue trough depth to 0.91 for I630 intensity, indicating that amplitude-related features were more repeatable than some local shape features. Operator-specific AUROCs were similar, but each node was assessed by only 1 operator, so a direct paired estimate of interoperator reproducibility was not possible. Future studies should use locked surface-preparation and timing procedures, record probe pressure and contact angle prospectively, and include repeated assessments of the same nodes by multiple operators.

This observation may also have a biological explanation. The ESS probe interrogates only a small superficial tissue volume, approximately 500×300×300 µm, meaning that each spectrum represents a local optical microsample rather than the entire node. Metastatic deposits may be focal, peripheral, central, or microscopic, and individual spectra may therefore capture slightly different aspects of nodal architecture. Aggregating multiple spectra from the same lymph node is consequently more than a statistical exercise; it represents repeated sampling of biological heterogeneity. The improvement observed with node-level aggregation suggests that repeated optical interrogation stabilizes classification by integrating information across multiple microenvironments within the same node. This interpretation is consistent with the probe geometry described in the original ESS nodal study, where collected photons originated from a defined superficial scattering volume rather than a whole-node measurement.

This node-level framing is particularly relevant when compared with imaging. Clinical examination, CT, MRI, PET-CT, and ultrasound remain essential in preoperative staging, yet each

XSL•FO
RenderX

can struggle with small-volume or biologically ambiguous nodal disease. Historical clinical and radiological studies have shown substantial false-negative and false-positive rates in the neck, and contemporary meta-analyses continue to show that advanced imaging is useful but imperfect [6,7,22,23]. ESS addresses a different point in the pathway. It is not a preoperative staging test. Its most realistic role is ex vivo assessment of dissected lymph nodes, helping to identify nodes that warrant closer histopathological attention, expedited review, or targeted deeper sectioning.

Patient-level sensitivity was 96% and specificity was 90% when a patient was classified as positive if at least 1 ESS-assessed node exceeded the cross-validation–derived summary threshold. This analysis provides an exploratory indication of performance at patient level, but it was based only on the ESS-assessed node set and should not be interpreted as complete pathological staging of the neck. The maximum-node approach may also be influenced by the number of nodes assessed per patient because patients contributing more nodes have more opportunities to generate a high node-level probability. Prospective evaluation in which the clinically relevant nodal yield is assessed according to a predefined sampling protocol will be required to establish patient-level clinical performance.

The potential clinical role of ESS should be framed as augmentation rather than replacement. Histopathology remains the reference standard and must continue to determine the final diagnosis. ESS may warrant prospective evaluation as an investigational triage tool for identifying high-probability nodes that could be prioritized for review or additional pathological assessment, while all nodes continue to undergo the required routine histopathological pathway. Whether this approach improves turnaround time, workload, or diagnostic safety remains to be demonstrated prospectively.

Any future clinical system would need to provide automated quality control and a simple probability or indeterminate output within a time compatible with gross specimen assessment. Evaluation should compare routine pathology with and without ESS-guided specimen prioritization and, where relevant, benchmark diagnostic performance and turnaround time against rapid sentinel-node or frozen-section pathways [24].

The observed node-level NPV of 96.6% at the cross-validation–derived summary threshold is potentially relevant to future triage applications but remains dependent on disease prevalence and should be interpreted as a secondary descriptive operating characteristic of this internally validated cohort. The prevalence of metastatic nodes will differ between elective and therapeutic neck dissections, tumor stages, anatomical sites, sampling strategies, and referral settings. Calibration was acceptable within the present cohort, but both predictive values and probability calibration may change in an independent population. External validation should therefore use a prospectively locked threshold, evaluate whether recalibration is required, and consider an indeterminate probability range rather than assuming that every node should receive a forced binary classification.

Strengths

The study has several strengths. Histopathology was used as the reference standard, repeated spectra were acquired from each node, and explicit quality-control criteria were applied before modeling. The primary validation kept all nodes and spectra from each patient together, and model fitting, hyperparameter selection, probability calibration where required, and threshold optimization were restricted to training data within the nested validation process. Performance was reported with patient-clustered CIs, calibration measures, patient-level results, case-level error review, and technical repeatability analyses. Moreover, 6 node-level index-test no-results were also reported explicitly rather than being treated as model classifications. The lower discrimination of area-normalized and shape-focused analyses suggests that amplitude information contributed materially to model performance in this cohort, although this finding reinforces the need for careful control of acquisition conditions and independent external validation.

Limitations

Several limitations remain. First, this was a single-center model-development and internal-validation study involving 45 patients. Although the analytical dataset contained 274 analyzable nodes and 760 retained spectra, observations were clustered within patients; the effective number of independent biological units was substantially smaller than the number of spectra or nodes, and spectrum-level fitting gave greater influence to patients contributing more nodes and to nodes retaining 3 rather than 2 spectra. Future work should compare patient- or node-weighted fitting with direct node-level spectral aggregation. Prospective external validation is, therefore, required. Second, the patient-level analysis represented the ESS-assessed node set rather than complete pathological staging of every node recovered from each neck dissection. The ESS-assessed nodes represented a pragmatic availability-based sample rather than the complete pathological nodal yield, so the observed node-level prevalence and diagnostic performance may not transfer directly to an unselected routine pathology population. Third, 6 of 280 ESS-assessed nodes produced no analyzable node-level results, and all 6 were pathology negative. The small number of no-results precluded reliable assessment of whether acquisition failure was associated with pathology, node characteristics, or other technical factors. Fourth, deeper levels and pancytokeratin immunohistochemistry were used selectively rather than systematically, so very small metastatic deposits could theoretically have been missed by the reference-standard assessment. Fifth, some technical variables were derived from archival records, while probe pressure and precise contact angle were not measured quantitatively. Sixth, each lymph node was assessed by only 1 operator, preventing paired assessment of interoperator reproducibility. Operator, acquisition batch, and case mix were not fully crossed within the same specimens, so operator-stratified and batch-based analyses could not completely separate technical effects from differences in patient or nodal characteristics. Seventh, amplitude contributed materially to discrimination and may be influenced by surface hydration, blood contamination, tissue handling, probe contact, system response, and elapsed time from excision. Eighth, the hybrid model family was selected because

XSL•FO
RenderX

it had the highest pooled patient-grouped out-of-fold AUROC point estimate among the candidate families. Its AUROC of 0.966 is therefore an internally validated, exploratory estimate of the selected model's performance, rather than independent validation of a prospectively fixed final model. Using the same out-of-fold predictions for model-family selection and performance reporting may introduce selection optimism; independent evaluation of a prospectively specified and locked model remains necessary. Ninth, the pooled binary operating characteristics were calculated using a summary threshold derived from the distribution of training-derived thresholds after repeated internal cross-validation and should be interpreted as descriptive rather than as validation of a prospectively locked clinical threshold. Tenth, the specimens were collected between 2016 and 2021 using a research ESS platform, so differences in contemporary devices, specimen handling, staging practice, and pathology workflows may affect transportability. Finally, 3 fixed acquisitions sampled only small superficial tissue volumes and may not consistently capture focal, peripheral, or heterogeneous metastatic deposits. Prospective studies should evaluate predefined acquisition protocols, denser or scanned sampling, automated quality control, and an independently locked model and decision threshold.

Future Work

A particularly important next step will be site-matched optical-pathological correlation. Future studies could coregister ESS acquisition points with digital histopathology and quantitative image analysis to determine whether wavelength-specific changes correlate with tumor area, nuclear density, keratinization, fibrosis, inflammatory infiltrates, vascularity, or other measurable tissue features. Such work would allow ESS findings to move beyond diagnostic discrimination towards biologically interpretable optical pathology and may clarify the cellular and subcellular origins of the wavelength patterns identified in this study.

Future evaluation should use a prospective multicenter design in which the acquisition protocol, preprocessing sequence, quality-control rules, classifier, probability-calibration procedure, and decision threshold are specified and locked before evaluation of the external validation cohort. The study should include multiple devices, operators, pathology teams, and clinically relevant patient groups. Outcomes should include discrimination, calibration, operator reproducibility, acquisition failure, indeterminate results, turnaround time, pathology workload, missed disease, and costs. The appropriate comparison is not ESS vs histopathology as competing diagnostic methods, but routine pathology with and without ESS-guided specimen prioritization.

Conclusion

In this single-center ex vivo model development and internal validation study, ESS showed promising node-level discrimination of metastatic cervical lymph nodes after repeated spectral acquisition and node-level aggregation. Patient-grouped validation produced lower performance estimates than validation grouped only at lymph-node level, emphasizing the importance of accounting for clustering of multiple nodes within patients. The findings remain exploratory and do not establish clinical utility. ESS should therefore be regarded as an investigational adjunct for pathology triage rather than a clinically ready diagnostic test. Prospective multicenter external validation using a prespecified and locked analytical pipeline and decision threshold, together with explicit assessment of no-results, calibration, paired operator reproducibility, acquisition variability, and workflow performance, is required before clinical implementation.

Acknowledgments

No generative AI was used during any stage of this article preparation.

Funding

The authors declared no financial support was received for this work.

Data Availability

The deidentified analytical dataset, data dictionary, and analysis code underlying the reported analyses may be made available by the corresponding author on reasonable request, subject to approval by the relevant institutional data custodian and confirmation that the proposed use is compatible with participant consent, confidentiality requirements, and applicable research-governance arrangements.

Authors' Contributions

Conceptualization: WJ, ZH, DR, CH
Methodology: WJ, ZH, DR, CH
Data acquisition and curation: WJ, ZH, DR, CH
Formal analysis and interpretation: WJ, ZH, DR, CH
Writing – original draft: WJ, ZH, DR, CH
Writing – review and editing: WJ, ZH, DR, CH
All authors reviewed and approved the final manuscript.

XSL•FO
RenderX

Conflicts of Interest

WJ is an Associate Editor for JMIR Medical Education at the time of this publication. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Supplementary tables and figures.
[DOCX File, 611 KB - cancer_v12i1e103960_app1.docx]

Checklist 1
TRIPOD+AI completed checklist.
[DOCX File, 19 KB - cancer_v12i1e103960_app2.docx]

Checklist 2
STARD-AI completed checklist.
[DOCX File, 18 KB - cancer_v12i1e103960_app3.docx]

Checklist 3
STARD 2015 completed checklist.
[DOCX File, 16 KB - cancer_v12i1e103960_app4.docx]

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, risk factors, and prevention of head and neck squamous cell carcinoma. Med Sci (Basel) 2023 Jun 13;11(2):42. [doi: 10.3390/medsci11020042] [Medline: 37367741]
3. Dunn LA, Ho AL, Pfister DG. Head and neck cancer: a review. JAMA 2026 Feb 10;335(6):531-541. [doi: 10.1001/jama.2025.21733] [Medline: 41396597]
4. Alqutub S, Alqutub A, Bakhshwin A, et al. Histopathological predictors of lymph node metastasis in oral cavity squamous cell carcinoma: a systematic review and meta-analysis. Front Oncol 2024;14:1401211. [doi: 10.3389/fonc.2024.1401211] [Medline: 38835393]
5. Yu YF, Cao LM, Li ZZ, et al. Frequency of lymph node metastases at different neck levels in patients with oral squamous cell carcinoma: a systematic review and meta-analysis. Int J Surg 2025 Jan 1;111(1):1285-1300. [doi: 10.1097/JS9.0000000000001953] [Medline: 39037727]
6. Alsibani A, Alqahtani A, Almohammadi R, et al. Comparing the efficacy of CT, MRI, PET-CT, and US in the detection of cervical lymph node metastases in head and neck squamous cell carcinoma with clinically negative neck lymph node: a systematic review and meta-analysis. J Clin Med 2024 Dec 14;13(24):7622. [doi: 10.3390/jcm13247622] [Medline: 39768545]
7. Guedj D, Neveü S, Becker M, Mermod M. FDG PET-CT for the detection of occult nodal metastases in head and neck cancer: a systematic review and meta-analysis. Cancers (Basel) 2024 Aug 24;16(17):2954. [doi: 10.3390/cancers16172954] [Medline: 39272812]
8. Jiang S, Locatello LG, Maggiore G, Gallo O. Radiomics-based analysis in the prediction of occult lymph node metastases in patients with oral cancer: a systematic review. J Clin Med 2023 Jul 28;12(15):4958. [doi: 10.3390/jcm12154958] [Medline: 37568363]
9. Khong B, Ferlito S, Quek S, et al. Past, present, and future diagnostic methods for the early noninvasive detection of oral premalignant lesions: a state of the art and systematic review. Ear Nose Throat J 2024 May 2:1455613241245204. [doi: 10.1177/01455613241245204] [Medline: 38695398]
10. Zhang Y, Li Z, Zhang C, et al. Recent advances of photodiagnosis and treatment for head and neck squamous cell carcinoma. Neoplasia 2025 Feb;60:101118. [doi: 10.1016/j.neo.2024.101118] [Medline: 39721461]
11. Hanna K, Asiedu AL, Theurer T, et al. Advances in Raman spectroscopy for characterising oral cancer and oral potentially malignant disorders. Expert Rev Mol Med 2024 Oct 8;26:e25. [doi: 10.1017/erm.2024.26] [Medline: 39375841]
12. Wu IC, Chen YC, Karmakar R, et al. Advancements in hyperspectral imaging and computer-aided diagnostic methods for the enhanced detection and diagnosis of head and neck cancer. Biomedicines 2024 Oct 11;12(10):2315. [doi: 10.3390/biomedicines12102315] [Medline: 39457627]
13. Alajaji SA, Sabzian R, Wang Y, Sultan AS, Wang R. A scoping review of infrared spectroscopy and machine learning methods for head and neck precancer and cancer diagnosis and prognosis. Cancers (Basel) 2025 Feb 26;17(5):796. [doi: 10.3390/cancers17050796] [Medline: 40075644]

XSL•FO
RenderX

14. Jerjes W, Swinson B, Pickard D, Thomas GJ, Hopper C. Detection of cervical intranodal metastasis in oral cancer using elastic scattering spectroscopy. Oral Oncol 2004 Aug;40(7):673-678. [doi: 10.1016/j.oraloncology.2004.01.009] [Medline: 15172636]
15. Bigio IJ, Mourant JR. Ultraviolet and visible spectroscopies for tissue diagnostics: fluorescence spectroscopy and elastic-scattering spectroscopy. Phys Med Biol 1997 May;42(5):803-814. [doi: 10.1088/0031-9155/42/5/005] [Medline: 9172260]
16. Bigio IJ, Bown SG, Briggs G, et al. Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results. J Biomed Opt 2000 Apr;5(2):221-228. [doi: 10.1117/1.429990] [Medline: 10938787]
17. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
18. Sounderajah V, Guni A, Liu X, et al. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. Nat Med 2025 Oct;31(10):3283-3289. [doi: 10.1038/s41591-025-03953-8] [Medline: 40954311]
19. Moons KGM, Damen JAA, Kaul T, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ 2025 Mar 24;388:e082505. [doi: 10.1136/bmj-2024-082505] [Medline: 40127903]
20. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. BMJ 2022 May 18;377:e070904. [doi: 10.1136/bmj-2022-070904] [Medline: 35584845]
21. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015 Oct 28;351:h5527. [doi: 10.1136/bmj.h5527] [Medline: 26511519]
22. Woolgar JA, Vaughan ED, Scott J, Brown JS. Pathological findings in clinically false-negative and false-positive neck dissections for oral carcinoma. Ann R Coll Surg Engl 1994 Jul;76(4):237-244. [Medline: 8074384]
23. Wide JM, White DW, Woolgar JA, Brown JS, Vaughan ED, Lewis-Jones HG. Magnetic resonance imaging in the assessment of cervical nodal metastasis in oral squamous cell carcinoma. Clin Radiol 1999 Feb;54(2):90-94. [doi: 10.1016/s0009-9260(99)91066-8] [Medline: 10050735]
24. Abukhder M, Sahovaler A, Vrakas P, McGurk M, Thavaraj S, Schilling C. Frozen section analysis of sentinel nodes in patients with oral squamous cell carcinoma: a systematic review and meta-analysis. JAMA Otolaryngol Head Neck Surg 2024 Oct 10;150(11):1021-1028. [doi: 10.1001/jamaoto.2024.3094] [Medline: 39388171]

Abbreviations

AUROC: area under the receiver operating characteristic curve
CT: computed tomography
ESS: elastic scattering spectroscopy
ICC: intraclass correlation coefficient
MRI: magnetic resonance imaging
NPV: negative predictive value
PET-CT: positron emission tomography-computed tomography
PPV: positive predictive value
SNV: standard normal variate
STARD: Standards for Reporting of Diagnostic Accuracy Studies
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

Edited by M Balcarras; submitted 07.Jun.2026; peer-reviewed by K Bogusiak, T Hayashi; revised version received 06.Sep.2026; accepted 09.Sep.2026; published 13.Oct.2026.

Please cite as:
Jerjes W, Hamdoon Z, Rashed D, Hopper C
Detection of Cervical Nodal Metastasis in Oral Squamous Cell Carcinoma Using Elastic Scattering Spectroscopy: Ex Vivo Diagnostic Modeling Study
JMIR Cancer 2026;12:e103960
URL: https://cancer.jmir.org/2026/1/e103960
doi:10.2196/103960

© Waseem Jerjes, Zaid Hamdoon, Dara Rashed, Colin Hopper. Originally published in JMIR Cancer (https://cancer.jmir.org), 13.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License

XSL•FO
RenderX

(https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Readiness for Digital Health Literacy Transformation and Intervention Preferences From the Perspectives of Patients With Cancer, Caregivers, and Health Care Professionals: Qualitative Interview Study

Hind Mohamed[1], MBBS, MRCPI, MSc; Turki Alanzi[2], BSc, MSc, PhD; Jon Salsberg[1], BA, MA, PhD; Mudathir Mohamed[3], MBBS, FRCEM; Dervla Kelly[1], BSc, MPSI, MSc, PhD

[1]Faculty of Education and Health Sciences, School of Medicine, University of Limerick, Henry Street, Limerick, Ireland

[2]Health Information Management and Technology Department, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

[3]Emergency Medicine Department, E1 Health Cluster, King Fahad Specialist Hospital, Dammam, Saudi Arabia

Corresponding Author:
Hind Mohamed, MBBS, MRCPI, MSc
Faculty of Education and Health Sciences, School of Medicine, University of Limerick, Henry Street, Limerick, Ireland

Abstract

Background: Technology is changing the way the world communicates and how we learn, remember, and transform information. The ascendancy of the internet has dramatically altered the landscape of health information access and seeking behaviors. This transformation is embodied by the concept of digital health literacy (DHL) and the need for interventions that improve DHL.

Objective: This study aims to explore readiness for DHL transformation and intervention preferences from the perspectives of patients with cancer, caregivers, and health care professionals.

Methods: We conducted semistructured telephone and on-site interviews with 19 patients with cancer, 6 caregivers, and 10 oncology health care professionals. Purposive sampling was used to recruit the participants. We followed the 7 stages of the Framework Method analysis: transcription, familiarization with the interview, coding, developing a working analytical framework, applying the analytical framework, charting the data into the framework matrix, and interpreting the data. This was used to investigate participants' beliefs about technology adoption, their preferences for DHL training, and the facilitating conditions for adopting such training. We used a hybrid deductive-inductive approach to data analysis, starting with a priori themes and allowing emergent themes to develop as the analysis progressed. The Unified Theory of Acceptance and Use of Technology informed our data generation and analysis.

Results: The following 6 themes emerged from the analysis: introducing technology-driven solutions, simplifying technology training materials, providing user-friendly training materials, patient-centered care, partnership, and addressing cultural and linguistic barriers. We found that patients with cancer and caregivers were self-sufficient and motivated to use new technology to improve their DHL; however, health care professionals were concerned about the reliability of online information. By mapping interview themes to the Unified Theory of Acceptance and Use of Technology, we identified specific recommendations for the creation of a DHL intervention: content should be concise, easy to understand, and web-based; content should include training on how to identify reliable cancer information; patients would like to be involved in content creation in addition to medical and government stakeholders; and content should be accessible in multiple languages.

Conclusions: Patients with cancer and their caregivers were motivated to use modern technology to improve cancer DHL, despite the quality issues raised by health care professionals. The participants' preferences regarding DHL training align with the innovative microlearning pedagogy. Microlearning could leverage technology to deliver tailored DHL training for patients with cancer. Collaboration with multiple medical and nonmedical stakeholders could facilitate the delivery of cancer DHL training. Future work should focus on designing and assessing the feasibility of implementing a microlearning-based DHL training program that involves end users and diverse stakeholders.

(JMIR Cancer 2026;12:e77738) doi:10.2196/77738

KEYWORDS

digital health; eHealth; internet; mobile health; mHealth; digital health literacy; intervention; training; microlearning; microlearning-based; Saudi Arabia

XSL•FO
RenderX

Introduction

Online technologies have changed societies and affect our daily lives in many ways. This is reflected in the growing use of the internet and mobile devices by consumers to access health information and services [1]. Saudi Arabia's Vision 2030 outlines an ambitious plan to transform the health care sector through digitalization [2]. This transformation includes aspects of digital health literacy and technology-driven health engagement, focusing on improving access, efficiency, and patient empowerment through digital health [3].

In Saudi Arabia, approximately two-thirds or more of individuals get their health-related information from social media or the internet [4]. Internet use can be particularly effective when there are barriers to accessing traditional health information, especially for stigmatized conditions such as cancer [5,6]. Patients with cancer widely seek digital health information, as demonstrated in a study reporting that 76.4% of survivors of cancer use the internet for health-related information [7].

Digital health is now considered a complement to cancer care to meet key unmet needs, such as "support to help cope with a diagnosis" and "cancer education and information" [8]. Indeed, one study reported "information and education" as one of the most needed digital health functions for patients with cancer, caregivers, and health care providers [9].

Digital health literacy (DHL) is the ability to use information and communication technologies to find, evaluate, create, and communicate health information, and it requires both cognitive and technical skills [10]. Information searching is often described as a problem-solving task that involves a series of complex cognitive processes, such as inferences and decision-making [10]. Complex or confusing digital health information is especially difficult for patients with cancer since their cognitive abilities decrease due to the effects of the cancer and chemotherapy treatment [11]. A 2023 systematic scoping review found that patients with cancer and caregivers struggle to identify reliable information resources and can be misled or misinterpret information online [12]. A study in Saudi Arabia demonstrated that participants with higher DHL scores were more likely to have increased breast cancer literacy [13]. Another study reported that almost one-third of participants face challenges formulating precise health-related inquiries, highlighting the critical need for specialized DHL education initiatives in Saudi Arabia [14]. Mackert et al [15] and Zarcadoolas et al [16] have advocated for user-friendly educational interventions to bridge the DHL gap.

Health education needs to adapt and develop innovative ways of relating appropriately to the way we live, work, and learn today. A recent systematic review reported many interventions to improve laypeople's DHL, including interactive workshops, community outreach programs, and educational initiatives. However, no one has used a novel educational approach or targeted patients with cancer [17]. Low DHL in cancer care contributes to adverse health outcomes, hinders self-management of patients' conditions, and is considered a barrier to engaging with digital health technologies [18].

Digital health readiness is defined as a patient's ability and comfort with using digital tools for health care engagement [19]. This concept expands on existing literature describing DHL to include factors such as health care trust, acceptance of technology, and relevance to care. Current approaches to DHL intervention design are evolving from viewing patients as merely passive consumers of care to viewing them as active coproducers in their care management [20,21]. Coproduction in health care delivery is a reciprocal, collaborative process in which patients and health care providers work together to deliver care, sharing information and setting goals to create value [20]. Digital health interventions in cancer care have been reported to reduce symptoms [22], improve health behaviors [23], and promote psychological well-being and quality of life [24].

Nutbeam [25] outlined the need to invest in research to improve our understanding of how effective health interventions should be implemented. Eccles et al [26] further argued that we need to see greater use of theoretical approaches underpinning the design and implementation of interventions. Although theories have been used to evaluate digital health interventions [27], there is a lack of theoretically informed research in the field of digital health readiness.

The Unified Theory of Acceptance and Use of Technology (UTAUT) is a comprehensive digital health readiness framework that integrates and extends various theories to explain the acceptance and use of technology [28]. It incorporates elements from 8 previous models of technology acceptance, namely the Theory of Reasoned Action, the Technology Acceptance Model, the Motivational Model, the Theory of Planned Behavior, a combined model of Technology Acceptance Model and Theory of Planned Behavior, the Model of Personal Computer utilization, the Diffusion of Innovations theory, and Social Cognitive Theory [28]. It has 4 determinants (performance expectancy, effort expectancy, social influence, and facilitating conditions) and 4 moderators (gender, age, experience, and voluntariness of use) [28]. The UTAUT model has demonstrated its broad applicability across diverse fields, including the Internet of Things [29], artificial intelligence products [30], and electronic health technology [31,32], but has not yet been applied to educational interventions targeting patients with cancer.

The aim of the paper was to explore the readiness for DHL transformation and intervention preferences from the perspectives of patients with cancer, caregivers, and health care professionals using the UTAUT model. To our knowledge, our paper is the first to use the UTAUT to explore readiness for DHL transformation and intervention preferences from the perspectives of patients with cancer, caregivers, and health care professionals. The UTAUT was chosen to help us understand the following objectives: participants' beliefs about technology adoption, preferences regarding DHL training adoption, and the conditions that facilitate DHL training adoption. Knowledge of technology adoption is crucial to ensure the delivery of quality cancer care education.

XSL•FO
RenderX

Methods

The COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [33] was used to guide reporting of the study (Checklist 1). It also provides a structured account of the study's methodological rigor, transparency, and researcher reflexivity.

Study Design

The study design domain describes a qualitative content analysis approach [34] informed by the UTAUT. The UTAUT was relevant because it enables in-depth exploration of the unconscious behaviors that drive the uptake of health education initiatives. Semistructured on-site and telephone interviews were conducted with patients with cancer attending the oncology department for chemotherapy treatment and their caregivers, as well as oncology health care professionals at King Fahad Specialist Hospital (KFSH), Dammam, Saudi Arabia. The researchers selected the semistructured interview format because it allowed them to preformulate questions, compare data across participants, and adjust questions as necessary during the interview [35].

Sampling and Recruitment

Purposive sampling was used to provide rich, relevant, and diverse data that were not intended to be representative of a broader population [36]. We deliberately sampled a mix of ages, genders, and socioeconomic statuses from different participant types, including patients with cancer and their caregivers and oncology health care professionals, based on preidentified inclusion and exclusion criteria (Textbox 1). Patients and caregivers were recruited face to face. Paper-based posters, recruitment emails, and telephone calls were also used to facilitate the recruitment. Oncology health care professionals were recruited through gatekeepers in the oncology department. We recruited 44 participants, including 24 patients, 6 caregivers, and 14 health care professionals; 5 patients and 4 health care professionals declined to participate. In total, 35 participants agreed to participate in the study. Saturation was used as a guiding principle to assess the adequacy of the purposive sample [37].

Textbox 1. Inclusion and exclusion criteria for the different participant types, including people with cancer, their caregivers, and health care professionals involved in cancer care.

Inclusion criteria

1. Patients with a personal history of cancer who are considered in remission
2. Family/friend caregivers involved in helping care for patients with cancer for 2 or more hours per week, providing emotional or other support
3. Health care professionals involved in cancer care (assistant consultants/consultants, nurses, health educators)
4. ≥18 years old, male or female, and able to understand and answer in Arabic
5. Must have used the internet to search for cancer information

Exclusion criteria

1. Patients with no personal history of cancer; patients with a personal history of cancer but not considered in remission
2. Caregivers with no significant involvement in helping with care as defined by less than 2 hours per week of emotional or other support
3. Health care professionals not involved in cancer care
4. <18 years old, unable to understand and answer in Arabic
5. Never used the internet to search for cancer information

Interview Question Design

A panel of 5 stakeholders agreed on the interview questions. The panel included 2 clinicians and 3 academics with expertise in DHL. The semistructured interview questions were open-ended focusing on the following areas: (1) digital learnability, (2) DHL training mode of delivery, (3) DHL training information content and display, (4) DHL training interval and duration, (5) DHL training design, and (6) facilitators of the DHL training adoption. The questionnaire initially had 10 questions. An additional question was added following a small number of pilot interviews to better understand the facilitators for adopting DHL training. This question and the DHL training design question received the most interest from patients with cancer and caregivers. Health care professionals were most interested in discussing the DHL training content; however, differences emerged among patients with cancer, their caregivers, and health care professionals' preferences. We developed the interview guide in English to allow discussion with the English-speaking researchers. The interview guide was translated into Arabic, as the interviews were conducted in Arabic per the protocol and inclusion criteria. Participants were provided with an Arabic version of the topic guide before consultation. Multimedia Appendix 1 includes the English and Arabic versions of the interview guide. Sociodemographic data were collected via a brief pre-interview questionnaire that was posed by the interviewer and answered by the participants.

Data Collection

A female PhD researcher (HM) with no prior interaction with participants carried out the interviews in Arabic. No one else was present besides the participants and the researcher. The researcher disclosed that the study was part of her doctoral research on DHL. The researcher has professional-level fluency in Arabic and English and is a DHL expert, with experience in

XSL•FO
RenderX

qualitative research, the research topic, and Saudi culture. The researcher completed the Good Clinical Practice course before conducting the study. The researcher conducted 35 telephone and on-site (KFSH) interviews with patients with cancer, caregivers, and oncology health care professionals following written and verbal informed consent. The interviews were conducted in cycles between August 15, 2024, and February 15, 2025, at times and locations convenient to the participants. Interviews were audio-recorded, transcribed, and coded using a thematic analysis approach within the Framework Method described in the following paragraphs [38]. No field notes were made during and/or after the interview. No repeat interviews were carried out. At the close of each interview, the audio recordings were transcribed into Microsoft Word. The description of the coding tree was not included in the main manuscript nor its appendices but can be shared upon reasonable request. Most of the interviews lasted 30 minutes to 45 minutes. However, interviews with patients with cancer and caregivers lasted between 30 minutes and 35 minutes, while health care professionals ' interviews lasted longer (35 - 45 min). There were no differences in interview quality between patients with cancer or caregivers and health care professionals or between telephone and on-site interviews. We reached data saturation after conducting the 35 interviews, as no new themes were emerging during interviews 34 and 35. We first assessed code saturation to ensure it was sufficient to understand the issues identified. Second, we assessed meaning saturation to ensure that no further dimensions, nuances, nor insights into the issues could be found. We also evaluated whether specific characteristics of codes influenced code or meaning saturation to provide parameters for estimating saturation based on the nature of codes developed in a study [37]. These approaches provided confidence that the collected data adequately captured the breadth of relevant perspectives within the study population, thereby ensuring the depth and credibility of the findings.

Data Analysis

Audio recordings of the interviews were made with the participants' consent and transcribed verbatim in Arabic. The quotes were translated forward and backward into English by the principal investigator (HM; researcher translator). A bilingual health literacy expert consultant reviewed the translation (MM). The transcriptions were analyzed by iteratively reading transcripts. Thematic analysis was conducted using the Arabic versions of the transcripts; the results section used English translations of the quotations for readability by a wider audience. We used a hybrid deductive-inductive approach to data analysis [38], starting with a priori themes and allowing emergent themes to develop as the analysis progressed. This blend helped balance theoretical grounding with a rich, data-driven understanding of the topic. Given the potential value of theory as a heuristic for qualitative analysis processes [39], data generation and analysis were informed by the UTAUT. Emergent themes were subsequently mapped onto the 4 determinants of the UTAUT [28]. The mapping process relied on iteratively moving between the emergent themes and the UTAUT determinants to build knowledge about applying UTAUT to the project data.

The Framework Method sits within a broad family of analysis methods often termed thematic analysis or qualitative content analysis. The Framework Method can be adapted for use with qualitative analyses that are deductive, inductive, or combined [40]. It also facilitated constant comparative techniques by reviewing data across the matrix; as such, it was found to be appropriate for our paper as it involves different types of participants (patients, caregivers, and health care professionals) [40]. The following 7 stages of the Framework Method analysis were followed: transcription, familiarization with the interview, coding, developing a working analytical framework, applying the analytical framework, charting the data into the framework matrix, and interpreting the data [40].

Reflexivity measures were applied to mitigate potential bias arising from the researcher's background in DHL. Two researchers from different professional backgrounds independently conducted coding in NVivo 14 to explore the development of inductive themes and the application of deductive ones. This approach helped to ensure intercoder reliability and to heighten reflexivity [41].

Transcript analysis was carried out at primary and secondary levels. Whereas first-level coding created descriptive, low-inference codes directly from the data, second-level analysis allowed for the reorganization of primary codes into appropriate categories, subthemes, and themes [41].

Member checking is crucial in qualitative research to ensure credibility and validity, reduce researcher bias, increase trustworthiness of the findings, and ensure participants' expertise was fully used [42]. A researcher prepared a draft of the findings, including anonymous quotations, 1 month after data collection and shared it with one-half of the participants individually (an equal number of health care professionals and patients). The researcher asked questions about accuracy, interpretation, completeness, fairness and respect, and clarification. The participants provided their answers through written feedback. The feedback was used to revise interpretations, clarify ambiguities, and add missing information.

To ensure the awareness of reflexivity in that one's own beliefs can affect interpretation [41], 3 additional members of the research team, a cancer survivor, and a consultant oncology doctor, reviewed the themes, subthemes, and categories and refined them, seeking clarity on language and suggesting additional work. The research team worked in a group throughout the analysis, comparing coding, discussing how the data related to the UTAUT coding frame, and refining the coding frame. Data analysis clinics were also used to strengthen the initial inductive analysis and to map themes onto the UTAUT. A narrative description was provided for each of the themes speaking to readiness for DHL transformation and intervention preferences.

Ethical Considerations

All the procedures performed in the study involved human participants. Ethical approval was obtained before data collection from the KFSH research ethics committee (April 23, 2024/EXT0429). The method of collecting and analyzing data was carried out in a manner that was compliant with all

XSL•FO
RenderX

applicable ethical norms. Full disclosure was made regarding the investigation's objective and the legal rights of those participating. Each participant was given a pseudonym to preserve confidentiality and maintain anonymity. Every participant provided informed consent before the interview began. Participation in the study was entirely voluntary throughout the process. Participants did not receive compensation for their participation in the study. Data will be stored on the principal investigator's (HM) password-protected Outlook OneDrive account for 2 years then destroyed.

Results

Participants' Characteristics

The first 24 patients who met the study criteria were invited to participate in the study; 5 patients declined (3 men and 2 women) because they were too ill to be interviewed on the day of the appointment. In addition, 6 caregivers agreed to participate in the study. Of the 14 health care professionals recruited, 4 refused to participate (3 nurses and 1 health educator) because they were too busy during the data collection period. Only 10 health care professionals (assistant consultants and consultants) agreed to participate in the study. The total number of patients, caregivers, and health care professionals interviewed was 35. Table 1 presents the participants' characteristics including age, gender, and participant category. Questions were asked regarding demographics, including education, and socioeconomic status; however, many participants declined to disclose these, most likely as they were considered sensitive information. Given the extent of the missing data, these variables have not been reported.

Table . Participant characteristics for those who participated (n=35) in the semistructured interviews in King Fahad Specialist Hospital, Dammam, Saudi Arabia, between August 15, 2024, and February 15, 2025.

Participant category	Participants, n	Male, n	Female, n	Age (years), range
Patients with cancer	19	9	10	30 - 65
Cancer caregivers	6	1	5	30-50
Oncology health care professionals	10	7	3	40-55

Qualitative Findings

Findings on readiness for the DHL transformation and intervention preferences are presented in Textbox 2 and framed around the 4 determinants of the UTAUT. We identified 6 themes and 9 subthemes under the UTAUT determinants. For the first 2 subthemes, there were categories that were considered individually as indicated in the following sections.

XSL•FO
RenderX

Textbox 2. The 4 determinants of the Unified Theory of Acceptance and Use of Technology and their definitions, themes, subthemes, and categories describing the factors that influence digital health literacy training preferences in people with cancer in a sample size of 35 participants.

Performance expectancy (user's belief that using technology will enhance their search performance and productivity)

1. Introduction of technology-driven solutions
    a. Digital literacy learnability
        1. Self-efficacy for using the technology to search for cancer information
        2. Willingness to engage with new technology to improve cancer digital health literacy
    b. Demand for technology-driven solutions to improve cancer digital health literacy
        1. Addressing the digital divide
        2. Designing an educational digital health literacy training

Effort expectancy (the degree of ease associated with the use of technology)

1. Simplifying technology training materials
    a. Bite-sized information
    b. Desire for reliable information
2. Providing user friendly training materials
    a. Self-paced training
    b. Short-duration training

Social influence (the impact of external factors on an individual's decision to adopt technology)

1. Patient-centered care
    a. Willingness to be involved in the design of the training material

Facilitating conditions (availability of resources, support, and infrastructure necessary for technology adoption and use)

1. Partnership
    a. Collaboration with multiple stakeholders (medical and nonmedical)
2. Addressing cultural and linguistic barriers
    a. Providing bilingual training material

Performance Expectancy

There were 2 subthemes under performance expectancy: digital literacy learnability and demand for technology-driven solutions to improve cancer DHL.

Digital Literacy Learnability

Digital literacy learnability refers to the self-efficacy of patients with cancer and their willingness to engage with technology.

Patients with cancer and their caregivers spoke about their self-efficacy in navigating online cancer information. Inversely, many health care professionals expressed concerns about the quality of cancer information retrieved by patients online.

> *I feel confident when opening web browsers to search for cancer information.* [Patient 5]
>
> *I feel confident to click on a link to visit a cancer website.* [Caregiver 2]
>
> *Nowadays, patients are increasingly dependent on the internet to search for cancer information; however, most of the information they brought to me was unreliable.* [Health care professional 3]

Patients with cancer and caregivers expressed being motivated to engage with new technologies, with a view to learning and improving their cancer care. Health care professionals also noted patients' eagerness to search online, often before their doctor's visit.

> *I am interested to learn about my health through the new technology channels.* [Patient 8]
>
> *I believe learning through the new technology can affect patients' health positively.* [Caregiver 5]
>
> *Most of the patients nowadays prefer searching online even before visiting the doctor.* [Health care professional 7]

XSL•FO
RenderX

Demand for Technology-Driven Solutions to Improve Cancer DHL

Demand for technology-driven solutions to improve cancer DHL refers to addressing the digital divide to ensure that technology is accessible to all users, regardless of their geographical location or socioeconomic status and designing educational DHL training.

Patients, caregivers, and health care professionals reported a preference for web-based DHL training. They believed that online training would help address the digital divide and facilitate digital inclusion.

> *Going far away from my home to get the training is costly, and I don't have the energy like before.* [Patient 5]
>
> *I won't be able to attend training conducted even 5 minutes away from my house, as it requires a lot of dedication.* [Caregiver 4]
>
> *I would say if the educational training is provided online, it will be easily accessible for the consumers and more engaging.* [Health care professional 2]

Patients and caregivers spoke about designing a smartphone app to deliver the DHL training. On the other hand, many health care professionals demonstrated a need for an educational video to improve cancer DHL.

> *I feel more engaged with smartphone applications as I hold my phone most of the day.* [Patient 15]
>
> *If the educational materials are provided through the applications, they will be easily accessible.* [Caregiver 6]
>
> *I believe designing an educational video to be available for the general public would be of great value, as they provide easy-to-catch information.* [Health care professional 7]

Effort Expectancy

There were 4 subthemes under effort expectancy: bite-sized information, desire for reliable information, self-paced training, and short-duration training.

Bite-Sized Information

Patients with cancer and caregivers emphasized the importance of providing concise educational content to reduce cognitive effort and time. Health care providers discussed the impact of chemotherapy on the learning abilities of patients with cancer, necessitating the provision of small-bundle information.

> *I am a busy person with my family, work, and now a cancer patient; I don't have that energy like before to learn complex educational materials.* [Patient 11]
>
> *I think providing educational material in small bites may reduce the learning burden.* [Caregiver 1]
>
> *Patients with cancer face many difficulties concerning learning, especially after receiving chemotherapy; they would benefit from small pieces or guides of information.* [Health care professional 8]

Desire for Reliable Information

Patients and caregivers expressed a need for training on identifying reliable cancer information. They recommended learning tips for evaluating the reliability of online information and identifying high-quality websites, while health care providers spoke about the importance of educating laypeople about common cancer misconceptions.

> *I want to learn about excellent quality websites for lay people to search cancer information.* [Patient 13]
>
> *I want to learn how to evaluate the reliability of online health information.* [Caregiver 1]
>
> *Many patients with cancer come to my clinic with worsening symptoms and poor prognosis due to misinformation. It is essential to increase public awareness about common cancer misconceptions.* [Health care professional 3]

Self-Paced Training

Most patients and caregivers stated that they like the flexibility of online training and being able to go at their own pace.

> *Accessing the training material with no fixed schedule and the ability to set my times would help me to engage.* [Patient 12]
>
> *I want to reach the training when I want, at the time I want, and with the time I have.* [Caregiver 6]

Short-Duration Training

Patients, caregivers, and health care professionals spoke about providing short-duration training to reduce cognitive load and optimize attention span for memorizing information.

> *It could even be an educational training that does not exceed 10 minutes to be able to focus and memorize it.* [Patient 11]
>
> *They can be 10 minutes, but then they are almost too long.* [Caregiver 6]
>
> *Short training, 5 minutes maximum to reduce the learning load.* [Health care professional 8]

Social Influence

There was 1 subtheme under social influence: willingness to be involved in the design of the DHL training.

Patients with cancer expressed a strong desire to be involved in the design of the training material.

Caregivers and health care professionals believed that the patient voice should be heard to ensure relevance and acceptability of training materials.

> *Nobody knows it better than yourselves, having gone through it… You can set up all the processes on paper: This happens, this happens, this happens, but until you actually press play and go through it.* [Patient 6]
>
> *People who set the training materials have never passed through a cancer journey. They see it from the other side of the desk, but to actually involve*

XSL•FO
RenderX

> *people who have been on the receiving end.* [Caregiver 15]
>
> *Getting patients with cancer involved could help to cover a huge gap that comes between the delivery and acceptance of the training materials.* [Health care professional 6]

Facilitating Conditions

There were 2 subthemes under facilitating conditions: collaboration with multiple stakeholders (medical and nonmedical) and providing bilingual training materials.

Collaboration With Multiple Stakeholders (Medical and Nonmedical)

Patients with cancer, caregivers, and health care professionals spoke about the essential role of collaborative work in the design of the training materials. They emphasized the role of technology assistants, governmental authorities, and health care professionals in designing the educational materials.

> *I trust the oncology doctors, nurses, health educators, researchers, and governmental authorities. Why don't they put their hands together to deliver such training?* [Patient 3]
>
> *It is not an easy game that one person can play; it is the health of cancer patients… it needs different expertise in different sectors. It is worth connecting to deliver such a lifesaving game.* [Caregiver 1]
>
> *I would see it as a cooperative work; one hand will not clap. We are not in front of the computer to see how this work can be done. We want to work as one unit to see this work delivered properly.* [Health care professional 1]

Providing Bilingual Training Materials

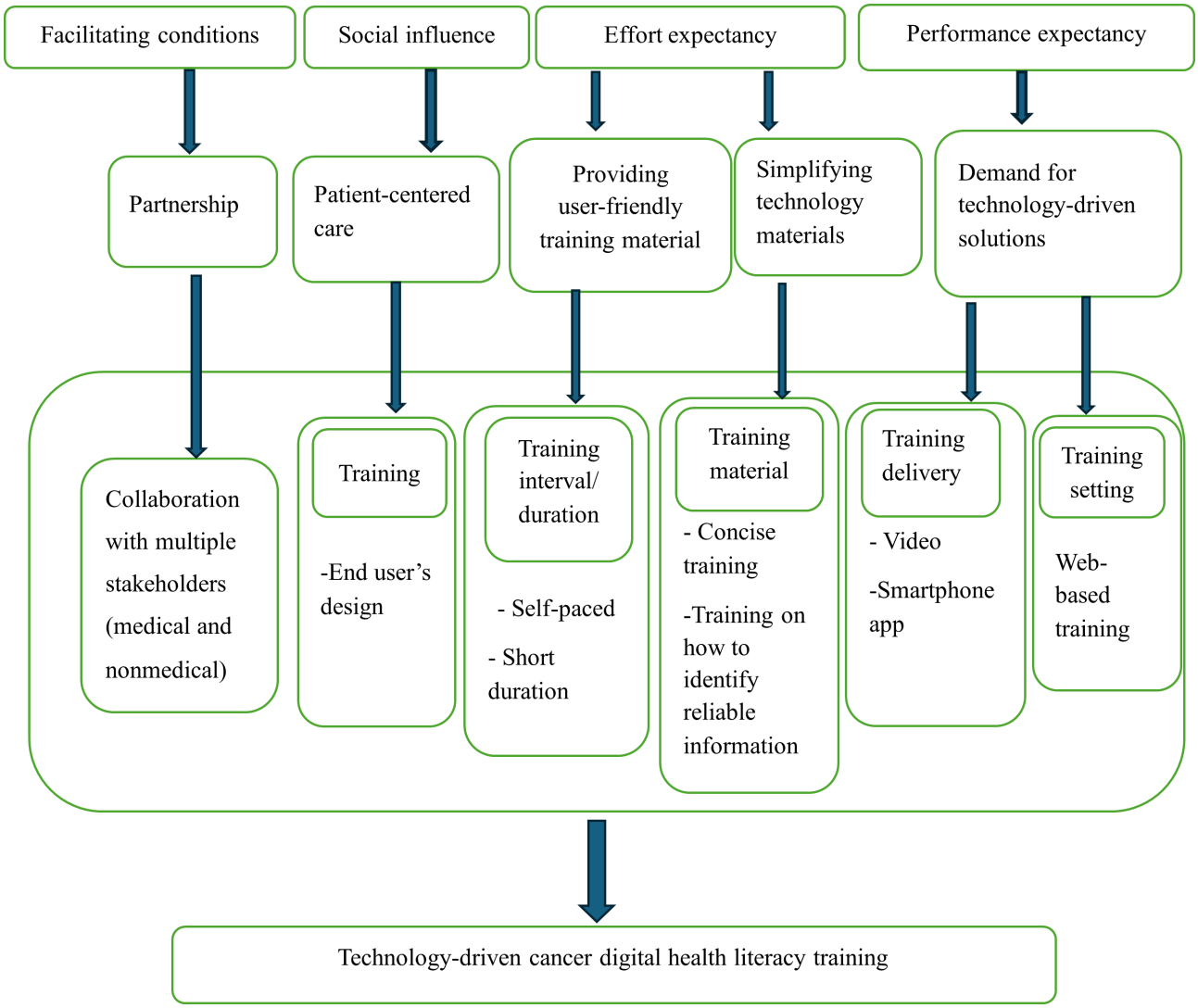
Patients with cancer, caregivers, and health care providers expressed a need for providing bilingual DHL training to deliver culturally appropriate material.

> *I think it would be worthwhile if the training is provided in Arabic and English.* [Patient 19]
>
> *Our native language in Saudi Arabia is Arabic, while some residents from different cultures speak English also. I believe providing an educational training in both languages would be useful for both.* [Caregiver 4]
>
> *Providing bilingual training in both Arabic and English may facilitate the use of the educational training by the end consumers from different cultures.* [Health care professional 6]

We integrated the findings into the commonly identified themes to generate a conceptual framework for cancer DHL training (Figure 1). The framework explores how the pre-identified themes and subthemes were mapped onto the 4 determinants of the UTAUT to generate technology-driven cancer DHL training.

XSL•FO
RenderX

Figure 1. Cancer digital health literacy conceptual framework based on the main findings of the study in conjunction with the Unified Theory of Acceptance and Use of Technology, including the useful training components derived from the thematic analysis and theory determinants.



Discussion

Principal Findings

We found that patients with cancer and caregivers were self-sufficient and motivated to engage with new technology to improve cancer DHL; however, health care professionals were concerned about the reliability of online information. We identified specific recommendations for creating cancer DHL training: Content should be concise, easy to understand, web-based, and delivered through technology-driven solutions; content should include training on how to identify reliable cancer information; patients would like to be involved in content creation; and training should be self-paced and of a short duration. Our findings demonstrated collaboration between medical and nonmedical stakeholders to facilitate DHL training adoption.

Comparison With Prior Work

Participants' Beliefs About Technology Adoption

Our findings indicated that patients with cancer and caregivers were self-sufficient and motivated to use novel technology to improve cancer health literacy. Previous studies in Saudi Arabia reported that most participants prefer to use the internet to search for health information [43-45]. A qualitative focus group study found that most participants recognized the benefits of using technology to manage their health care, despite expressing concerns about its use [16].

Car et al [46] identified several barriers to the use of digital health information, including a lack of motivation and cognitive skills to find and use information. The expectation that online health information will be useful is one of the strongest predictors of internet use [47], and a lack of skills was debated as one possible reason for not using the internet [46]. Specifically, potential users may lack the skills to effectively locate, evaluate, and use online health information [48]. Motivation was reported as an essential step in line with education for DHL training adoption [46].

Participants' Preferences Considering DHL Training Adoption

Our findings indicated that most participants preferred web-based DHL training. This approach has been reported in previous DHL interventions [49-53]; however, those interventions were not microlearning-based. An online learning

XSL•FO
RenderX

environment has certain limitations; as such, learning materials should be as explicit and purposeful as possible, thus microdesigned. Microlearning could leverage the use of technology to deliver tailored DHL training but not a technology-dependent format of learning [54].

Our findings indicated that the participants preferred to receive DHL training through technology-driven solutions, namely smartphone apps and videos. A study in Saudi Arabia used an educational video to improve participants' ability to detect health-related misinformation in WhatsApp messages [55]. Nokes and Reyes [56] targeted patients with HIV through an educational video about identifying reliable internet sites. As a cost-effective approach, mobile health apps can also appropriately deliver interventions to patients with cancer [57] and patients with chronic conditions [58]. The "SUCCESS app" is a DHL cross-platform application developed to support Australian adults with chronic kidney failure to improve patients' critical health literacy skills [59]. The "WebChoice" application was developed to provide reliable and useful links and resources for patients with cancer [60].

Effective microdesigned videos and health apps have recently been reported in health professional education studies [61-63]. However, to our knowledge, no microdesigned DHL apps or videos targeting patients with cancer have been developed yet. More research is needed to understand participants' views on how smartphone apps and videos could be microdesigned to improve DHL for patients with cancer.

We noted that the preferred style for information display for patients with cancer was a bite-sized format. Buchem and Hamelmann [64] contended that large bundles of information on the internet are often ignored, whereas small pieces of the whole are tagged and linked in ways that create new patterns, ideas, and meaning. Microlearning, therefore, has the capacity to present information in the most common way people learn today [65]. The individual learning units had a clear, focused message that reinforced two of the aims of microlearning: reducing cognitive load on working memory [66] and adapting to the pressures of a limited attention span [67].

Our paper identified demand for training on how to identify reliable cancer information, including learning to identify high-quality cancer websites and tips for evaluating online information. Both approaches were observed in prior DHL interventions and were reported to be effective in a systematic review [17]. A DHL workshop introduced participants to cancer websites of various top-level domains and web pages accredited by the Health on the Net Foundation [68]. A DHL tutorial also introduced participants to the MedlinePlus.gov website to search for health information [69]. Previous interventions have used several tools to teach people how to evaluate online information, including the SEEK (Source, Evidence, Explanation, and Knowledge) criteria [70], the DISCERN critical appraisal tool [49], and the PILOT (purpose, information, links, originator, timeliness) criteria [71]. However, Bernstam et al [72] reported that commonly cited website quality criteria are not effective at identifying inaccurate online information about breast cancer. Future research could focus on using the quality criteria to assess their effectiveness in evaluating other cancer web pages.

Health care professionals in our study identified a need to increase public awareness of common cancer misconceptions. A cross-sectional study conducted in Saudi Arabia explored the 7 commonly identified cancer misconceptions in Saudi Arabia [45]. However, the results of the study are not available in lay language. The next essential step is translating and disseminating the research findings into lay language, which could be facilitated through technology-driven DHL interventions.

Our findings indicated that the participants preferred self-paced, short-duration DHL training. This approach was seen in prior DHL interventions [56,70]. However, none of those interventions fit within the microlearning time frame. Unlike formal learning, the time spent on microlearning ranges from a few seconds to 15 minutes [64]. With microlearning's ability to space out the learning process over time, a more extended period is often required to learn the content, thereby leveraging retrieval spacing effects to improve information retention and support perpetual, sustainable lifelong learning [65].

Our findings indicated that cancer patients were interested in being involved in the design of cancer DHL training. Mohamed et al [17] reported using this approach in a few studies and considered it a facilitator of intervention use by end consumers. Kemp et al [18] called for using participatory design principles to address DHL in cancer care. Microlearning focuses on users' social interactions to drive content creation and can transform consumers into producers (prosumers) [64]. We identified a small body of literature focusing on users' feedback in the design of microlearning interventions in cancer care [73]. Future research could focus on developing microlearning-based DHL interventions from end users' perspectives.

Several educational approaches have been trialed to improve consumers' DHL, such as didactic lessons [71,74] and collaborative learning [75]. However, driven by the patient-centered care concept and the era of rapidly advancing technologies, Lee et al [76] called for the use of student-centered pedagogical approaches to educate health consumers. One such innovative pedagogy that could meet our participants' demand for bite-size, discrete learning bundles is microlearning [77].

Facilitating Conditions for DHL Training Adoption

In our study, the participants called for a collaborative approach between medical and nonmedical staff to deliver DHL training. Previous research reported the delivery of several DHL interventions through the collaboration of multiple stakeholders [53,56,68,74]. Most of those interventions were reported to be effective in a systematic review study [17]. Partnerships have become increasingly relevant in the delivery of health care services over the last few decades [78]. For instance, a 2012 Institute of Medicine report listed calls for enhanced collaboration among the public, private, health care, and non-health care sectors to improve chronic disease prevention [79]. Other studies have echoed recommendations for cross‐sector collaboration to address social determinants of health [80,81]. Recently, the infosphere has been recognized by Morley et al [82] as a social determinant of health, defining it as the entire information environment that influences health-related behaviors and outcomes, alongside socioeconomic status, income, education, age, race, ethnicity, and gender.

XSL•FO
RenderX

Strengths and Limitations

This study has several limitations. We only explored the perspectives of patients with cancer, caregivers, and oncology health care professionals. Our paper did not consider the opinions of policymakers, technology advisers, and other health care professionals (nurses and health educators). Future research can include diverse stakeholder types. Only 6 caregivers were included in our study. Future research could include more caregivers to compare their experiences with those of patients with cancer. The study focused on participants from a single hospital, limiting generalizability to broader populations. Future research should include a diverse sample across multiple health care settings. The study participants were all deliberate internet users, selected to understand their experiences with online navigation needs. Addressing noninternet users and considering their preferences for DHL intervention were not considered in our paper. Implementing digital health in cancer care must address the variability of DHL in recipients to enhance digital inclusion and digital health equity. The Framework Method used for thematic analysis was time-consuming and resource-intensive and requires a high level of training to be used successfully within a multidisciplinary team.

The UTAUT was found to be a good fit for our study, as it captured the content discussed in the interviews and helped explore which technology-driven DHL training works well from the perspectives of end users and health care providers. To our knowledge, our study is the first to apply the UTAUT to propose a microlearning-based DHL training program. However, we did not examine the effect of the UTAUT moderators (age, gender, experiences, voluntariness of use) on participants' intention to adopt technology. Future research could use a health literacy lens to examine the interactions among those factors and participants' technology adoption.

Implications for Practice

To ensure patient-centered care, it is essential to involve the end users in designing future DHL interventions. Collaborating with multiple stakeholders across relevant fields could facilitate shared decision-making. Education, combined with targeted design, can help effectively implement digital health in cancer care.

Driven by the issue raised in our paper regarding the digital health literacy gap for patients with cancer, microlearning could help address the need to learn infrequently used or new skills, thereby improving patient safety [64]. A previous feasibility study of microlearning to deliver information on chemotherapy side effects to patients with lung cancer demonstrated improvements in participants' cancer literacy and self-management coping strategies [73]. However, to our knowledge, our paper is the first to propose microlearning as a subscription model for designing a DHL intervention. Future work should focus on creating a prototype and examining its feasibility for short-term implementation.

Despite using microlearning as an instructional educational strategy, it may eventually be superseded by emerging innovations such as nanolearning, metaverse learning, or artificial intelligence–assisted learning, which promise opportunities to improve the quality and accessibility of education and to facilitate lifelong learning.

Conclusion

Patients with cancer and caregivers were motivated to use modern technology-driven solutions to improve cancer DHL. The participants' preferences regarding DHL training aligned with the innovative microlearning pedagogy. Microlearning could leverage technology to deliver tailored digital health literacy training for patients with cancer. Collaboration with multiple medical and nonmedical stakeholders could facilitate the delivery of cancer DHL training.

Future work should focus on designing and assessing the feasibility of implementing a microlearning-based DHL training program that involves end users and diverse stakeholders.

Acknowledgments

We are grateful to the journal editors and anonymous reviewers for their helpful comments and suggestions. We thank all patients and caregivers and acknowledge all health care professionals who participated in the study. No generative artificial intelligence, including tools like ChatGPT or other language models, was used in any portion of the manuscript generation. All content was created and written by the authors.

Funding

No external financial support nor grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The research data are unavailable to access because the data include sensitive and confidential information.

Authors' Contributions

Conceptualization: HM (lead), DK (supporting)
Data curation: HM
Formal analysis: HM

XSL•FO
RenderX

Investigation: HM
Methodology HM
Project administration: HM
Software: HM
Supervision: DK (lead), TA (supporting), JS (supporting)
Translation: HM (lead), MM (supporting)
Validation: HM
Visualization: HM
Writing – original draft: HM
Writing – review & editing: HM (lead), DK (supporting), MM (supporting), TA (supporting), JS (supporting)
HM (email: hind-awdw@hotmail.com; alternative email: 20138148@studentmail.ul.ie)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Arabic and English versions of the interview guide.
[DOCX File, 24 KB - cancer_v12i1e77738_app1.docx]

Checklist 1
COREQ (Consolidated Criteria for Reporting qualitative research) checklist.
[PDF File, 541 KB - cancer_v12i1e77738_app2.pdf]

References

1. Vance K, Howe W, Dellavalle RP. Social internet sites as a source of public health information. Dermatol Clin 2009 Apr;27(2):133-136. [doi: 10.1016/j.det.2008.11.010] [Medline: 19254656]
2. Al-Kahtani N, Alruwaie S, Al-Zahrani BM, et al. Digital health transformation in Saudi Arabia: a cross-sectional analysis using Healthcare Information and Management Systems Society' digital health indicators. Digit Health 2022;8:20552076221117742. [doi: 10.1177/20552076221117742] [Medline: 35959196]
3. Rahman R, Al-Borie HM. Strengthening the Saudi Arabian healthcare system: role of Vision 2030. Int J Healthc Manag 2021 Oct 2;14(4):1483-1491. [doi: 10.1080/20479700.2020.1788334]
4. AlMuammar SA, Noorsaeed AS, Alafif RA, Kamal YF, Daghistani GM. The use of internet and social media for health information and its consequences among the population in Saudi Arabia. Cureus 2021 Sep;13(9):e18338. [doi: 10.7759/cureus.18338] [Medline: 34722089]
5. Fujisawa D, Hagiwara N. Cancer stigma and its health consequences. Curr Breast Cancer Rep 2015 Sep;7(3):143-150. [doi: 10.1007/s12609-015-0185-0]
6. Al-Azri M, Al-Awisi H, Al-Rasbi S, et al. Psychosocial impact of breast cancer diagnosis among Omani women. Oman Med J 2014 Nov;29(6):437-444. [doi: 10.5001/omj.2014.115] [Medline: 25584162]
7. Fareed N, Swoboda CM, Jonnalagadda P, Huerta TR. Persistent digital divide in health-related internet use among cancer survivors: findings from the Health Information National Trends Survey, 2003-2018. J Cancer Surviv 2021 Feb;15(1):87-98. [doi: 10.1007/s11764-020-00913-8] [Medline: 32671557]
8. Penedo FJ, Natori A, Fleszar-Pavlovic SE, et al. Factors associated with unmet supportive care needs and emergency department visits and hospitalizations in ambulatory oncology. JAMA Netw Open 2023 Jun 1;6(6):e2319352. [doi: 10.1001/jamanetworkopen.2023.19352] [Medline: 37342038]
9. Yoo SH, Sung JH, Lee K, et al. The needs for digital health and eHeath literacy of cancer patients, caregivers, and healthcare providers: a multicenter, descriptive correlational study. Eur J Oncol Nurs 2024 Jun;70:102581. [doi: 10.1016/j.ejon.2024.102581] [Medline: 38749385]
10. Sanchiz M, Amadieu F, Fu WT, Chevalier A. Does pre-activating domain knowledge foster elaborated online information search strategies? Comparisons between young and old web user adults. Appl Ergon 2019 Feb;75(201-213):201-213. [doi: 10.1016/j.apergo.2018.10.010] [Medline: 30509528]
11. Ahles TA, Root JC. Cognitive effects of cancer and cancer treatments. Annu Rev Clin Psychol 2018 May 7;14:425-451. [doi: 10.1146/annurev-clinpsy-050817-084903] [Medline: 29345974]
12. Mohamed H, O'Malley L, Kelly D. An infodemiology study on exploring the quality and reliability of colorectal cancer immunotherapy information. Digit Health 2023;9:20552076231205286. [doi: 10.1177/20552076231205286] [Medline: 37808242]
13. Almoajel A, Alshamrani S, Alyabsi M. The relationship between e-Health literacy and breast cancer literacy among Saudi women. Front Public Health 2022;10:841102. [doi: 10.3389/fpubh.2022.841102] [Medline: 35462821]

XSL•FO
RenderX

14. Alhur A, Alhur A, Alshammari M, et al. Digital health literacy and web-based health information-seeking behaviors in the Saudi Arabian population. Cureus 2023 Dec;15(12):e51125. [doi: 10.7759/cureus.51125] [Medline: 38274905]
15. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health literacy and health information technology adoption: the potential for a new digital divide. J Med Internet Res 2016 Oct 4;18(10):e264. [doi: 10.2196/jmir.6349] [Medline: 27702738]
16. Zarcadoolas C, Vaughon WL, Czaja SJ, Levy J, Rockoff ML. Consumers' perceptions of patient-accessible electronic medical records. J Med Internet Res 2013 Aug 26;15(8):e168. [doi: 10.2196/jmir.2507] [Medline: 23978618]
17. Mohamed H, Kittle E, Nour N, et al. An integrative systematic review on interventions to improve layperson's ability to identify trustworthy digital health information. PLOS Digit Health 2024 Oct;3(10):e0000638. [doi: 10.1371/journal.pdig.0000638] [Medline: 39453891]
18. Kemp E, Trigg J, Beatty L, et al. Health literacy, digital health literacy and the implementation of digital health technologies in cancer care: the need for a strategic approach. Health Promot J Austr 2021 Feb;32 Suppl 1:104-114. [doi: 10.1002/hpja.387] [Medline: 32681656]
19. Rising KL, Guth A, Gentsch AT, et al. Development and preliminary validation of a screener for digital health readiness. JAMA Netw Open 2024 Sep 3;7(9):e2432718. [doi: 10.1001/jamanetworkopen.2024.32718] [Medline: 39254972]
20. Batalden M, Batalden P, Margolis P, et al. Coproduction of healthcare service. BMJ Qual Saf 2016 Jul;25(7):509-517. [doi: 10.1136/bmjqs-2015-004315] [Medline: 26376674]
21. McColl-Kennedy JR, Snyder H, Elg M, et al. The changing role of the health care customer: review, synthesis and research agenda. JOSM 2017 Mar 20;28(1):2-33. [doi: 10.1108/JOSM-01-2016-0018]
22. Kim SH, Sung JH, Yoo SH, et al. Effects of digital self-management symptom interventions on symptom outcomes in adult cancer patients: a systematic review and meta-analysis. Eur J Oncol Nurs 2023 Oct;66:102404. [doi: 10.1016/j.ejon.2023.102404] [Medline: 37517339]
23. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HWW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017 Dec;11(6):704-719. [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]
24. Shaffer KM, Turner KL, Siwik C, et al. Digital health and telehealth in cancer care: a scoping review of reviews. Lancet Digit Health 2023 May;5(5):e316-e327. [doi: 10.1016/S2589-7500(23)00049-3] [Medline: 37100545]
25. Nutbeam D. Getting evidence into policy and practice to address health inequalities. Health Promot Int 2004 Jun;19(2):137-140. [doi: 10.1093/heapro/dah201] [Medline: 15128705]
26. Eccles MP, Armstrong D, Baker R, et al. An implementation research agenda. Implementation Sci 2009 Dec;4(1):18. [doi: 10.1186/1748-5908-4-18]
27. Lee H, Oldewage-Theron W. Process evaluation of a theory-informed eHealth intervention through "Healthy Online Parental Education (HOPE)" study. J Nutr Educ Behav 2023 Jul;55(7):86. [doi: 10.1016/j.jneb.2023.05.187]
28. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
29. Scur G, da Silva AVD, Mattos CA, Gonçalves RF. Analysis of IoT adoption for vegetable crop cultivation: multiple case studies. Technol Forecast Soc Change 2023 Jun;191:122452. [doi: 10.1016/j.techfore.2023.122452]
30. Al-Sharafi MA, Al-Emran M, Arpaci I, et al. Generation Z use of artificial intelligence products and its impact on environmental sustainability: a cross-cultural comparison. Comput Human Behav 2023 Jun;143:107708. [doi: 10.1016/j.chb.2023.107708]
31. Gu D, Khan S, Khan IU, et al. Assessing the adoption of e-Health technology in a developing country: an extension of the UTAUT model. Sage Open 2021 Jul;11(3). [doi: 10.1177/21582440211027565]
32. Cobelli N, Cassia F, Donvito R. Pharmacists' attitudes and intention to adopt telemedicine: integrating the market-orientation paradigm and the UTAUT. Technol Forecast Soc Change 2023 Nov;196:122871. [doi: 10.1016/j.techfore.2023.122871]
33. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
34. Downe-Wamboldt B. Content analysis: method, applications, and issues. Health Care Women Int 1992;13(3):313-321. [doi: 10.1080/07399339209516006] [Medline: 1399871]
35. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Community Health 2019;7(2):e000057. [doi: 10.1136/fmch-2018-000057] [Medline: 32148704]
36. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544. [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
37. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 2017 Mar;27(4):591-608. [doi: 10.1177/1049732316665344] [Medline: 27670770]
38. Naeem M, Ozuem W, Howell K, Ranfagni S. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. Int J Qual Methods 2023 Oct;22. [doi: 10.1177/16094069231205789]

XSL•FO
RenderX

39. Giles EL, Harrison SL. Reflecting on the importance of theory-informed qualitative research in people with chronic respiratory disease and their carers. Chron Respir Dis 2023;20:14799731231185128. [doi: 10.1177/14799731231185128] [Medline: 37344979]
40. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13(117):117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
41. O'Connor C, Joffe H. Intercoder reliability in qualitative research: debates and practical guidelines. Int J Qual Methods 2020 Jan 1;19. [doi: 10.1177/1609406919899220]
42. McKim C. Meaningful member-checking: a structured approach to member-checking. Am J Qualitative Res 2023;7(2):41-52. [doi: 10.29333/ajqr/12973]
43. Bahkali S, Alfurih S, Aldremly M, Alzayyat M, Alsurimi K, Househ M. The prevalence of internet and social media based medication information seeking behavior in Saudi Arabia. Stud Health Technol Inform 2016;226:275-278. [doi: 10.3233/978-1-61499-664-4-275] [Medline: 27350524]
44. Alhur A. Exploring Saudi Arabia individuals' attitudes toward electronic personal health records. JCSTS 2022;4(1):80-87. [doi: 10.32996/jcsts.2022.4.1.10]
45. Marouf A, Tayeb R, Alshehri GD, et al. Public perception of common cancer misconceptions: a nationwide cross-sectional survey and analysis of over 3500 participants in Saudi Arabia. J Family Med Prim Care 2023 Jun;12(6):1125-1132. [doi: 10.4103/jfmpc.jfmpc_1753_22] [Medline: 37636192]
46. Car J, Lang B, Colledge A, Ung C, Majeed A. Interventions for enhancing consumers' online health literacy. Cochrane Database Syst Rev 2011 Jun 15;2011(6):CD007092. [doi: 10.1002/14651858.CD007092.pub2] [Medline: 21678364]
47. Mead N, Varnam R, Rogers A, Roland M. What predicts patients' interest in the Internet as a health resource in primary care in England? J Health Serv Res Policy 2003 Jan;8(1):33-39. [doi: 10.1177/135581960300800108] [Medline: 12683432]
48. Gray NJ, Klein JD, Noyce PR, Sesselberg TS, Cantrill JA. The internet: a window on adolescent health literacy. J Adolesc Health 2005 Sep;37(3):243. [doi: 10.1016/j.jadohealth.2004.08.023] [Medline: 16109345]
49. Austvoll-Dahlgren A, Bjørndal A, Odgaard-Jensen J, Helseth S. Evaluation of a web portal for improving public access to evidence-based health information and health literacy skills: a pragmatic trial. PLoS ONE 2012;7(5):e37715. [doi: 10.1371/journal.pone.0037715] [Medline: 22701531]
50. Pennycook G, McPhetres J, Zhang Y, Lu JG, Rand DG. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. Psychol Sci 2020 Jul;31(7):770-780. [doi: 10.1177/0956797620939054] [Medline: 32603243]
51. Mörelius E, Robinson S, Arabiat D, Whitehead L. Digital interventions to improve health literacy among parents of children aged 0 to 12 years with a health condition: systematic review. J Med Internet Res 2021 Dec 22;23(12):e31665. [doi: 10.2196/31665] [Medline: 34941559]
52. Xie B. Improving older adults' e-health literacy through computer training using NIH online resources. Libr Inf Sci Res 2012 Jan 1;34(1):63-71. [doi: 10.1016/j.lisr.2011.07.006] [Medline: 22639488]
53. Evans P, Aungst TD, Massey C, Bartlett D. Expanding clinical and information services to the ambulatory older adult through community outreach programs. Consult Pharm 2015 Jan;30(1):31-37. [doi: 10.4140/TCP.n.2015.31] [Medline: 25591029]
54. Sankaranarayanan R, Leung J, Abramenka-Lachheb V, Seo G, Lachheb A. Microlearning in diverse contexts: a bibliometric analysis. TechTrends 2023;67(2):260-276. [doi: 10.1007/s11528-022-00794-x] [Medline: 36254216]
55. Alsaad E, AlDossary S. Educational video intervention to improve health misinformation identification on WhatsApp among Saudi Arabian population: pre-post intervention study. JMIR Form Res 2024 Jan 17;8:e50211. [doi: 10.2196/50211] [Medline: 38231563]
56. Nokes KM, Reyes DM. Internet use for health-related information: self-care agency of lower income persons living with HIV/AIDS. Nurs Sci Q 2020 Jul;33(3):234-239. [doi: 10.1177/0894318420920605] [Medline: 32605479]
57. Kim H, Goldsmith JV, Sengupta S, et al. Mobile health application and e-Health literacy: opportunities and concerns for cancer patients and caregivers. J Canc Educ 2019 Feb;34(1):3-8. [doi: 10.1007/s13187-017-1293-5]
58. Mosa ASM, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012 Jul 10;12(1):67. [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]
59. Muscat DM, Lambert K, Shepherd H, et al. Supporting patients to be involved in decisions about their health and care: development of a best practice health literacy app for Australian adults living with chronic kidney disease. Health Prom J of Aust 2021 Feb;32(S1):115-127 [FREE Full text] [doi: 10.1002/hpja.416]
60. Ruland CM, Jeneson A, Andersen T, et al. Designing tailored Internet support to assist cancer patients in illness management. AMIA Annu Symp Proc 2007 Oct 11;2007:635-639. [Medline: 18693913]
61. Román-Sánchez D, De-La-Fuente-Rodríguez JM, Paramio A, Paramio-Cuevas JC, Lepiani-Díaz I, López-Millan MR. Evaluating satisfaction with teaching innovation, its relationship to academic performance and the application of a video-based microlearning. Nurs Open 2023 Sep;10(9):6067-6077. [doi: 10.1002/nop2.1828] [Medline: 37221960]

XSL•FO
RenderX

62. Flornoy-Guédon A, Fonzo-Christe C, Meier E, et al. Development and evaluation of a blended learning training programme for pharmacy technicians' continuing education. Eur J Hosp Pharm 2024 Aug 22;31(5):403-408. [doi: 10.1136/ejhpharm-2022-003679] [Medline: 36898765]
63. Zolfaghari M, Shirzadi S, Motamed M. Using a mobile application for psychiatry training in medical students: a quasi-experimental study. Australas Psychiatry 2023 Jun;31(3):389-394. [doi: 10.1177/10398562231159509] [Medline: 36840465]
64. Buchem I, Hamelmann H. Microlearning: a strategy for ongoing professional development. Microcontent and Microlearning 2010;21(7):1-5 [FREE Full text]
65. Carter JW, Youssef-Morgan C. Psychological capital development effectiveness of face-to-face, online, and micro-learning interventions. Educ Inf Technol (Dordr) 2022;27(5):6553-6575. [doi: 10.1007/s10639-021-10824-5] [Medline: 35068989]
66. Taylor A, Hung W. The effects of microlearning: a scoping review. Education Tech Research Dev 2022;70:363-395. [doi: 10.1007/s11423-022-10084-1]
67. Sun G, Cui T, Yong J, Shen J, Chen S. MLaaS: a cloud-based system for delivering adaptive micro learning in mobile MOOC learning. IEEE Trans Serv Comput 2018;11(2):292-305. [doi: 10.1109/TSC.2015.2473854]
68. Hoffman-Goetz L, Friedman DB, Celestine A. Evaluation of a public library workshop: Teaching older adults how to search the internet for reliable cancer information. J Consum Health Internet 2006;10(3):29-43. [doi: 10.1300/J381v10n03_03]
69. De Main AS, Xie B, Shiroma K, Yeh T, Davis N, Han X. Assessing the effects of eHealth tutorials on older adults' eHealth literacy. J Appl Gerontol 2022 Jul;41(7):1675-1685. [doi: 10.1177/07334648221088281] [Medline: 35466732]
70. Kammerer Y, Amann DG, Gerjets P. When adults without university education search the internet for health information: the roles of Internet-specific epistemic beliefs and a source evaluation intervention. Comput Human Behav 2015 Jul;48:297-309. [doi: 10.1016/j.chb.2015.01.045]
71. Kalichman SC, Cherry C, Cain D, et al. Internet-based health information consumer skills intervention for people living with HIV/AIDS. J Consult Clin Psychol 2006 Jun;74(3):545-554. [doi: 10.1037/0022-006X.74.3.545] [Medline: 16822111]
72. Bernstam EV, Walji MF, Sagaram S, Sagaram D, Johnson CW, Meric-Bernstam F. Commonly cited website quality criteria are not effective at identifying inaccurate online information about breast cancer. Cancer 2008 Mar 15;112(6):1206-1213. [doi: 10.1002/cncr.23308] [Medline: 18266210]
73. Janssen A, Shah K, Rabbets M, et al. Feasibility of microlearning for improving the self-efficacy of cancer patients managing side effects of chemotherapy. J Cancer Educ 2023 Oct;38(5):1697-1709. [doi: 10.1007/s13187-023-02324-6] [Medline: 37452227]
74. Gray K, Elliott K, Wale J. A community education initiative to improve using online health information: participation and impact. Inform Health Soc Care 2013 Sep;38(3):171-181. [doi: 10.3109/17538157.2012.705201] [Medline: 23324099]
75. Xie B. Older adults, e-health literacy, and collaborative learning: an experimental study. J Am Soc Inf Sci 2011 May;62(5):933-946. [doi: 10.1002/asi.21507]
76. Lee K, Hoti K, Hughes JD, Emmerton LM. Interventions to assist health consumers to find reliable online health information: a comprehensive review. PLoS ONE 2014;9(4):e94186. [doi: 10.1371/journal.pone.0094186]
77. Emerson LC, Berge ZL. Microlearning: knowledge management applications and competency-based training in the workplace. Knowledge Management & E-Learning 2018;10(2). [doi: 10.34105/j.kmel.2018.10.008]
78. Henderson S, Wagner JL, Gosdin MM, et al. Complexity in partnerships: a qualitative examination of collaborative depression care in primary care clinics and community-based organisations in California, United States. Health Soc Care Community 2020 Jul;28(4):1199-1208. [doi: 10.1111/hsc.12953] [Medline: 32052531]
79. Harris JR, Wallace RB. The Institute of Medicine's new report on living well with chronic illness. Prev Chronic Dis 2012;9:E148. [doi: 10.5888/pcd9.120126] [Medline: 22995102]
80. Miranda J, Ong MK, Jones L, et al. Community-partnered evaluation of depression services for clients of community-based agencies in under-resourced communities in Los Angeles. J GEN INTERN MED 2013 Oct;28(10):1279-1287. [doi: 10.1007/s11606-013-2480-7] [Medline: 23670566]
81. Predmore Z, Hatef E, Weiner JP. Integrating social and behavioral determinants of health into population health analytics: a conceptual framework and suggested road map. Popul Health Manag 2019 Dec;22(6):488-494. [doi: 10.1089/pop.2018.0151] [Medline: 30864884]
82. Morley J, Cowls J, Taddeo M, Floridi L. Public health in the information age: recognizing the infosphere as a social determinant of health. J Med Internet Res 2020 Aug 3;22(8):e19311. [doi: 10.2196/19311] [Medline: 32648850]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
DHL: digital health literacy
KFSH: King Fahad Specialist Hospital
PILOT: purpose, information, links, originator, timeliness
SEEK: Source, Evidence, Explanation, and Knowledge
UTAUT: Unified Theory of Acceptance and Use of Technology

XSL•FO
RenderX

Edited by M Balcarras, N Cahill; submitted 19.May.2025; peer-reviewed by L Wreyford, M Katapodi, S Ye; revised version received 22.Jan.2026; accepted 22.Jan.2026; published 04.Mar.2026.

Please cite as:
Mohamed H, Alanzi T, Salsberg J, Mohamed M, Kelly D
Exploring the Readiness for Digital Health Literacy Transformation and Intervention Preferences From the Perspectives of Patients With Cancer, Caregivers, and Health Care Professionals: Qualitative Interview Study
JMIR Cancer 2026;12:e77738
URL: https://cancer.jmir.org/2026/1/e77738
doi:10.2196/77738

© Hind Mohamed, Turki Alanzi, Jon Salsberg, Mudathir Mohamed, Dervla Kelly. Originally published in JMIR Cancer (https://cancer.jmir.org), 4.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Real-World Evidence of Treatment Outcomes in Small Cell Lung Cancer: A Bayesian Mixed Effects and Competitive Risk Approach

Luca Marzano[1*], PhD; Asaf Dan[2,3*], MD; Adam S Darwich[1], PhD; Luigi De Petris[2,3], MD, PhD; Salomon Tendler[4], MD, PhD; Rolf Lewensohn[2,3], MD, PhD; Jayanth Raghothama[1], PhD; Sebastiaan Meijer[1], PhD

[1]KTH Royal Institute of Technology, Hälsovägen 11C, Huddinge, Stockholm, Sweden

[2]Department of Oncology and Pathology, Karolisnka Institute, Stockholm, Sweden

[3]Thoracic Oncology Center, Karolinska University Hospital, Stockholm, Sweden

[4]Department of Medicine, Memorial Sloan Kettering Cancer Center, Ner York, NY, United States

[*]these authors contributed equally

Corresponding Author:
Adam S Darwich, PhD
KTH Royal Institute of Technology, Hälsovägen 11C, Huddinge, Stockholm, Sweden

Abstract

Background: Small cell lung cancer (SCLC) is a challenging disease to treat due to rapid progression, development of chemoresistance, and discrepancies in outcomes between real-world data and clinical trials. There is a lack of comprehensive analyses in other studies with regard to intermediate events and the treatment process, such as treatment decisions, progression of disease, and the occurrence of adverse events (AEs) over time.

Objective: The aim of this study was to apply advanced statistical methods to a longitudinal SCLC dataset in order to identify factors of importance for the risk of AEs and for survival.

Methods: Treatment pathways of 421 patients with SCLC collected from Karolinska University Hospital, located in Stockholm, Sweden, between 2016 and 2022, were analyzed with data-driven modeling. The analysis focused on the impact of dose adjustment on AEs, including neutropenia, by estimating odds ratios (ORs) using Bayesian mixed effects modeling. Covariates' effects on Eastern Cooperative Oncology Group performance status (ECOG PS) deterioration and early discontinuation of chemotherapy with cause-specific hazard ratios (csHR) were explored using competitive risk models. This approach was applied to patient cohorts receiving combinatorial first-line platinum/etoposide and second-line platinum/etoposide or platinum/irinotecan.

Results: At the end of the first-line treatment, most patients exhibited tumor regression (n=167). Patients with neutropenia had longer overall survival (hazard ratio 0.70, 95% CI 0.53-0.92). Higher etoposide dose levels were associated with subsequent occurrences of AEs (OR 5.97, 95% CI 1.41-30.5) and neutropenia (OR 3.55, 95% CI 1.03-13.3). Dose adjustment did not affect overall survival if the patient completed the 4-dose regimen treatment. With regard to second-line therapy, fewer patients completed 4 treatment cycles, and the most common reason for early discontinuation was tumor progression (n=72, 58%). Male patients (n=118) experienced fewer AEs and better first-line treatment response compared to females (csHR 0.51, 95% CI 0.25-0.90). High-risk patients (defined as ECOG PS 2 - 3 or age >75 years) with early discontinuation of therapy had survival outcomes similar to those who did not receive any therapy.

Conclusions: Our results indicate that first-line therapies may benefit from more individualized dosing strategies. It would also be beneficial to assess the risk-benefit of treating specific subgroups, including patients receiving second-line therapy. Real-world data proved beneficial for studying therapy response and risk-benefit of treating patient groups that are underrepresented in clinical trials.

(JMIR Cancer 2026;12:e84042) doi:10.2196/84042

KEYWORDS

real-world data; small cell lung cancer; SCLC; real-world evidence; adverse effects; survival; toxicity; chemotherapy

Introduction

Small cell lung cancer (SCLC) is an aggressive and challenging disease to treat, mainly due to the lack of knowledge surrounding chemoresistance mechanisms [1,2]. Standard SCLC treatment consists of defined treatment lines, each comprising 4-6 cycles of platinum-based chemotherapy (cisplatin or carboplatin) with etoposide (PE) or irinotecan (IP). For patients with cancer stages I-III, this is often combined with concurrent thoracic radiotherapy. In stage IV cancer, immunotherapy with checkpoint inhibitors can be combined with chemotherapy, though with a modest median overall survival gain of 2 - 3

XSL•FO
RenderX

months in randomized controlled trials (RCTs) and with even smaller survival benefits reported in real-world studies [3,4]. However, most patients with SCLC experience relapse within 6 months after initial treatment, necessitating further therapy, though with limited survival. Real-world studies further indicate that outcomes beyond second-line therapy are poor [5], and the 5-year overall survival for patients with SCLC remains below 10% [6].

In the setting of disease relapse, the distinction between sensitive-relapse and resistance-relapse plays a crucial role in predicting survival. Patients with sensitive disease, defined by relapse occurring more than 90 days after the completion of first-line platinum-based chemotherapy, might be rechallenged with the same chemotherapy regimen administered as first-line and demonstrate significantly better survival outcomes compared to those with resistance disease, where relapse occurs within 90 days [7,8]. In this latter case, the most common approach is to change to other chemotherapy agents if the patient is still in acceptable general condition.

The rapid progression of SCLC often makes it difficult to manage [9], which results in discrepancies in observed outcomes between clinical trials and real-world patients [10]. The knowledge gap between clinical trials and the real-world underrepresented patient populations can be overcome by treatment adjustments based on clinical experience, including choice of chemotherapy agents, dose adjustments, and end-of-life care [9,11,12]. The focus of previous studies has mainly been on survival outcomes, often overlooking patients who received subsequent treatment lines after first-line chemotherapy. Further, there is a lack of studies that analyze events that occur over time in the treatment process [9,13-15].

Real-world data, here defined as health care data produced for the primary purpose of treatment in a clinical setting, offer an opportunity to retrospectively explore the impact of treatment decisions and pathways on patient outcomes, given the variability in the real-world patient population [16,17]. As such, real-world data have the potential to inform the individualization of treatments for specific subpopulations [18,19], such as patients with poor Eastern Cooperative Oncology Group performance status (ECOG PS), older patients, and patients with brain metastases who tend to be underrepresented in clinical trials [10].

However, evidence generation from real-world data is challenging because of complex clinical workflows and issues related to its secondary use, such as inherent biases in data collection and decision-making processes, and nonobservable effects [20,21]. In SCLC treatment, the nuances of dose adjustments, planned treatment pauses, assessment of response, and early discontinuation of treatment due to severe toxicity or health deterioration exemplify the intricacies involved [22,23].

The aim of this study is to describe the time dynamics of patient trajectories in a clinical patient with SCLC cohort in order to identify factors of importance for dose adjustments and survival outcomes. This approach necessitates an appropriate choice of data analysis [24]. While artificial intelligence and machine learning have shown much promise for analyzing real-world data, issues of transparency and interpretability pose a potential risk of propagating biases from the data, including not-captured information and operational confounders [10,25].

Here, the focus was on selecting transparent modeling approaches [25,26], allowing for clinical contextualization and interpretation [10,21]. Then, the analysis consisted of a combination of Bayesian [27-29] and mixed effects models [11,30,31] for studying dose adjustments and their impact on adverse events (AEs; Bayesian logistic regression [32]) and the grade of toxicity (Bayesian cumulative link ordinal regression [32]). Competitive risk Fine-Gray regression [33] was applied to compute cause-specific hazard ratios (csHR) and predict the time course of the therapy, ECOG PS deterioration, and early discontinuation of treatment, defined as discontinuation before completing 4 planned chemotherapy cycles due to AEs, or radiological and clinical signs of tumor progression.

Methods

Collection of Longitudinal Data

This study included a total of 421 patients diagnosed with SCLC at the Karolinska University Hospital in Stockholm, Sweden, between 2016 and 2022. Baseline patient characteristics such as age, sex, smoking status, TNM (tumor node metastasis) stage (*International Association for the Study of Lung Cancer, 8th Edition*), Charlson Comorbidity Index, brain metastasis (CNS), and previous cancer prior to SCLC were recorded. ECOG PS was recorded at baseline and at the start and cessation of each treatment line. Blood values, including hemoglobin, creatinine, sodium, lactate dehydrogenase (LDH), C-reactive protein (CRP), and albumin, were computed at baseline and at the start of each treatment line.

The specific drug regimen received by each patient was recorded. Body surface area (BSA) and estimated glomerular filtration rate (GFR) were recorded before the start of treatments. These two variables were used to compute the dose of chemotherapy administered for each patient (eg, GFR for carboplatin dosing using the Calvert formula [34] and BSA for etoposide and irinotecan; see Table S1 in Multimedia Appendix 1). The date and percentage of dose administered for each chemotherapy treatment were registered, including the reason for any modifications or adjustments to the administered doses, for the calculation of dose interval and intensity.

The reason for the dose reduction was recorded. AEs were classified by type and grade of severity of toxicity according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5. In this work, high toxicity was defined as AEs with CTCAE G3-4 [35].

Details on radiotherapy, including the treatment indication (curative, consolidating, as a separate treatment line, or palliative) and fractionation, were reported. At the end of each treatment line, the tumoral response was assessed: complete response, partial response, stable disease, or signs of tumor progression.

The overall survival (OS) was calculated from the date of diagnosis until the patient's death or the end of the study period, whichever came first. For each treatment line, progression-free

XSL•FO
RenderX

survival (PFS) was calculated from the date of the administration of the first dose until signs of progression or death.

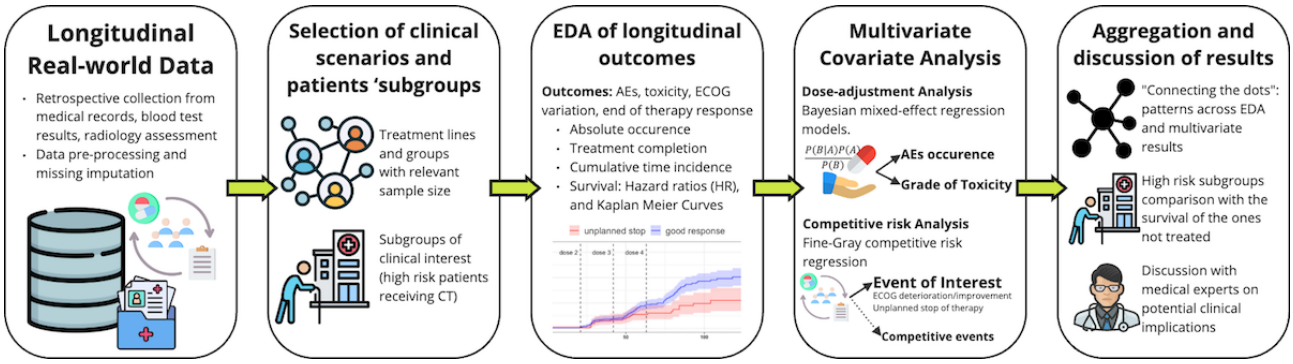
Analysis Description

Overview of the Real-World Data Pipeline

Several aspects of the treatment process were analyzed for each treatment line, including dose adjustments due to AEs, changes in ECOG PS from the start to the end of the treatment line, and early treatment discontinuation (stopping treatment before 4 cycles of chemotherapy due to signs of tumor progression, toxicity, deteriorating medical condition, or unwillingness of the patient to receive the treatment) versus no early discontinuation, defined as cessation of a treatment after 4 cycles of chemotherapy due to partial response or stable disease, followed by a pause.

The analytical pipeline (Figure 1) can be summarized in the following steps: selection of clinical scenarios and patient subgroups, exploratory data analysis of the longitudinal cohort, multivariate analysis of dose adjustments, multivariate analysis of competitive risks, analysis of high-risk subgroups, and aggregation and discussion of results.

Figure 1. Analytical pipeline description. AE: adverse event; ECOG: Eastern Cooperative Oncology Group; EDA: exploratory data analysis.



Selection of Clinical Scenarios and Patient Subgroups

Prior to commencing the analysis, the collected data were analyzed with a broad perspective to identify key subpopulations of interest. The cohort was explored in terms of baseline characteristics at the first-line treatment decision and the administered chemotherapy lines using summary statistics and Sankey flows for the subsequent treatment lines. Treatment lines with relevant sample size (cohort ≥20 individuals [36]) were detected for the next analysis steps. Moreover, special subgroups of patients receiving treatment but with highly fragile conditions (underrepresented in RCTs) or subjected to a high risk of poor survival outcomes were selected. The continuous covariates with high skewness were log-transformed before starting the analysis. Furthermore, in case of missing values, for each treatment line analysis, values were imputed using the machine learning algorithm MissForest [37]. MissForest was chosen because it is particularly effective for datasets with mixed-type variables and complex interactions, maintaining low imputation errors and preserving relationships among variables even with high percentages of missing information [38].

Exploratory Analysis of the Longitudinal Cohort

CTCAE G3-4, ECOG PS change at the end of treatment, and early discontinuation of therapy were assessed by exploring absolute count and relative percentages. Hazard ratios (HRs) and Kaplan-Meier curves were computed to estimate the impact on survival outcomes [39]. The cumulative incidence-time curves of dose reduction due to AEs, separating patients with good therapy response and the ones with unplanned stops, were analyzed [40]. Moreover, the survival outcomes were estimated by stratifying patients by those receiving 4 doses (full course) without dose reduction, patients who received 4 cycles but with dose reduction, and patients who received fewer than 4 cycles.

Multivariate Analysis of Dose Adjustments

For the dose adjustment analysis, the association between dose reductions caused by an AE and the grade of toxicity was explored. This was done by training 2 Bayesian mixed-effect regression models: a binary logistic regression for the AE incidence [32,41] and a cumulative link regression for the grade of toxicity modeled as an ordinal outcome [32,42]. The priors used in these Bayesian models were the empirical ones computed from the analyzed cohort. The longitudinal variables (dose level, dose interval, and their mutual interaction) were analyzed with respect to the key baseline characteristics. The Bayesian logistic regression can be summarized as follows:

$$Y_{AE}=\log[P(AE_{i+1})/(1-P(AE_{i+1}))]=b_0+b_D D_i+b_{\Delta t}\Delta t_i+b_{D\Delta t} D_i\Delta t_i+\sum_{n=1}^{x} b_n x_n+\sum_{m=1}^{} \text{random}_m$$

where $P(AE_{i+1})$ is the probability of a dose reduction due to an adverse event at dose i+1 caused by the ith dose D_i, with interval Δt_i between the doses, x baseline covariates with b its estimated regression coefficient, and random effects (unexplained individual variability term). The cumulative link ordinal Bayesian regression can be summarized as follows:

$$Y_{GT}=\log[P(GT_{i+1}\le j)/(1-P(GT_{i+1}\le j))]=b_{0j}+b_D D_i+b_{\Delta t}\Delta t_i+b_{D\Delta t} D_i\Delta t_i+\sum_{n=1}^{x} b_n x_n+\sum_{m=1}^{} \text{random}_m$$

With similar notation for the Bayesian logistic regression, where $P(GT_{i+1}\le j)$ is the probability of grade of toxicity lower than or equal to j (with j=0,..., 4), and b_{0j} is the baseline intercept for the jth level. The Bayesian models were trained by running 4 independent Markov chain Monte Carlo simulations with 2500 iterations for each chain. The convergence of the chains was assessed through trace plots, density plots, Gelman-Rubin statistics, and estimates of the Markov chain effective sample size. In case of lack of convergence, the model was retrained by increasing the iterations (see Multimedia Appendix 2). The odds ratios (exponentials of the coefficients b) with the 95% CIs were estimated for all the covariates. The random effects

XSL•FO
RenderX

were estimated for the individual patients, administered doses, and the dose interval. This allowed for individual differences and the time-dependent nature of the dose variables to be considered.

Multivariate Analysis of Competitive Risks

For the variation of ECOG PS at the end of the treatment line and the early discontinuation of the therapy, a competitive risk time-to-event model was developed. Fine-Gray competitive risk regression was used with the outcomes [32], the PFS associated with the investigated treatment line, and the event of interest to estimate the distributed HRs of the covariates. The model can be summarized as follows:

$$hj(t)=h0(t)exp(b1x1+\ldots+bmxm)$$

where h_j is the cause-specific hazard of the event of interest, h_0 is the baseline hazard, and the exponential of the mth coefficient b is the distributed hazard ratio of the mth covariate x. The csHRs were computed for ECOG PS deterioration/improvement (in competition with unchanged PS or improvement and deterioration) and the early discontinuation of the therapy in competition with the planned stop of treatment due to partial response or stable disease. Here, we indicate the OS or PFS Cox hazard ratios with HR, csHR for the cause-specific hazard ratios computed with the Fine-Gray model, and OR for the odds ratios obtained by the Bayesian mixed effects models.

High-Risk Subgroups Analysis

Multivariate analysis was performed for the subgroups of patients with detected fragile conditions receiving treatment. The OS of these patients was compared to the OS of patients with similar characteristics that did not receive chemotherapy. This allowed the comparison of survival outcomes of patients with poor responses with the ones who did not receive chemotherapy, thus comparing the benefits and risks of the received treatment.

Aggregation and Discussion of Results

The considerations obtained during the exploratory data analysis are discussed together with the results obtained from the multivariate analysis. This allows for creating a connection between covariate effects on the longitudinal outcomes that can impact the OS or treatment decision. Moreover, the high-risk subgroups of patients are compared with the ones that did not receive treatment stratified by the significant covariates or longitudinal outcome discovered during the multivariate subgroup analysis to assess the risk-benefit of receiving treatment. Finally, the overall outcomes are discussed with the clinical experts regarding the potential clinical implications.

Ethical Considerations

The study received ethical approval from the Swedish Ethical Review Authority (approval number: 2023-07389-01). Patients still alive at the time of data collection signed a written informed consent. The data was anonymized prior to use, and the approval covers secondary analysis without additional consent. Research was performed in accordance with the Declaration of Helsinki.

Results

Selection of Clinical Scenarios and Patient Subgroups

The cohorts taken into consideration were the patients receiving PE as first-line therapy and the patients receiving either rechallenge with PE or IP as second-line therapy. The analysis included stage III-IV patients and excluded patients with ECOG PS of 4 (n=29). Figure S1 in Multimedia Appendix 3 reports the inclusion/exclusion flowchart of the cohort. In terms of toxicity, only neutropenia was analyzed in detail in this study because it was the most commonly reported treatment-related AE (see Table S2 in Multimedia Appendix 1). Figures S1-S4 in Multimedia Appendix 3 show Sankey flows of the subsequent treatment lines and the dose adjustments.

The identified high-risk subgroups included the first-line patients receiving PE with ECOG PS 2‐3 (n=114), with CNS metastases (n=42), over the age of 75 years (n=99), and with stage IVB (n=135). These subgroups were selected since they are known to be either underrepresented in RCTs or are associated with relatively poor OS compared to the general population [10]. The age threshold was chosen to define a sharp subgroup of older patients from the cohort. The multivariate analysis was not performed for patients with CNS metastases due to wide confidence intervals of ORs and csHRs caused by the small sample size of the subgroup (n=42).

Laboratory covariates referred to throughout the results include hemoglobin, sodium, LDH, CRP, albumin, creatinine, BSA, and GFR. For some individuals, blood values and lab tests were missing (neutrophil counts 44.9%, LDH 22.9%, and CRP 18.3%). The percentages of missing values for neutrophil counts at the start of the treatment lines were as follows: 50.9% in the first-line PE, 34.8% in the second-line PE, and 39.3% in the second line of IP. The high percentage of missing neutrophil values was due to these not being routinely measured at baseline before the initiation of chemotherapy treatment. Therefore, these values were imputed using the MissForest algorithm as described in the Methods section.

A summary of the analyzed cohorts for the first- and the second-line is reported in Table 1.

XSL•FO
RenderX

Table . Summary of the analyzed cohort.

Covariates	First-line	Second-line	
	Platinum etoposide (n=289)	Platinum etoposide (n=69)	Platinum irinotecan (n=56)
Baselines			
Age (years), n (%)			
<75	190 (65.7)	51 (73.9)	47 (83.9)
≥75	99 (34.3)	18 (26.1)	9 (16.1)
Sex, n (%)			
Female	171 (59.2)	34 (49.3)	29 (51.8)
Male	118 (40.8)	35 (50.7)	27 (48.2)
ECOG PS[a] baseline, n (%)			
0	58 (20.1)	16 (23.2)	15 (26.8)
1	117 (40.5)	35 (50.7)	29 (51.8)
2	75 (26.0)	13 (18.8)	9 (16.1)
3	39 (13.5)	5 (7.2)	3 (5.4)
Charles Comorbidity Index			
Mean (SD)	8.11 (2.30)	7.22 (2.18)	8.07 (1.87)
Median (min-max)	9 (2-15)	8 (3-12)	8 (3-12)
Missing, n (%)	4 (1.4)	2 (2.9)	2 (3.6)
Brain metastasis, n (%)			
No	247 (85.5)	66 (95.7)	46 (82.1)
Yes	42 (14.5)	3 (4.3)	10 (17.9)
Previous cancer, n (%)			
No	158 (54.7)	50 (72.5)	44 (78.6)
Yes	60 (20.8)	19 (27.5)	12 (21.4)
Missing	71 (24.6)	— 0 (0)	0 (0)
TNM[b] stage			
Overall stage, n (%)			
IIIA-B	61 (21.1)	19 (27.5)	8 (14.3)
IIIC	19 (6.6)	6 (8.7)	1 (1.8)
IVA	74 (25.6)	20 (29.0)	15 (26.8)
IVB	135 (46.7)	24 (34.8)	32 (57.1)
TNM-T, n (%)			
T1-2	70 (24.2)	16 (23.2)	9 (16.1)
T3	34 (11.8)	9 (13.0)	8 (14.3)
T4	185 (64.0)	44 (63.8)	39 (69.6)
TNM-N, n (%)			
N0	20 (6.9)	9 (13.0)	6 (10.7)
N1	12 (4.2)	3 (4.3)	2 (3.6)
N2	101 (34.9)	24 (34.8)	17 (30.4)
N3	155 (53.6)	33 (47.8)	31 (55.4)
Missing	1 (0.3)	0 (0)	0 (0)
TNM-M, n (%)			

XSL•FO
RenderX

Covariates	First-line	Second-line	
	Platinum etoposide (n=289)	Platinum etoposide (n=69)	Platinum irinotecan (n=56)
M0	80 (27.7)	25 (36.2)	9 (16.1)
M1a	37 (12.8)	13 (18.8)	6 (10.7)
M1b	37 (12.8)	7 (10.1)	9 (16.1)
M1c	135 (46.7)	24 (34.8)	32 (57.1)
Blood values and weight at the start of the treatment after MissingForrest Imputation			
Hemoglobin (g/L)			
Mean (SD)	130 (16.5)	127 (19.5)	122 (14.1)
Median (min-max)	130 (87-170)	128 (14-156)	120 (100-156)
C-reactive protein[c] (mg/L)			
Mean (SD)	34.7 (49.6)	10.3 (11.8)	21.8 (38.9)
Median (min-max)	15 (1-331)	7 (1-69)	10 (1-267)
Sodium (mmol/L)			
Mean (SD)	137 (4.97)	138 (4.93)	137 (4.41)
Median (min-max)	137 (112-148)	139 (127-151)	137 (120-146)
LDH[d,e] (µkat/L)			
Mean (SD)	26.1 (44.4)	23.4 (46.9)	22.7 (31.5)
Median (min-max)	9.30 (2-607)	10.8 (2.60 - 377)	11.2 (3-193)
Creatinine[c] (µmol/L)			
Mean (SD)	68.6 (22.9)	71.7 (21.9)	68.9 (20.3)
Median (min-max)	66 (26-204)	69 (38-139)	65 (32-123)
Albumin[c] (g/L)			
Mean (SD)	32.7 (14.3)	34.8 (4.33)	34.1 (7.62)
Median (min-max)	32 (3-250)	35 (24-43)	33 (25-82)
Neutrophil count[c] (10^9 /L)			
Mean (SD)	34.8 (44.5)	5.92 (2.66)	6.84 (3.66)
Median (min-max)	19.2 (1.8 - 514)	5.59 (0.5 - 14.5)	6.23 (0.9 - 26.4)
Weight (kg)			
Mean (SD)	70.6 (15.3)	69.8 (12.5)	75.8 (16.7)
Median (min-max)	68.5 (37-125)	70 (45-103)	78 (48-113)
Missing, n (%)	3 (1.0)	2 (2.9)	3 (5.4)
Dose calculation			
Body surface area (m^2)			
Mean (SD)	1.79 (0.21)	1.84 (0.38)	1.84 (0.25)
Median (min-max)	1.78 (1.18 - 2.37)	1.81 (1.31 - 4.53)	1.86 (1.09 - 2.31)
Missing, n (%)	4 (1.4)	2 (2.9)	1 (1.8)
GFR[e] (mL/min/1.7)			
Mean (SD)	78.6 (28.7)	78.2 (25.4)	86.4 (41.0)
Median (min-max)	75 (8-235)	75 (30-157)	80 (36-299)
Missing, n (%)	11 (3.8)	1 (1.4)	2 (3.6)

XSL•FO
RenderX

Covariates	First-line	Second-line	
	Platinum etoposide (n=289)	Platinum etoposide (n=69)	Platinum irinotecan (n=56)
Etoposide dose administration, n (%)			
10 days	69 (23.9)	9 (13.0)	0 (0)
3 days	220 (76.1)	60 (87.0)	0 (0)
Adverse effects of dose reduction			
Neutropenia, n (%)			
No	205 (70.9)	43 (62.3)	53 (94.6)
Yes	84 (29.1)	26 (37.7)	3 (5.4)
Other adverse events, n (%)			
No	225 (77.9)	52 (75.4)	33 (58.9)
Yes	64 (22.1)	17 (24.6)	23 (41.1)
CTCAE[f] G3-4, n (%)			
No	185 (64.0)	39 (56.5)	38 (67.9)
Yes	104 (36.0)	30 (43.5)	18 (32.1)
End of treatment reported response			
ECOG PS variation, n (%)			
Stable	87 (30.1)	36 (52.2)	28 (50.0)
Deterioration	141 (48.8)	21 (30.4)	17 (30.4)
Improvement	61 (21.1)	12 (17.4)	11 (19.6)
Treatment line response, n (%)			
Early discontinuation	110 (38.1)	36 (52.2)	37 (66.1)
Complete response, partial response, or stable disease	179 (61.9)	33 (47.8)	19 (33.9)
Radiotherapy			
Treatment intention, n (%)			
No radiotherapy	172 (59.5)	47 (56.0)	50 (70.4)
Curative	45 (15.6)	0 (0)	0 (0)
Consolidating	39 (13.5)	4 (4.8)	0 (0)
Palliative	33 (11.4)	30 (35.7)	19 (26.8)
Prophylactic cranial irradiation (PCI)	0 (0)	3 (3.6)	2 (2.8)
Survival outcomes			
Progression-free survival (days)			
Mean (SD)	185 (174)	157 (168)	127 (93)
Median (min-max)	163 (0 - 1980)	125 (0 - 1170)	110 (9-473)
Missing, n (%)	43 (14.9)	8 (11.6)	4 (7.1)
Overall survival (days)			
Mean (SD)	311 (303)	607 (391)	355 (157)
Median (min-max)	245 (11 - 2310)	516 (140 - 2310)	318 (144-911)
Status at the last reported information, n (%)			
Dead	262 (90.7)	63 (91.3)	54 (96.4)

XSL•FO
RenderX

Covariates	First-line	Second-line	
	Platinum etoposide (n=289)	Platinum etoposide (n=69)	Platinum irinotecan (n=56)
Alive (censored)	27 (9.3)	6 (8.7)	2 (3.6)

[a]ECOG PS: Eastern Cooperative Oncology Group performance status.

[b]TNM: tumor node metastasis.

[c]Log transformation was applied during the analysis to handle high skewness.

[d]LDH: lactate dehydrogenase.

[e]GFR: estimated glomerular filtration rate.

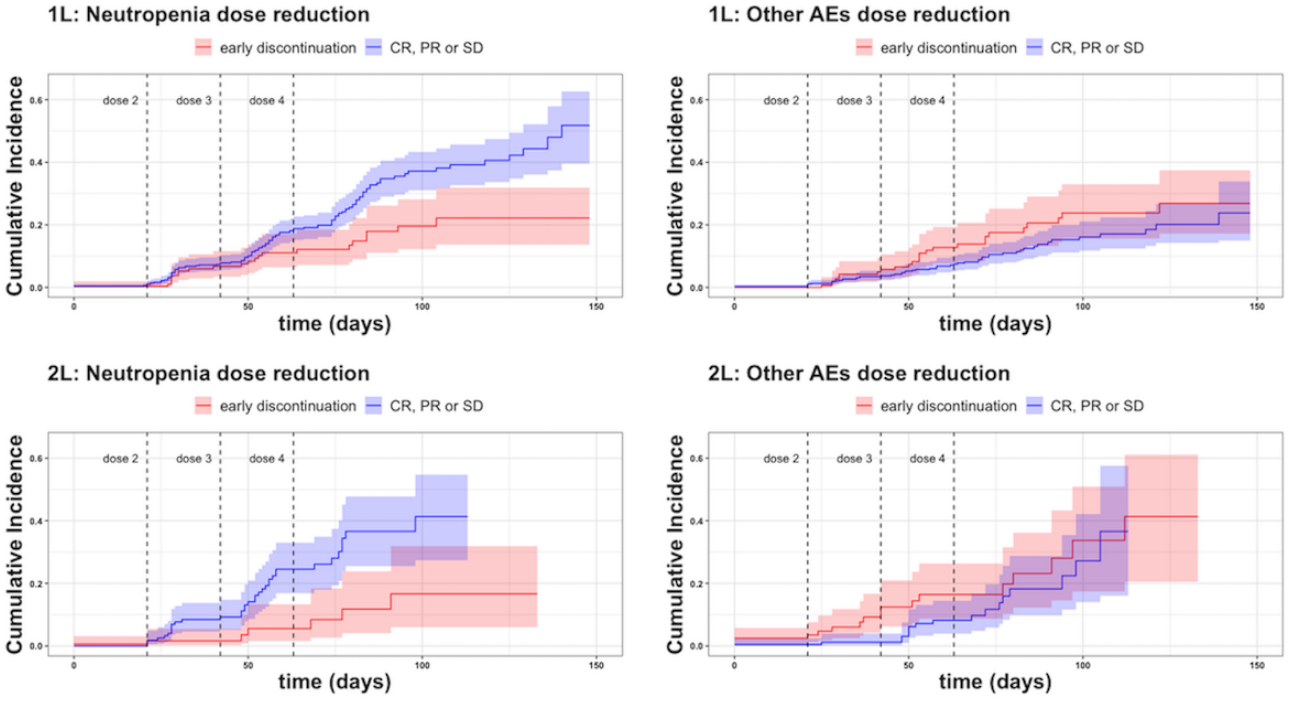
[f]CTCAE: Common Terminology Criteria for Adverse Events.

Exploratory Survival Analysis

In the first-line setting, chemotherapy doses were reduced in 127 out of 289 (43.9%) patients, mainly due to any grade of neutropenia (n=89) and AE G3-4 (n=104) (Table 1). Similar figures were observed for the second-line setting (PE n=69, IP n=56), with neutropenia mainly associated with PE treatment (n=26). At the end of the first-line setting, 179 (out of which 167 partial responses) patients did not experience early discontinuation of treatment. Early discontinuation of therapy happened in 110 cases due to clinical (n=48) or radiological (n=23) signs of disease progression. In the second-line setting, early discontinuation (n=72) was more common than completion of treatment (n=52).

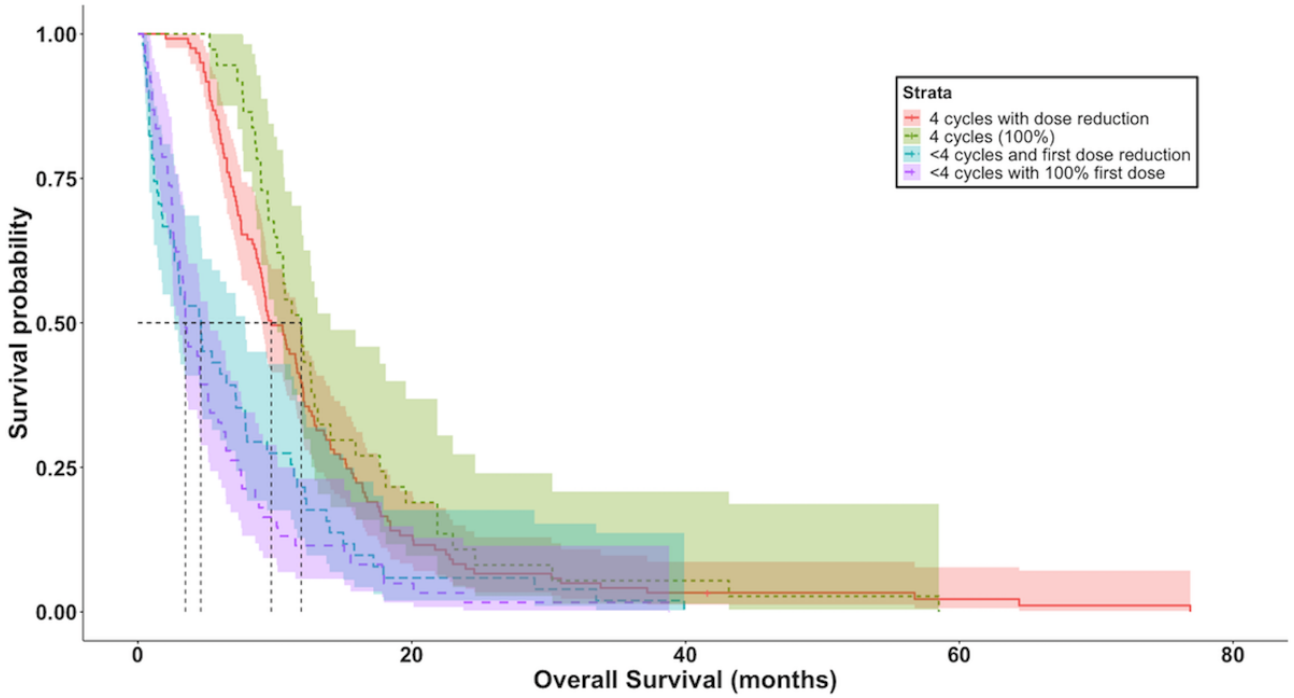
Figure 2 shows the cumulative incidences of dose adjustment due to neutropenia. The incidence was higher for patients in both lines who achieved treatment completion without premature discontinuation. In the first-line scenario, occurrences of neutropenia and other AEs were associated with more favorable OS (HR 0.70, 95% CI 0.53-0.92 and HR 0.72, 95% CI 0.57-0.95, respectively). In the second-line scenario, only neutropenia was associated with better OS (HR 0.51, 95% CI 0.33-0.80). OS of patients completing 4 cycles of first-line treatment was not affected by dose reduction due to AEs (HR 0.85, 95% CI 0.59-1.23; reference: patients completing 4 cycles without dose reduction) (Figure 3).

Figure 2. Cumulative incidences of dose reduction (%) due to neutropenia and other adverse events (AEs) for the first- and the second-line therapy. The curves are stratified by early discontinuation of chemotherapy (red curves) and therapy response (CR: complete; PR: partial; SD: stable disease) (blue curves). Dotted lines correspond to the 3-week intervals corresponding to planned administration of the second, third, and fourth doses. 1L: first-line therapy; 2L: second-line therapy.



XSL•FO
RenderX

Figure 3. Kaplan-Meier curves for overall survival of patients receiving first-line treatment stratified by administered doses. Patients receiving fewer than 4 doses experienced early discontinuation of the therapy.



Patients with early discontinuation of therapy in the first-line setting showed an expected markedly worse OS compared to patients who completed 4 cycles of chemotherapy (HR 4.27, 95% CI 3.27-5.57). Similarly, worse OS was observed for patients with early discontinuation during second-line therapy (HR 2.21, 95% CI 1.50-3.27). Moreover, ECOG PS deterioration was significantly associated with a poorer prognosis, even if this was only seen in the first-line setting (HR 2.05, 95% CI 1.54-2.72; reference: stable ECOG PS). These patients resulted in early discontinuation of therapy, of which 45 resulted in PFS less than 3 months (median 61, IQR 43-68 days).

Patients receiving PE as a second line completed the 4-dose cycle more often than the IP patients (73.36% vs 41.07%, respectively). Interestingly, the PFS of PE patients was similar compared to patients receiving IP (HR 0.84, 95% CI 0.57-1.23). However, as expected, second-line PE patients showed better OS (HR 0.51, 95% CI 0.33-0.78) compared to the ones receiving IP, when overall survival was computed from the start of the second-line treatment.

Multivariate Analysis

First-Line Therapy PE

Tables 2 and 3 detail the results of the dose adjustment and the competitive risk analysis for the first-line therapy. All the longitudinal variables presented significant variability (see Table S3 in Multimedia Appendix 1). Interindividual variability was higher than interoccasion variability. This high variability combined with the presence of patients with little longitudinal information (ie, patients receiving only one or two doses before interrupting the therapy) impacted some of the estimates resulting in large intervals.

XSL•FO
RenderX

Table . Bayesian mixed effects analysis for the first-line platinum etoposide patients.

Covariates	Adverse effects, OR[a] (95% CI)	Neutropenia grade, OR (95% CI)
Dose interval	1.00 (0.94 - 1.05)	0.99 (0.95 - 1.03)
Dose platinum	0.60 (0.07 - 4.33)	0.56 (0.14 - 2.18)
Dose etoposide	5.97 (1.41 - 30.5)	3.55 (1.03 - 13.3)
Dose platinum × dose interval	1.00 (0.97 - 1.03)	1.01 (0.98 - 1.05)
Dose etoposide × dose interval	1.02 (0.99 - 1.06)	1.04 (1.00 - 1.08)
Etoposide dose load		
3 days	—	—
10 days	1.69 (0.31 - 9.23)	5.03 (1.25 - 23.9)
Weight	0.28 (0.03 - 1.88)	1.52 (0.30 - 7.47)
Estimated glomerular filtration rate	3.14 (0.47 - 24.8)	1.59 (0.47 - 5.39)
Body surface area	0.53 (0.04 - 5.95)	0.16 (0.02 - 1.12)
Hemoglobin	0.77 (0.37 - 1.50)	0.88 (0.52 - 1.46)
Creatinine	3.98 (1.45 - 12.5)	1.32 (0.67 - 2.56)
C-reactive protein	0.83 (0.40 - 1.76)	0.99 (0.57 - 1.72)
Sodium	0.97 (0.50 - 1.91)	1.11 (0.68 - 1.81)
Lactate dehydrogenase	1.59 (0.78 - 3.41)	1.56 (0.91 - 2.82)
Albumin	0.61 (0.30 - 1.20)	0.83 (0.48 - 1.40)
Neutrophil counts	0.39 (0.16 - 0.82)	0.67 (0.36 - 1.21)
Age (years)		
≥75	—	—
<75	0.92 (0.20 - 4.42)	0.77 (0.26 - 2.32)
Gender		
Female	—	—
Male	0.27 (0.05 - 1.39)	1.96 (0.47 - 8.81)
ECOG PS[b] baseline		
0	—	—
1	0.34 (0.07 - 1.56)	0.89 (0.31 - 2.61)
2	0.49 (0.07 - 3.22)	1.14 (0.29 - 4.71)
3	0.45 (0.04 - 4.05)	4.24 (0.66 - 36.5)
TNM-T[c]		
T4	—	—
T1-2	1.87 (0.44 - 8.35)	0.68 (0.23 - 1.97)
T3	0.50 (0.07 - 3.54)	0.73 (0.19 - 3.04)
TNM-N		
N3	—	—
N0	0.84 (0.07 - 10.1)	1.92 (0.30 - 12.2)
N1	5.65 (0.17 - 190)	37.4 (2.46 - 721)
N2	0.97 (0.25 - 3.86)	1.03 (0.37 - 2.61)
TNM-M		
M1c	—	—
M0	6.01 (0.72 - 61.9)	1.47 (0.31 - 6.97)

XSL•FO
RenderX

Covariates	Adverse effects, OR[a] (95% CI)	Neutropenia grade, OR (95% CI)
M1a	2.95 (0.45 - 21.2)	1.14 (0.30 - 4.64)
M1b	3.22 (0.43 - 25.6)	0.47 (0.10 - 2.28)
Smoking status		
Smoker	—	—
Previous smoker	0.92 (0.24 - 3.49)	1.11 (0.39 - 3.21)
Nonsmoker	0.62 (0.02 - 18.5)	1.05 (0.07 - 14.8)
Charles Comorbidity Index	1.49 (0.55 - 3.92)	1.26 (0.62 - 2.63)
Brain metastasis		
No	—	—
Yes	0.70 (0.10 - 4.85)	0.86 (0.21 - 3.63)
Previous cancer		
No	—	—
Yes	0.66 (0.13 - 2.84)	0.58 (0.20 - 1.57)

[a]OR: odds ratio.

[b]ECOG PS: Eastern Cooperative Oncology Group performance status.

[c]TNM: tumor node metastasis.

XSL•FO
RenderX

Table . Competitive risk analysis results for the first-line treatment of platinum etoposide.

	Early discontinuation of treatment				
Covariates	n	csHR[a] (95% CI)	n	ECOG PS[b] deterioration, csHR (95% CI)	ECOG PS improvement, csHR (95% CI)
Age (years)					
<75	149	—	140	—	—
≥75	82	1.01 (0.55 - 1.84)	64	1.57 (0.75 - 3.27)	0.43 (0.14 - 1.33)
Gender					
Female	131	—	111	—	—
Male	100	0.51 (0.29 - 0.90)	93	0.89 (0.41 - 1.96)	0.52 (0.13 - 2.04)
ECOG PS baseline					
0	49	—	—	—	—
1	94	1.08 (0.53 - 2.19)	—	—	—
2	61	2.93 (1.31 - 6.52)	—	—	—
3	27	4.15 (1.68 - 10.2)	—	—	—
Charles Comorbidity Index	231	1.10 (0.92 - 1.32)	204	0.98 (0.80 - 1.22)	1.09 (0.83 - 1.45)
Previous cancer					
No	181	—	159	—	—
Yes	50	0.75 (0.38 - 1.50)	45	0.77 (0.34 - 1.71)	1.34 (0.40 - 4.46)
Brain metastasis					
No	196	—	176	—	—
Yes	35	0.83 (0.40 - 1.75)	28	0.64 (0.24 - 1.71)	1.50 (0.34 - 6.50)
TNM-T[c]					
T4	149	—	130	—	—
T1-2	52	1.55 (0.82 - 2.93)	47	1.33 (0.67 - 2.62)	0.79 (0.29 - 2.11)
T3	30	0.78 (0.42 - 1.45)	27	0.96 (0.50 - 1.85)	0.54 (0.13 - 2.20)
TNM-N					
N3	129	—	113	—	—
N0	17	0.17 (0.04 - 0.70)	15	0.73 (0.27 - 1.98)	0.48 (0.13 - 1.81)
N1	8	1.05 (0.36 - 3.08)	6	0.13 (0.04 - 0.49)	0.00 (0.00 - 0.00)
N2	77	0.80 (0.47 - 1.36)	70	0.77 (0.40 - 1.51)	0.59 (0.22 - 1.54)
TNM-M					
M1c	119	—	102	—	—
M0	52	1.52 (0.53 - 4.41)	47	0.49 (0.11 - 2.14)	1.07 (0.17 - 6.67)
M1a	29	1.25 (0.58 - 2.69)	25	0.55 (0.18 - 1.69)	3.05 (1.04 - 8.93)
M1b	31	1.03 (0.53 - 1.97)	30	1.21 (0.54 - 2.72)	4.26 (0.81 - 22.5)
Hemoglobin	231	1.01 (0.99 - 1.03)	204	1.00 (0.98 - 1.01)	0.99 (0.97 - 1.02)
Lactate dehydrogenase	231	1.44 (1.14 - 1.82)	204	1.19 (0.94 - 1.49)	0.98 (0.58 - 1.64)
C-reactive protein	231	1.13 (0.92 - 1.38)	204	1.02 (0.80 - 1.31)	1.22 (0.89 - 1.68)
Neutrophil counts	231	0.99 (0.73 - 1.35)	204	1.40 (1.06 - 1.85)	0.51 (0.32 - 0.82)
Sodium	231	1.02 (0.97 - 1.07)	204	1.01 (0.95 - 1.07)	1.05 (0.96 - 1.15)
Creatinine	231	2.40 (0.84 - 6.83)	204	2.54 (0.68 - 9.48)	0.80 (0.09 - 7.39)
Weight	231	0.99 (0.95 - 1.03)	204	1.06 (1.00 - 1.12)	0.93 (0.85 - 1.02)

XSL•FO
RenderX

Covariates	Early discontinuation of treatment				
	n	csHR[a] (95% CI)	n	ECOG PS[b] deterioration, csHR (95% CI)	ECOG PS improvement, csHR (95% CI)
Estimated glomerular filtration rate	231	1.00 (0.99 - 1.02)	204	1.00 (0.99 - 1.01)	0.99 (0.97 - 1.01)
Body surface area	231	2.97 (0.14 - 62.5)	204	0.01 (0.00 - 1.58)	144 (0.07 - 286,657)
Etoposide dose load					
10 days	57	—	51	—	—
3 days	174	0.39 (0.22 - 0.69)	153	0.72 (0.35 - 1.48)	3.65 (0.91 - 14.7)
Radiotherapy					
No radiotherapy or palliative	175	—	151	—	—
Curative	24	0.41 (0.14 - 1.24)	22	1.35 (0.42 - 4.28)	0.31 (0.06 - 1.47)
Consolidating	32	0.20 (0.06 - 0.69)	31	0.50 (0.16 - 1.50)	0.53 (0.24 - 1.17)

[a]csHR: cause-specific hazard ratio.

[b]ECOG PS: Eastern Cooperative Oncology Group performance status.

[c]TNM: tumor node metastasis.

The results show a strong association between higher etoposide dose levels and the subsequent occurrence of AEs (OR 5.97, 95% CI 1.41-30.5), specifically neutropenia (OR 3.55, 95% CI 1.03-13.3). Early discontinuation of therapy was associated with higher LDH levels (csHR 1.44, 95% CI 1.14-1.82), baseline ECOG PS 2 (csHR 2.93, 95% CI 1.31-6.52), and baseline ECOG PS 3 (csHR 4.15, 95% CI 1.68-10.20).

Males showed less probability of early discontinuation of the first treatment line compared to females (csHR 0.51, 95% CI 0.29-0.90). As did subgroups of patients with M0 and those receiving consolidating radiotherapy (n=29) (Table 3). Similarly, patients with M1a status regardless of T and N stage were associated with ECOG PS improvement (Table 3).

Second-Line Therapy PE-IP

Results of the multivariate analysis for the second-line treatment are reported in Tables S4 and S5 in Multimedia Appendix 1. Due to the large confidence intervals of the ORs, the ORs are reported on a log scale.

A higher dose of platinum was associated with a higher incidence of AEs and a higher risk of neutropenia. Patients receiving IP experienced fewer AEs. Male patients were less likely to experience AEs and neutropenia. However, male patients were associated with worse PFS (HR 1.62, 95% CI 1.10-2.38) due to early discontinuation of the second line, regardless of PE or IP (n=44). Higher sodium levels were instead associated with lower risk of neutropenia and ECOG PS deterioration. Patients with TNM-M M1a had a higher probability of ECOG PS deterioration and more AEs. CNS metastasis was linked to increased risk for ECOG PS deterioration and early discontinuation of therapy.

High-Risk Subpopulation Analysis: To Treat or Not to Treat?

As defined at the beginning of the Results section, high-risk subgroups included patients with ECOG PS 2 - 3, aged ≥75 years, and TNM stage IVB disease. Multivariate analyses were performed separately within each subgroup to identify predictors of toxicity, ECOG PS variability, early discontinuation, and survival outcomes.

The results of these analyses are reported in Tables S6-S11 in Multimedia Appendix 1. Table 4 summarizes the significant covariates for each subgroup. Similar to the second-line therapy analysis, the ORs were reported in log scale due to the wide confidence intervals.

XSL•FO
RenderX

Table . Qualitative summary of statistically significant covariates identified in the multivariate analysis for high-risk subgroups[a].

High-risk subgroup	Adverse events occurrence	Neutropenia	ECOG PS[b] deterioration	ECOG PS improvement	Early discontinuation
TNM-IVB[c]	Albumin (–) Neutrophil counts (–)	Dose × interval etoposide (+) BSA[d] (–)	Neutrophil counts (+)	Neutrophil counts (–) LDH[e] (+)	LDH (+) Creatinine (+)
ECOG PS 2‐3	Etoposide 10 days regimen (+) Neutrophil counts (–) Age ≥75 years (+) Male (–)	Dose × interval platinum (+)	Neutrophil counts (+)	Neutrophil counts (–) LDH (+)	Etoposide 10-days regime (+) TNM-N0 (–) TNM-M0 (+) HB[f] (+) LDH (+) CRP[g] (+) CCI[h] (+)
Age ≥75 years	Male (–) ECOG 3 (–) CNS[i] (–)	Dose × interval etoposide (+) Weight (–) Male (–) TNM-N0 (+) TNM-M1a (+)	TNM-N2 (–)	——	ECOG 2 (+) TNM-N0 (–) LDH (+)

[a]For each outcome, covariates with statistically significant associations are shown. "+" indicates odds ratio (OR) or cause-specific hazard ratio (csHR) with a direct association with the studied outcome. "–" indicates OR or csHR with an inverse association. Detailed results are reported in Tables S6-S11 in Multimedia Appendix 1.

[b]ECOG PS: Eastern Cooperative Oncology Group performance status.

[c]TNM: tumor node metastasis.

[d]BSA: body surface area.

[e]LDH: lactate dehydrogenase.

[f]HB: hemoglobin.

[g]CRP: C-reactive protein.

[h]CCI: Charles Comorbidity Index.

[i]CNS: brain metastasis.

Results observed in the subgroup analyses that aligned with the overall cohort were the consistent association of elevated LDH and ECOG PS 2‐3 with an increased risk of early discontinuation of therapy across all subgroups.

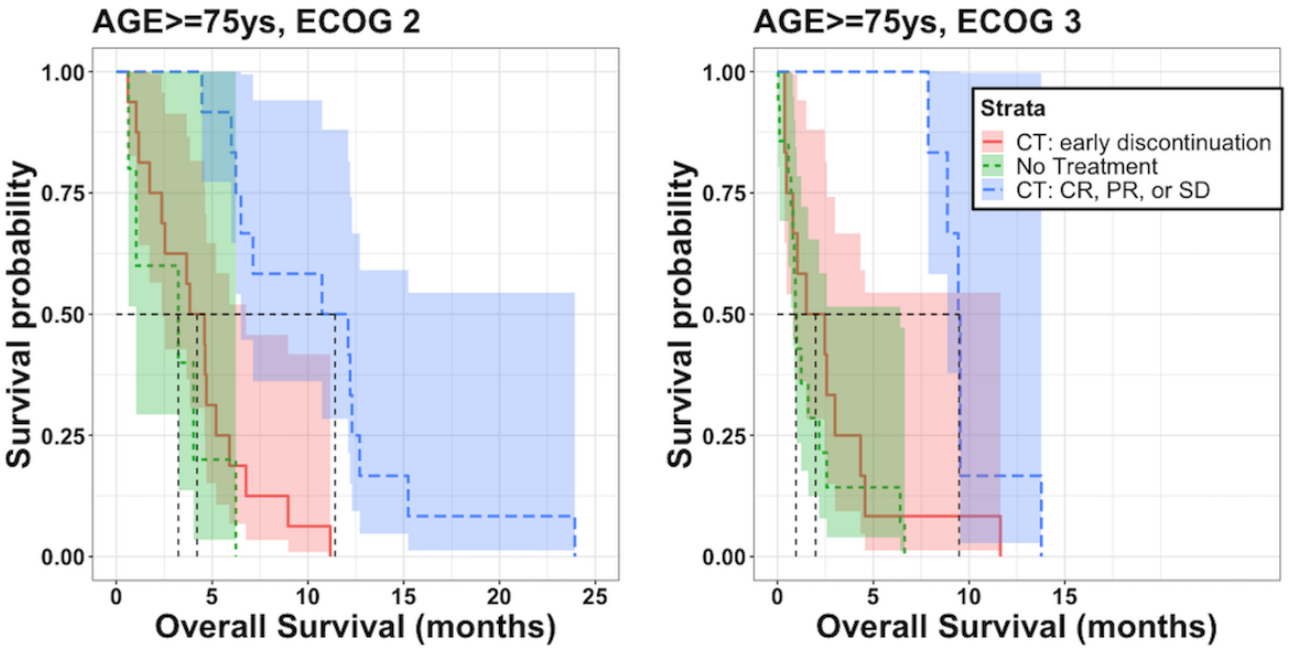
In patients with TNM stage IVB and in patients aged ≥75 years, a higher etoposide dose was associated with an increased risk of neutropenia, while older male patients (aged ≥75 years) and male patients with ECOG PS 2‐3 were associated with a lower risk for AEs. For patients with ECOG PS 2‐3, the platinum dose was associated with an increased risk of neutropenia, while the etoposide 10-day regimen was associated with an increased risk of AEs.

The comparison with patients who did not receive chemotherapy showed that, although the occurrence of AEs was associated with longer survival in the overall cohort, patients with ECOG PS 2‐3 or aged ≥75 years who experienced early discontinuation had HRs closer to those of the nontreated subgroup rather than to those patients who achieved disease control (Table S12 in Multimedia Appendix 1).

When further stratifying the high-risk subgroups using the significant variables reported in Table 4, the subgroups showed similar survival to the nontreated subgroup. For example, Figure 4 shows the Kaplan-Meier curves of the comparative analysis for patients with ECOG PS 2‐3 and age ≥75 years.

XSL•FO
RenderX

Figure 4. Kaplan-Meier curves of patients older than 75 years with ECOG performance status (PS) 2 - 3 stratified by good response to first-line chemotherapy, early discontinuation of treatment, and patients who did not receive chemotherapy. CR: complete response; CT: chemotherapy; ECOG: Eastern Cooperative Oncology Group; PR: partial response; SD: stable disease.



Discussion

In this study, we analyzed a longitudinal real-world cohort of patients with SCLC and their treatment pathways. The analysis allowed the study of the relationships between treatment decisions, phases of progression of disease, occurrence of AEs, and overall survival. To the best of our knowledge, this is one of the most detailed longitudinal datasets of SCLC real-world treatment decisions. This type of data aims to add novel insights on dosing effects over time, the emergence of toxicity, and the decision points relative to the clinical outcomes.

The improved prognosis of patients who experienced neutropenia aligns with previous findings [43,44]. The analysis allowed the investigation of treatment-related AEs in relation to the efficacy of therapy or to detect when this was instead linked to deterioration of the patient's conditions. For example, the cumulative incidence curves of AEs stratified by responding patients against early discontinuation of the therapy in Figure 2 showed the role of neutropenia when the time windows of AEs may overlap with the efficacy of treatment (Figure 3).

We identified significant ORs relating to the dose of etoposide and increased occurrence of AEs. The analysis showed a similar OS for patients receiving 4 doses during first-line therapy, regardless of whether the patient had a dose adjustment/reduction or not (Figure 3). A worse prognosis was observed when there was a deterioration in ECOG PS or early discontinuation of treatment.

We interpret our findings as indicating that the occurrence of neutropenia reflects a longer duration of treatment exposure, rather than being a causal factor for improved survival. Patients who completed all 4 treatment cycles—often at reduced doses—demonstrated better outcomes than those receiving fewer cycles at higher doses. This is supported by the higher incidence of neutropenia among patients with treatment response, suggesting that extended exposure, not intensity, is the key contributor to therapeutic benefit.

SCLC treatment effects vary widely between individuals. This could suggest the need to identify key covariates and clinically relevant subgroups that might inform a more tailored dosing, or in some cases, guide decisions against treatment altogether, in order to improve survival while avoiding AEs in patients unlikely to benefit from therapy [11,12,45]. Indeed, etoposide dose calculation is based on BSA, which has been shown to present such limitations [46]. Similar analyses could be done for the second-line therapy. In fact, only a few patients received the full expected dose without dose reduction, and the platinum dose and PE treatment were strongly associated with AEs.

ECOG PS 2 - 3 and elevated LDH serum levels are known to be prognostic factors of worse OS in SCLC [9,47]. Our results, obtained by the careful analysis of the longitudinal data line by line with the clinical experts, added a new insight by showing their role in predicting early discontinuation of treatment, which was associated with significantly poorer OS.

Another finding was the high probability of a favorable response to first-line therapy in male patients, along with their lower likelihood of experiencing AEs (Tables 3 and 4; Tables S4 and S5 in Multimedia Appendix 1). Moreover, when stratifying older patients by sex, women who developed neutropenia exhibited survival rates comparable to those in the same subgroup who did not receive chemotherapy. A possible explanation is that female patients have historically been underrepresented in clinical trials [10,48,49]. In addition, in the second-line setting, the worse PFS and high frequency of early discontinuation for male patients might suggest the need for further studies regarding different therapy responses for males and females.

The high-risk subgroup analysis allowed detection of key groups of patients such as older patients with poor ECOG PS, where

XSL•FO
RenderX

an eventual poor response to therapy is associated with similar survival as compared to patients not receiving chemotherapy. This finding is clinically relevant, as it highlights the importance of balancing the benefit and risk of treating such fragile patients and may help clinicians avoid overtreating this vulnerable population. These outcomes underline the potential role that real-world data may fill in providing evidence on treatment in groups of patients that are underrepresented in clinical trials [10].

The design of the analytic pipeline relied on the consideration that the data should be analyzed using models capable of capturing the complexity of the patient treatment pathways without sacrificing interpretability (such as black box machine learning models [25]). We believe that this is an important consideration when approaching real-world data. In fact, the discussions with the clinicians on the obtained results allowed the assessment of the validity of the results, the detection of potential inherent bias, and confounders related to operations and nonobservable effects not captured from the data. For example, it was observed that the 10-day dosing regimen of etoposide resulted in poorer survival. At the time of the present analysis, this dosing regimen had, however, already been abandoned due to toxicity caused by the limited time for patients to recover between the cycles (Tables 2 and 3).

The study had several limitations. The reason for missing values was not at random for some of the variables, thus introducing some biases in the results. Particularly for baseline neutrophil counts, where missing observations were high across all treatment lines. This bias is partially mitigated by the MissForest algorithm because it is a nonparametric model, and the algorithm does not need an explicit assumption about the missing mechanism (ie, missing at random) [38].

Further, these variables were available only at the start of the treatment line, meaning that we were not able to analyze their time profiles (ie, neutrophil counts over time between doses). Moreover, it should be stated that AEs are not always described in medical notes, thus introducing investigator bias in the collected data.

Another limitation was the sample size of the high-risk subgroups and patients receiving the second-line treatment, resulting in higher estimates of variability of the outcomes as compared to the first-line treatment cohort. In detail, the odds ratios of the mixed effects models resulted in wide confidence intervals. The reason for this high variability in outcomes could be that these small cohorts were also associated with fewer longitudinal observations per patient (the number of administered doses). Increasing the sample sizes of these subgroups in the future by collecting more data will help to provide more precise estimates of the covariate effects. Moreover, patients receiving more than two lines were not included in the multivariate analysis, as this was only observed in a few instances.

The median age of our population was 72 years, yet we chose 75 years to better isolate a more vulnerable subset of patients by moving further from the median. This aligns with existing studies that identify this age group as particularly fragile, often requiring specialized considerations in treatment planning [50]. Future studies exploring different criteria and strategies to separate older patients would be required.

This was a single-center study, and expanding the analysis by including other SCLC cohorts from other centers would be beneficial. A comparison of the dose adaptation and the follow-up processes in clinical trials would be of high importance [10]. A future aspect to explore would be to expand the collection of longitudinal real-world data, designed similarly to the study presented here, in larger initiatives such as multicenter studies or cancer registries [51-53], and recent new proposed therapies (ie, durvalumab after chemoradiotherapy [54], tarlatamab/lubri in second- and third-line settings [55], and immune-oncology therapies [9]).

This study could inform the design of future works, such as multistate models of patient survival in function of different therapies and events [56,57], or model-informed precision dosing models to explore individualized treatment and dosing to reduce the risk of AEs (eg, neutropenia) without compromising the efficacy of the treatment in terms of survival outcomes [58,59]. Furthermore, the outcomes from the approaches presented herein can inform more reliable future analyses using, for example, machine learning algorithms suitable for modeling longitudinal processes, such as physics-based neural networks [60].

A continued fruitful collaboration between clinical and technical experts when analyzing complex data and contextualizing the clinical processes, as we did in this work, would enhance the impact of real-world data in clinical practice and clinical drug development.

Acknowledgments

The authors declare that they used the generative AI tool Copilot GPT-4 for proofreading and editing, under full human supervision. The authors take full responsibility for the final manuscript.

Funding

This study was supported by the Swedish Cancer Society (grants CAN 2018/597 and CAN2021/1469 Pj01) to RL and by the Stockholm Cancer Society (grant #201202 to RL and grants #174063, #201103, and #231123 to LDP).

Data Availability

The data that support the findings of this study are not openly available due to reasons of sensitivity but are available from the authors upon reasonable request.

XSL•FO
RenderX

Authors' Contributions

LM and AD wrote the manuscript. LM, ASD, AD, LDP, ST, JR, RL, and SM designed the research. LM, AD, LDP, and ASD performed the research. LM, AD, ASD, and LDP analyzed the data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Result tables.
[XLSX File, 65 KB - cancer_v12i1e84042_app1.xlsx]

Multimedia Appendix 2
Diagnostics and convergence assessment of Monte Carlo Markov Chains.
[DOCX File, 2419 KB - cancer_v12i1e84042_app2.docx]

Multimedia Appendix 3
Sankey flows of the subsequent treatment lines and the dose adjustments.
[DOCX File, 3363 KB - cancer_v12i1e84042_app3.docx]

References

1. Pietanza MC, Byers LA, Minna JD, Rudin CM. Small cell lung cancer: will recent progress lead to improved outcomes? Clin Cancer Res 2015 May 15;21(10):2244-2255. [doi: 10.1158/1078-0432.CCR-14-2958] [Medline: 25979931]
2. Reddy HG, Qin A, Kalemkerian GP. Emerging drugs for small cell lung cancer: a focused review on immune checkpoint inhibitors. Expert Opin Emerg Drugs 2020 Sep;25(3):353-366. [doi: 10.1080/14728214.2020.1798929] [Medline: 32683991]
3. Horn L, Mansfield AS, Szczęsna A, et al. First-line atezolizumab plus chemotherapy in extensive-stage small-cell lung cancer. N Engl J Med 2018 Dec 6;379(23):2220-2229. [doi: 10.1056/NEJMoa1809064] [Medline: 30280641]
4. Shaw J, Pundole X, Balasubramanian A, et al. Recent treatment patterns and real-world survival following first-line anti-PD-L1 treatment for extensive-stage small cell lung cancer. Oncologist 2024 Dec 6;29(12):1079-1089. [doi: 10.1093/oncolo/oyae234] [Medline: 39349396]
5. Bar J, Xu Q, Karve S, Hingorani P, Virani AA, Ready NE. Real‐world response and survival outcomes and treatment patterns in patients with extensive stage small‐cell lung cancer receiving third‐line treatment. Cancer Rep 2025 Aug;8(8):e70289. [doi: 10.1002/cnr2.70289]
6. Dingemans AMC, Früh M, Ardizzoni A, et al. Small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2021 Jul;32(7):839-853. [doi: 10.1016/j.annonc.2021.03.207] [Medline: 33864941]
7. Tendler S, Zhan Y, Pettersson A, et al. Treatment patterns and survival outcomes for small-cell lung cancer patients – a Swedish single center cohort study. Acta Oncol 2020 Apr;59(4):388-394. [doi: 10.1080/0284186X.2019.1711165] [Medline: 31910696]
8. Ardizzoni A, Tiseo M, Boni L. Validation of standard definition of sensitive versus refractory relapsed small cell lung cancer: a pooled analysis of topotecan second-line trials. Eur J Cancer 2014 Sep;50(13):2211-2218. [doi: 10.1016/j.ejca.2014.06.002] [Medline: 24981975]
9. Johal S, Hettle R, Carroll J, Maguire P, Wynne T. Real-world treatment patterns and outcomes in small-cell lung cancer: a systematic literature review. J Thorac Dis 2021 Jun;13(6):3692-3707. [doi: 10.21037/jtd-20-3034] [Medline: 34277061]
10. Marzano L, Darwich AS, Dan A, et al. Exploring the discrepancies between clinical trials and real-world data: a small-cell lung cancer study. Clin Transl Sci 2024 Aug;17(8):e13909. [doi: 10.1111/cts.13909] [Medline: 39113428]
11. Darwich AS, Ogungbenro K, Vinks AA, et al. Why has model-informed precision dosing not yet become common clinical reality? Lessons from the past and a roadmap for the future. Clin Pharmacol Ther 2017 May;101(5):646-656. [doi: 10.1002/cpt.659] [Medline: 28182269]
12. Darwich AS, Polasek TM, Aronson JK, et al. Model-informed precision dosing: background, requirements, validation, implementation, and forward trajectory of individualizing drug therapy. Annu Rev Pharmacol Toxicol 2021 Jan 6;61:225-245. [doi: 10.1146/annurev-pharmtox-033020-113257] [Medline: 33035445]
13. Povsic M, Enstone A, Wyn R, Kornalska K, Penrod JR, Yuan Y. Real-world effectiveness and tolerability of small-cell lung cancer (SCLC) treatments: A systematic literature review (SLR). PLOS ONE 2019;14(7):e0219622. [doi: 10.1371/journal.pone.0219622] [Medline: 31318909]
14. Moser SS, Bar J, Kan I, et al. Real world analysis of small cell lung cancer patients: prognostic factors and treatment outcomes. Curr Oncol 2021 Jan 8;28(1):317-331. [doi: 10.3390/curroncol28010036] [Medline: 33435584]

XSL•FO
RenderX

15. Zellweger NM, Schmid S, Bertschinger M, et al. 1540P Real-world analysis of outcomes of first-line chemo-immunotherapy in patients with extensive disease small cell lung cancer (ED-SCLC). Ann Oncol 2022 Sep;33:S1251. [doi: 10.1016/j.annonc.2022.07.1634]
16. Derksen JWG, May AM, Koopman M. The era of alternative designs to connect randomized clinical trials and real-world data. Nat Rev Clin Oncol 2019 Sep;16(9):589. [doi: 10.1038/s41571-019-0250-0] [Medline: 31296966]
17. Baumfeld Andre E, Carrington N, Siami FS, et al. The current landscape and emerging applications for real-world data in diagnostics and clinical decision support and its impact on regulatory decision making. Clin Pharmacol Ther 2022 Dec;112(6):1172-1182. [doi: 10.1002/cpt.2565] [Medline: 35213741]
18. Schurman B. The framework for FDA's real-world evidence program. Appl Clin Trials 2019;28(4):15-17 [FREE Full text]
19. Hiramatsu K, Barrett A, Miyata Y, PhRMA Japan Medical Affairs Committee Working Group 1. Current status, challenges, and future perspectives of real-world data and real-world evidence in Japan. Drugs Real World Outcomes 2021 Dec;8(4):459-480. [doi: 10.1007/s40801-021-00266-3] [Medline: 34148219]
20. Dang A. Real-world evidence: a primer. Pharmaceut Med 2023 Jan;37(1):25-36. [doi: 10.1007/s40290-022-00456-6] [Medline: 36604368]
21. Greenhalgh T, Papoutsi C. Studying complexity in health services research: desperately seeking an overdue paradigm shift. BMC Med 2018 Jun 20;16(1):95. [doi: 10.1186/s12916-018-1089-4] [Medline: 29921272]
22. Thatcher N, Girling DJ, Hopwood P, et al. Improving survival without reducing quality of life in small-cell lung cancer patients by increasing the dose-intensity of chemotherapy with granulocyte colony-stimulating factor support: results of a British Medical Research Council multicenter randomized trial. J Clin Oncol 2000 Jan 1;18(2):395-395. [doi: 10.1200/JCO.2000.18.2.395]
23. Socinski MA, Smit EF, Lorigan P, et al. Phase III study of pemetrexed plus carboplatin compared with etoposide plus carboplatin in chemotherapy-naive patients with extensive-stage small-cell lung cancer. J Clin Oncol 2009 Oct 1;27(28):4787-4792. [doi: 10.1200/JCO.2009.23.1548] [Medline: 19720897]
24. Ibrahim JG, Chu H, Chen LM. Basic concepts and methods for joint models of longitudinal and survival data. J Clin Oncol 2010 Jun 1;28(16):2796-2801. [doi: 10.1200/JCO.2009.25.0654] [Medline: 20439643]
25. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019 May;1(5):206-215. [doi: 10.1038/s42256-019-0048-x] [Medline: 35603010]
26. Alabi RO, Youssef O, Pirinen M, et al. Machine learning in oral squamous cell carcinoma: current status, clinical concerns and prospects for future – a systematic review. Artif Intell Med 2021 May;115:102060. [doi: 10.1016/j.artmed.2021.102060] [Medline: 34001326]
27. Aubel P, Antigny M, Fougeray R, Dubois F, Saint-Hilary G. A Bayesian approach for event predictions in clinical trials with time-to-event outcomes. Stat Med 2021 Dec 10;40(28):6344-6359. [doi: 10.1002/sim.9186] [Medline: 34541701]
28. Yin J, Qin R, Ezzalfani M, Sargent DJ, Mandrekar SJ. A Bayesian dose-finding design incorporating toxicity data from multiple treatment cycles. Stat Med 2017 Jan 15;36(1):67-80. [doi: 10.1002/sim.7134] [Medline: 27633877]
29. Zhao D, Zhu J, Wang L. Bayesian interval-based oncology dose-finding design with repeated quasi-continuous toxicity model. Contemp Clin Trials 2021 Mar;102:106265. [doi: 10.1016/j.cct.2021.106265] [Medline: 33418097]
30. Xu R. Measuring explained variation in linear mixed effects models. Stat Med 2003 Nov 30;22(22):3527-3541. [doi: 10.1002/sim.1572] [Medline: 14601017]
31. Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982 Dec;38(4):963-974. [Medline: 7168798]
32. Coz E, Fauvernier M, Maucort-Boulch D. An overview of regression models for adverse events analysis. Drug Saf 2024 Mar;47(3):205-216. [doi: 10.1007/s40264-023-01380-7] [Medline: 38007401]
33. Dignam JJ, Zhang Q, Kocherginsky M. The use and interpretation of competing risks regression models. Clin Cancer Res 2012 Apr 15;18(8):2301-2308. [doi: 10.1158/1078-0432.CCR-11-2097] [Medline: 22282466]
34. Calvert AH, Newell DR, Gumbrell LA, et al. Carboplatin dosage: prospective evaluation of a simple formula based on renal function. J Clin Oncol 1989 Nov;7(11):1748-1756. [doi: 10.1200/JCO.1989.7.11.1748] [Medline: 2681557]
35. Jiang S, Huang L, Zhen H, Jin P, Wang J, Hu Z. Carboplatin versus cisplatin in combination with etoposide in the first-line treatment of small cell lung cancer: a pooled analysis. BMC Cancer 2021 Dec 7;21(1):1308. [doi: 10.1186/s12885-021-09034-6] [Medline: 34876060]
36. Ogungbenro K, Aarons L. How many subjects are necessary for population pharmacokinetic experiments? Confidence interval approach. Eur J Clin Pharmacol 2008 Jul;64(7):705-713. [doi: 10.1007/s00228-008-0493-7] [Medline: 18483725]
37. Tang F, Ishwaran H. Random forest missing data algorithms. Stat Anal Data Min 2017 Dec;10(6):363-377. [doi: 10.1002/sam.11348] [Medline: 29403567]
38. Shah AD, Bartlett JW, Carpenter J, Nicholas O, Hemingway H. Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. Am J Epidemiol 2014 Mar 15;179(6):764-774. [doi: 10.1093/aje/kwt312] [Medline: 24589914]
39. Harrell FE. Cox proportional hazards regression model. In: Regression Modeling Strategies: Springer; 2015:475-519. [doi: 10.1007/978-3-319-19425-7_20]
40. Suchmacher M, Geller M. Chapter 2: disease frequency measures. In: Practical Biostatistics: A Step-by-Step Approach for Evidence-Based Medicine, 2nd edition: Academic Press; 2021:9-15. [doi: 10.1016/B978-0-323-90102-4.00017-5]

XSL•FO
RenderX

41. O'Brien SM, Dunson DB. Bayesian multivariate logistic regression. Biometrics 2004 Sep;60(3):739-746. [doi: 10.1111/j.0006-341X.2004.00224.x] [Medline: 15339297]
42. Liu LC, Hedeker D. A mixed-effects regression model for longitudinal multivariate ordinal data. Biometrics 2006 Mar;62(1):261-268. [doi: 10.1111/j.1541-0420.2005.00408.x] [Medline: 16542254]
43. Kasi PM, Grothey A. Chemotherapy-induced neutropenia as a prognostic and predictive marker of outcomes in solid-tumor patients. Drugs (Abingdon Engl) 2018 May;78(7):737-745. [doi: 10.1007/s40265-018-0909-3]
44. Miyagi T, Tsuji D, Kawasakai Y, et al. Chemotherapy-induced neutropenia as a prognostic factor in patients with extensive-stage small cell lung cancer. Eur J Clin Pharmacol 2023 Mar;79(3):407-414. [doi: 10.1007/s00228-023-03451-1] [Medline: 36645467]
45. Sachs JR, Mayawala K, Gadamsetty S, Kang SP, de Alwis DP. Optimal dosing for targeted therapies in oncology: drug development cases leading by example. Clin Cancer Res 2016 Mar 15;22(6):1318-1324. [doi: 10.1158/1078-0432.CCR-15-1295]
46. Sawyer M, Ratain MJ. Body surface area as a determinant of pharmacokinetics and drug dosing. Invest New Drugs 2001 May;19(2):171-177. [doi: 10.1023/a:1010639201787] [Medline: 11392451]
47. Zhang X, Guo M, Fan J, et al. Prognostic significance of serum LDH in small cell lung cancer: a systematic review with meta-analysis. Cancer Biomark 2016;16(3):415-423. [doi: 10.3233/CBM-160580] [Medline: 27062698]
48. Singh S, Parulekar W, Murray N, et al. Influence of sex on toxicity and treatment outcome in small-cell lung cancer. J Clin Oncol 2005 Feb 1;23(4):850-856. [doi: 10.1200/JCO.2005.03.171] [Medline: 15681530]
49. Pang HH, Wang X, Stinchcombe TE, et al. Enrollment trends and disparity among patients with lung cancer in national clinical trials, 1990 to 2012. J Clin Oncol 2016 Nov 20;34(33):3992-3999. [doi: 10.1200/JCO.2016.67.7088] [Medline: 27646951]
50. Surbone A, Kagawa-Singer M, Terret C, Baider L. The illness trajectory of elderly cancer patients across cultures: SIOG position paper. Ann Oncol 2007 Apr;18(4):633-638. [doi: 10.1093/annonc/mdl178] [Medline: 17028242]
51. Wormeli P, Mazreku J, Pine J, Damesyn M. Next generation of central cancer registries. JCO Clin Cancer Inform 2021 Mar;5(5):288-294. [doi: 10.1200/CCI.20.00177] [Medline: 33760641]
52. Zanetti R, Schmidtmann I, Sacchetto L, et al. Completeness and timeliness: cancer registries could/should improve their performance. Eur J Cancer 2015 Jun;51(9):1091-1098. [doi: 10.1016/j.ejca.2013.11.040] [Medline: 24393522]
53. Leon L, Kosty M, Jahanzeb M, et al. Effectiveness of bevacizumab exposure beyond disease progression in patients with non-small-cell lung cancer: analyses of the ARIES observational cohort study. Pharmacoepidemiol Drug Saf 2016 May;25(5):569-577. [doi: 10.1002/pds.3948] [Medline: 26748833]
54. Cheng Y, Spigel DR, Cho BC, et al. Durvalumab after chemoradiotherapy in limited-stage small-cell lung cancer. N Engl J Med 2024 Oct 10;391(14):1313-1327. [doi: 10.1056/NEJMoa2404873] [Medline: 39268857]
55. Ahn MJ, Cho BC, Felip E, et al. Tarlatamab for patients with previously treated small-cell lung cancer. N Engl J Med 2023 Nov 30;389(22):2063-2075. [doi: 10.1056/NEJMoa2307980] [Medline: 37861218]
56. Krishnan SM, Friberg LE, Mercier F, et al. Multistate pharmacometric model to define the impact of second-line immunotherapies on the survival outcome of the IMpower131 study. Clin Pharmacol Ther 2023 Apr;113(4):851-858. [doi: 10.1002/cpt.2838] [Medline: 36606486]
57. Krishnan SM, Friberg LE, Bruno R, Beyer U, Jin JY, Karlsson MO. Multistate model for pharmacometric analyses of overall survival in HER2 - negative breast cancer patients treated with docetaxel. CPT Pharmacometrics Syst Pharmacol 2021 Oct;10(10):1255-1266. [doi: 10.1002/psp4.12693]
58. Lombard A, Mistry H, Aarons L, Ogungbenro K. Dose individualisation in oncology using chemotherapy-induced neutropenia: example of docetaxel in non-small cell lung cancer patients. Br J Clin Pharmacol 2021 Apr;87(4):2053-2063. [doi: 10.1111/bcp.14614] [Medline: 33075149]
59. Liu H, Ibrahim EIK, Centanni M, Sarr C, Venkatakrishnan K, Friberg LE. Integrated modeling of biomarkers, survival and safety in clinical oncology drug development. Adv Drug Deliv Rev 2025 Jan;216:115476. [doi: 10.1016/j.addr.2024.115476] [Medline: 39577694]
60. Laurie M, Lu J. Explainable deep learning for tumor dynamic modeling and overall survival prediction using Neural-ODE. NPJ Syst Biol Appl 2023 Nov 18;9(1):58. [doi: 10.1038/s41540-023-00317-1] [Medline: 37980358]

Abbreviations

AE: adverse event
BSA: body surface area
CNS: brain metastasis
CRP: C-reactive protein
csHR: cause-specific hazard ratio
CTCAE: Common Terminology Criteria for Adverse Events
ECOG PS: Eastern Cooperative Oncology Group performance status
GFR: glomerular filtration rate

XSL•FO
RenderX

HR: hazard ratio
IP: irinotecan platinum
LDH: lactate dehydrogenase
OR: odds ratio
OS: overall survival
PE: platinum etoposide
PFS: progression-free survival
RCT: randomized controlled trial
SCLC: small cell lung cancer
TNM: tumor node metastasis

Edited by A Schwartz, M Balcarras; submitted 13.Sep.2025; peer-reviewed by F Pabon-Rodriguez, T Zeng; accepted 03.Mar.2026; published 10.Apr.2026.

Please cite as:
Marzano L, Dan A, Darwich AS, De Petris L, Tendler S, Lewensohn R, Raghothama J, Meijer S
Real-World Evidence of Treatment Outcomes in Small Cell Lung Cancer: A Bayesian Mixed Effects and Competitive Risk Approach
JMIR Cancer 2026;12:e84042
URL: https://cancer.jmir.org/2026/1/e84042
doi:10.2196/84042

© Luca Marzano, Asaf Dan, Adam S Darwich, Luigi De Petris, Salomon Tendler, Rolf Lewensohn, Jayanth Raghothama, Sebastiaan Meijer. Originally published in JMIR Cancer (https://cancer.jmir.org), 10.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Piloting a Clinical Decision Support System for Unintended Weight Loss in Primary Care: Mixed Methods Study on Early Cancer Detection

Javiera Martinez-Gutierrez[1,2,3], MD, MPH; Sophie Chima[1,2], PhD; Barbara Hunter[1], PhD; Alex Lee[1], PhD; Lucas De Mendonca[1], MPhil; Deborah Daly[1,4], MBA; George Fishman[4], MBA; Fong Seng Lim[5], MD; Benny Wang[6], PhD; Craig Nelson[7,8], PhD, FRACP, MBBS; Brian Nicholson[9], PhD; Jo-Anne Manski-Nankervis[1,10,11], PhD; Jon Emery[1,10], Ma, MBBCHIR, DPhil

[1]Department of General Practice and Primary Care, Melbourne Medical School, Faculty of Medicine, Dentistry, and Health Sciences, The University of Melbourne, 3rd Floor Medical Building, Cnr Grattam Street and Royal Parade, Melbourne, Victoria, Australia

[10]Primary Care and Family Medicine Department, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Chile

[11]NHG Polyclinics, National Healthcare Group, Singapore, Singapore

[2]Collaborative Centre for Genomic Cancer Medicine, The University of Melbourne, Melbourne, Victoria, Australia

[3]Department of Family Medicine, Faculty of Medicine, Pontificia Universidad Catolica de Chile, Santiago, Chile

[4]Primary Care Collaborative Clinical Trials Group, Melbourne, Victoria, Australia

[5]National University Polyclinic, Singapore, Singapore

[6]Singapore Primary Care Cancer Network (SPriNT), Singapore, Singapore

[7]Western Health Chronic Disease Alliance, Western Health Melbourne, Melbourne, Australia

[8]Department of Medicine – Western Health, The University of Melbourne, Melbourne, Victoria, Australia

[9]Department of Primary Care Health Sciences, University of Oxford, 10. Nuffield, Oxford, England, United Kingdom

Corresponding Author:
Javiera Martinez-Gutierrez, MD, MPH
Department of General Practice and Primary Care, Melbourne Medical School, Faculty of Medicine, Dentistry, and Health Sciences, The University of Melbourne, 3rd Floor Medical Building, Cnr Grattam Street and Royal Parade, Melbourne, Victoria, Australia

Abstract

Background: Delayed cancer diagnosis leads to poorer outcomes. Unintended weight loss (UWL) is a nonspecific symptom associated with cancer and other serious conditions, which can make it complex to identify the underlying cause. Clinical decision support systems (CDSSs) can provide evidence-based recommendations to facilitate timely investigation and diagnosis.

Objective: This study aimed to pilot the implementation of a CDSS for improving early cancer detection in primary care patients with UWL.

Methods: Five practices reviewed patients identified by the UWL CDSS. Staff were interviewed on the acceptability and feasibility of the CDSS. Interviews were analyzed thematically using 2 relevant frameworks: the acceptability of health care interventions and the sociotechnical model for evaluation of digital interventions framework. Clinical audits assessed the correct identification of UWL, follow-up rates, and patient characteristics.

Results: Of 60 patients identified by the CDSS as potentially having UWL, 36 (60%) had true UWL. Among the misclassified cases, most patients (16/55, 30%) were intentionally trying to lose weight; this intention was documented as free text in the clinical notes for 98% (58/60) of patients and in a structured field for only 20% of patients. Of the 36 patients, 5 (14%) were actively recalled by their practices for further follow-up; the others (31/36, 86%) were not recalled, as most were deemed already under appropriate follow-up. By 6 months, 94% (34/36) of the cohort had received follow-up care regardless of whether they had been formally recalled. One-third (12/36, 33%) of patients had no additional symptoms, while 36% (13/36) had recorded additional abdominal symptoms. Mental health conditions accounted for 19% of diagnoses linked to UWL consultation. Practice staff were generally receptive to the UWL CDSS concept. It was particularly appreciated by practices with strong quality improvement processes and prior CDSS experience. A high proportion of misclassified patients, poor workflow integration, and conflicting patients' agendas were cited as implementation barriers.

Conclusions: Our study revealed challenges and potential benefits of implementing a CDSS for identifying patients with UWL at risk of cancer in primary care. While integration issues and misclassification of patients were noted, high rates of follow-up care were observed regardless of the CDSS. These high follow-up rates prompt consideration of whether they reflect Australian

XSL•FO
RenderX

primary care and whether a CDSS with these characteristics is the most appropriate approach to support early cancer detection. Future research should focus on improving CDSS integration with existing workflows, enhancing correct identification through access to clinical notes and the use of more sophisticated digital methods (eg, AI). These findings highlight the potential of CDSS in improving patient care and the complexities involved in their successful implementation.

(JMIR Cancer 2026;12:e90885) doi:10.2196/90885

KEYWORDS

digital health; primary care; cancer early detection; general practice; clinical decision support

Introduction

Early detection of cancer is crucial for improving patient outcomes, offering several significant benefits. It can substantially enhance a patient's chance of survival, broaden the range of available treatment options, and contribute to an improved quality of life [1,2]. Moreover, timely diagnosis can help reduce the overall burden on health care systems by potentially decreasing the complexity and cost of treatments required for advanced-stage cancers [3]. General practice plays a pivotal role in this early detection process, often serving as the first point of contact for patients experiencing new signs or symptoms of cancer [4,5]. However, the diagnosis of cancer can be particularly challenging in many cases, given the nonspecific nature of presenting symptoms. These symptoms can often be associated with a number of other, more common and less severe conditions, making it difficult for general practitioners (GPs) to immediately identify potential malignancies [6].

Unintended weight loss (UWL) is an example of a nonspecific symptom associated with cancer. While UWL can also be associated with various benign conditions, including depression, thyroid disorders, and gastrointestinal diseases, it is also a potential indicator of several cancer types [6,7]. The diagnostic complexity is compounded by the fact that UWL can initially occur in isolation, without accompanying clear symptoms that might point toward a specific underlying cause [8]. This absence of additional clinical indicators makes it particularly challenging for health care professionals to determine the optimal route of investigation. The nonspecific nature of UWL as a symptom underscores the need for a comprehensive and systematic approach to patient evaluation, potentially incorporating advanced diagnostic tools and risk assessment models.

Given the diagnostic complexity of nonspecific symptoms' association with many cancers, clinical decision support systems (CDSSs) are being developed for use in general practice [9]. These tools are designed to assist GPs in making more informed decisions by integrating algorithms underpinned by patient data and clinical guidelines [10]. In the context of early cancer detection, CDSS can potentially streamline the diagnostic process by identifying patients who require further investigation or referral [11]. For patients presenting with nonspecific symptoms such as UWL, these tools could be particularly valuable in stratifying risk and supporting decision-making.

Studies evaluating CDSS in practice have demonstrated increased adherence to guidelines and successful diagnostic support [11]. However, the majority of developed tools fail to achieve widespread implementation, and those that are implemented often face low uptake rates [12]. While there are significant opportunities to support early cancer detection, designing tools that are useful, acceptable, and seamlessly integrated into primary care workflows is a complex, multistep process. This process necessitates continuous input and refinement from end users throughout the development cycle to ensure relevance and usability [13,14]. This study aimed to evaluate a novel CDSS designed for the early detection of cancer in patients presenting to general practice with UWL. Specifically, the study sought to assess whether the tool was acceptable to practitioners, met the diverse needs of general practice, and effectively supported the delivery of high-quality, evidence-based care.

Methods

Design

This study used mixed methods to evaluate the implementation of a complex CDSS for patients with UWL, using patient data collected in clinics through interviews with participating general practices and an electronic health record (EHR) audit.

The Future Health Today Tool

Future Health Today (FHT; The University of Melbourne) is a software platform that has been developed for use in general practice, streamlining the identification and management of chronic diseases. While the development of the tool has been described elsewhere, it underwent an extensive co-design process with GPs and consumers [15,16]. FHT is integrated with the EHR and has 2 core components: First, it is a CDSS that provides proactive, guideline-based recommendations at the point of care; the second component is a web-based portal, which supports quality improvement (QI) activities in general practice, has the ability to audit and recall patient groups within the practice population, and provides access to education and resources.

The UWL Module

The FHT tool supports a wide range of conditions, and different modules are developed for use in FHT [17,18]. This study explores the use of the FHT UWL module, which flags patient records with UWL indicated in the EHR in the past 6 months. Development of the UWL module was extensive, beginning with the development and validation of phenotyping and risk prediction algorithms to identify patients with "unintended weight loss" terms in structured EHR fields (reason for visit, diagnosis, and prescriptions) while also excluding those with indications of intentional weight loss. Additional exclusion

XSL•FO
RenderX

criteria, including recent cancer diagnoses, intentional weight loss reasons (eg, bariatric surgery or weight loss prescriptions), and/or increasing weight measurements, were further added [19]. We did not use actual weight loss as an inclusion criterion, given the low prevalence of weight measurements in Australian general practice data [20]. Algorithm criteria (age, sex, blood test results, and smoking status) determined the patient's risk profile and provided the recommendation delivered to the GP.

The phenotyping and validation process and the risk profile calculations have been described elsewhere [19]. Briefly, after applying inclusion and exclusion criteria to identify patients with UWL, the module provided 3 types of recommendations according to cancer risk:

1. For patients with low cancer risk (<2%): a recommendation to take new weight measurements and a reminder of the most common differential diagnosis according to age and sex
2. For patients with moderate to high risk (>2%) and absent or normal blood test results: recommendation to ask for symptoms and order (additional) pathology tests
3. For patients with moderate to high risk (>2%) and abnormal blood test results: recommendations to follow up with further testing (imaging, scopes, and referrals)

Figure 1 shows a simplified diagrammatic representation of the rule-based algorithm rules.

The content and wording of the recommendations and the supporting resources were developed with input from GP and consumer and were informed by evidence from a previous scoping review [21]. An example of a recommendation as it appears in the EHR is shown in Figure 2.

Figure 1. Unintended weight loss (UWL) module algorithm rules and recommendations. Black boxes represent the algorithm processes; green boxes are relevant recommendations; blue box provides further relevant extra information; red boxes represent no further recommendations. Alb: albumin; AP: alkaline phosphatase; CRP: C-reactive protein; Hb: hemoglobin; PLT: platelets; SR: sedimentation rate.

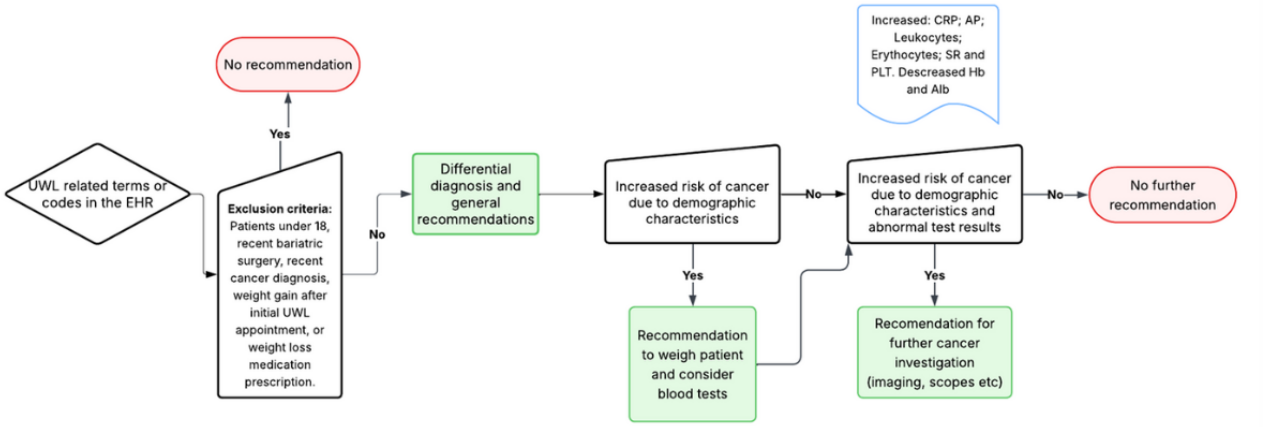
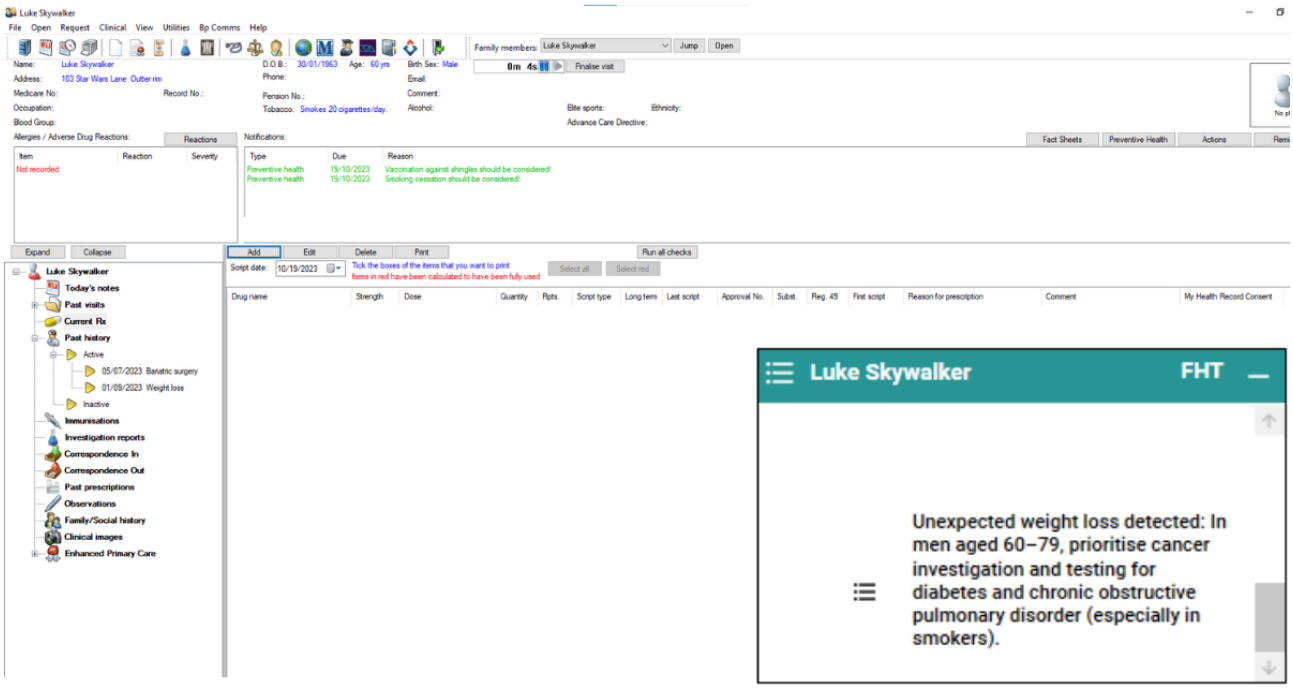


Figure 2. Example of a recommendation as it appears in a medical record (hypothetical patient).



XSL•FO
RenderX

Study Participants, Recruitment, and Implementation

Recruitment of general practices occurred by approaching practices participating in the FHT program (ie, practices that already had FHT installed). Practice managers and/or practice owners were contacted and provided with study information. All recruitment documentation was created with assistance from the FHT Consumer and General Practice Advisory Group. In agreeing to participate in the study, practices were required to review a list of patients with UWL, nominate 1 person from the practice to participate in an interview, and partake in a clinical audit.

We invited, via email, practices that had at least five patients identified by the UWL algorithm. Practices recruited to participate in the pilot UWL implementation study had the FHT UWL module switched on for at least 6 months. At study initiation, the FHT team created a list of patients (also referred to as "cohort") who met the criteria for UWL, which could be accessed via the FHT web-based portal. Practices were also offered a 30-minute, Zoom (Zoom Communications, Inc)–based training session once the module was switched on. Each practice nominated an appropriate person (eg, GP, general practice nurse, and QI lead) or team of people to review the cohort of patients with UWL. The purpose of the cohort review was to determine which patients should be recalled for further follow-up or whether the recommendations should be deferred. After review of the patient list, practices were asked to fill in a log with the kind of action performed (eg, patient recalled or deferred) and the reasons for this action (eg, weight loss was intentional, patient is already being followed up, and patient requires further investigation). After the cohort review was complete, practices were invited to use FHT as they chose for the remainder of the study.

Data Collection

Interviews

One month after the implementation of the UWL module and completion of the cohort review, practice staff who reviewed the cohort of patients flagged with UWL were contacted to participate in an interview. At least one person from each general practice participated in a semistructured interview conducted via Zoom. Interview schedules (Multimedia Appendix 1) were informed by relevant literature and prior qualitative work conducted as part of the FHT program [22,23]. In the interview, staff were asked about the FHT technology generally, their experience using the UWL module to review and recall patients, and the UWL recommendations. The interviews explored the themes of implementation (including barriers and facilitators to use), the content and clinical appropriateness of the UWL recommendations, and usability, acceptability, and feasibility of the module in the longer term. All interviews were audio recorded for analysis and were transcribed using Otter.AI. JMG conducted the interviews. Interview duration was 30 to 45 minutes.

JMG is a female mid-career researcher and academic GP with experience in qualitative research. SC is an early-career researcher with expertise in FHT and qualitative research. JMG did not have a previous relationship with any of the interviewees. SC had previous contact with 2 practice staff from 1 practice as a research coordinator for a previous study using FHT. We did not provide information regarding personal goals or additional reasons for doing the research other than the description in the plain language statement. Participants were not asked to review interview transcripts or study findings, and no repeat interviews were conducted. Field notes were not recorded during or after the interviews. To ensure methodological rigor and transparency, the presentation of findings adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Checklist 1).

Clinical Audits

An EHR audit was conducted approximately 1 to 4 months after the creation and review of the UWL cohort. EHR data of patients who were flagged by the FHT platform as having UWL were audited by a clinical researcher using a data extraction template. The extraction template was piloted and refined by clinician researchers in the team before data collection (Multimedia Appendix 2). Data extracted from the EHR included age, sex, details on the record of UWL and all UWL-related free text, prescriptions, laboratory results, referrals, images, weight measurements, and any UWL-related follow-up and were entered into the REDCap platform [24]. The time frame of extracted data included the 6 months prior to the initial UWL presentation until the date of the clinical audit. Due to the remoteness or location of participating practices, clinical audits were conducted using TeamViewer, a secure remote access software hosted by the University of Melbourne.

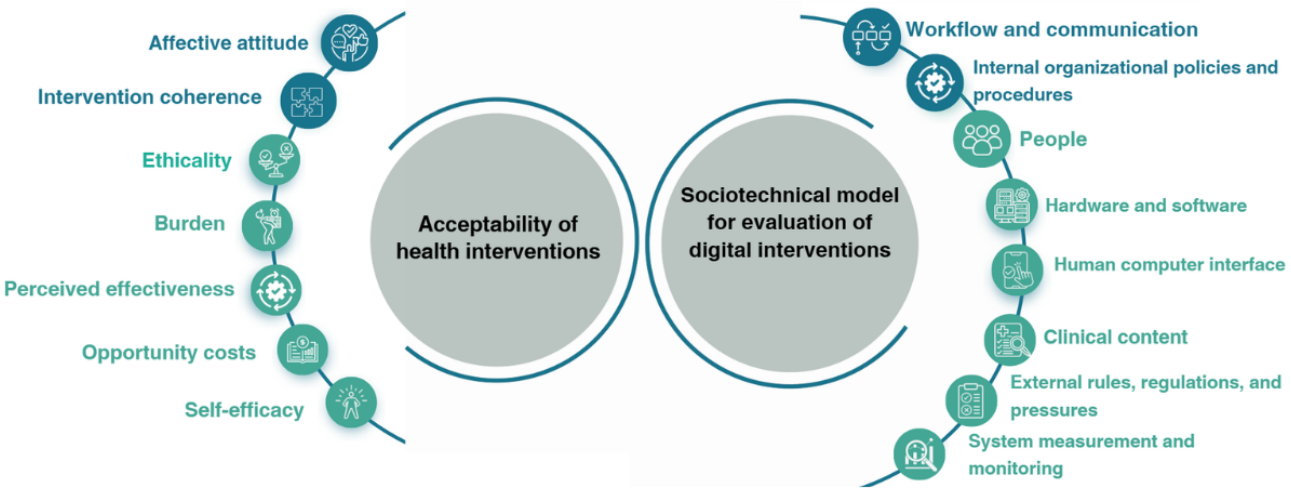
Data Analysis

Interviews

Interview transcripts were uploaded into NVivo 14 (Lumivero) and thematically analyzed independently by 2 researchers (JMG and SC) [25]. We used deductive thematic analysis using 2 overarching frameworks in the analysis: acceptability of health care interventions [26] and sociotechnical model for evaluation of digital interventions framework [22] (Figure 3).

XSL•FO
RenderX

Figure 3. Frameworks used in the analysis of semistructured interviews with general practice.



We chose these 2 frameworks for their health care–specific focus and comprehensive scope, encompassing both technical and human factors in CDSS implementation. These recent frameworks offer a multidimensional approach to assess user perceptions and evaluate digital interventions, allowing us to identify a broad range of barriers and facilitators for safe CDSS testing in primary care. In addition, inductive analysis was conducted to create de novo themes developed from the data. The qualitative themes mapped onto the 2 frameworks are reported based on dominant themes arising from the interviews. Differences in interpretation were discussed and resolved by consensus. Input was sought from the wider research team, including consumer investigators, in the development and interpretation of themes.

Main themes analyzed are shown in blue, which are adapted from studies by Sekhon et al [26] and Sittig and Singh [22].

Clinical Audits and Implementation Data

Data collected in the EHR audit were uploaded to REDCap and then downloaded into a master Excel (Microsoft) spreadsheet to be reviewed, cleaned, and analyzed by 2 researchers (JMG and SC), including 1 clinician (JMG). Disagreements were discussed with the wider clinical and implementation team. Data collected via the EHR audit and clinical review log for each practice were analyzed using descriptive statistics. This included determining the binary outcomes: “Does this patient have unexpected weight loss?” (yes/no) and “Has this patient been followed up for UWL?” (yes/no). True and false positives of UWL were calculated using the information collected in the EHR audit following a predetermined set of criteria (Multimedia Appendix 3). For patients with true UWL, descriptive statistics were used to describe follow-up variables in the coded and free-text fields (diagnosis, symptoms, investigations, and referrals).

Ethical Considerations

Ethical approval was granted by the University of Melbourne Human Research Ethics Committee STEMM 1 (Ethics ID: 25308). Participants were provided with a plain language statement explaining the study. Informed consent was obtained from all participating general practices. Practices were paid Aus $1000 (approximately US $660) for participation. Participation was voluntary, and all interview participants were reimbursed for their time (Aus $100, equivalent to US $67).

Results

Practice Characteristics

We recruited 5 general practices from a pool of 40 practices that had the FHT program installed. Practices were located in Queensland and Victoria, Australia. Four practices were located in metropolitan areas, and 1 practice was located in a rural area based on the Australian Modified Monash Model [27]. The mean number of patients identified as having UWL 6 months previous to cohort creation was 13 (SD 5) patients per practice. Practice characteristics are described in Table 1.

XSL•FO
RenderX

Table . Characteristics of 5 general practices participating in pilot study (n=5).

Practice characteristics	Values
Number of active patients[a] in each practice, mean (SD)	17,619 (3368)
Number of active patients by reported gender, mean (SD)	
Male	7623 (1325)
Female	9505 (1891)
Other	491 (310)
Number of active patients aged ≥40 y, mean (SD)	8933 (1905)
Number of physicians per practice, mean (SD)	10.4 (5)
Rurality classification of practice[b], n (%)	
Metropolitan	4 (80)
Rural	1 (20)
State or location, n (%)	
Victoria	2 (40)
Queensland	3 (60)
Cohort size, mean (SD)	13 (5)

[a]The Royal Australian College of General Practitioners defines an "active patient" as "a patient who has attended the practice 3 or more times in the past 2 years" [28].

[b]Rurality classification based on the Modified Monash Model [27].

Practice Staff Interviews

We conducted 5 interviews with 7 practice staff. Interviews were conducted via Zoom after the list of patients had been reviewed by practice staff and at least a month had passed to give patients the opportunity to come back for a follow-up appointment if they had been recalled by the practice.

Participant characteristics are presented in Table 2.

XSL•FO
RenderX

Table . Characteristics of interview participants.

Characteristics	Values, n (%)
Age	
20 - 29	1 (14)
40 - 49	3 (43)
50 - 59	1 (14)
≥60	1 (14)
Missing	1 (14)
Gender	
Male	1 (14)
Female	6 (86)
Role	
GP[a]	1 (14)
PN[b]	3 (43)
Pharmacist[c]	1 (14)
PM[d] or practice owner	1 (14)
Admin[c]	1 (14)
Practice setting	
Metro	5 (71)
Rural	2 (29)

[a]GP: general practitioner.

[b]PN: practice nurse.

[c]Role also involves leading the quality improvement activities in the practice.

[d]PM: practice manager.

We first describe "affective attitude" to contextualize the staff's general attitude toward the CDSS. Furthermore, 3 main deductive themes became evident from our analysis using the 2 frameworks: internal organizational policies, procedures, and culture; and workflow and communication and intervention coherence. Remaining relevant themes from the frameworks are summarized, with exemplary quotations provided in Multimedia Appendix 4.

Affective Attitude

Overall, participants liked the concept of the tool. For quality assurance purposes, the cohort tool, with the ability to create lists of patients with specific conditions and the detailed information it provided on the patient, was found to be very useful. Regarding the UWL cohort, they particularly liked that the cohorts were composed of only a few patients, as expected given the low prevalence (1.5%) of UWL reported in the literature.

Clinical and administrative staff generally favored QI initiatives and supporting software, with 1 staff member reporting frequent use of FHT in management plans and health assessments.

One clinical staff member mentioned finding both features useful:

> *I like the way that it pops up a reminder in real time when we open their record. And it's really useful to be able to develop the cohorts for auditing purposes and for purposes of developing, you know, PDSA (Plan, Do, Study, Act) cycle quality improvement activities and things like that. So, so overall, I've found FHT really, a really useful innovation.* [Clinical staff, practice 2]

> *I like when you click on the patient's name in the cohort, you know you can see the meds, there's a snapshot of their history. I like the MBS (Medicare Benefits Schedule) items that might apply, because part of my job is making sure we're sustainable and maximizing those, those chronic disease item opportunities.* [Clinical staff, practice 1]

Intervention Coherence

Despite FHT being installed for at least a year in all 5 practices, 5 (70%) participants were not well acquainted with the cohort tool in FHT. Even after co-design efforts, provision of written materials, and virtual training offers, several challenges emerged in the implementation of the UWL module. The inclusion and exclusion criteria of the algorithm were not well understood. Users were uncertain about how patients progressed through the patient management system (from "awaiting action" to

XSL•FO
RenderX

"deferred," "actioned," or "no further action needed"), indicating confusion about this component of FHT.

After reviewing the lists, clinicians identified that several patients on the list were intentionally trying to lose weight (see "Clinical Audits" section) or had only one recorded weight measurement and staff were confused by the inclusion of these patients. Technical issues further complicated matters, as the tool malfunctioned during use by the first practice. In this case, the practice staff reviewed the cohort and actioned the patients accordingly (marking them as "deferred": no action needed or "recalled"), but the system did not save these actions, causing all patients to remain in the "awaiting action" category, causing additional confusion.

Despite these challenges, there were positive aspects. The tool's visual elements and patient summaries were seen as helpful facilitators. Most users found reviewing patients in the cohort straightforward:

> *I didn't find it too hard. It was - once you went into the patient's file it was pretty self-explanatory if you just kept going through the flow of the appointments and what the doctors - because you can see where it said weight loss in the reason for visit, if you sort of read from then on, and then go back if you have to, but, it sort of flowed pretty well, so you could understand.* [Clinical staff, practice 3]

> *First of all, of the 19 patients, some of them, I'm not sure why they popped up on the list, some of them had only ever had one weight recorded. So I'm not quite sure how the tool came to conclude that they'd lost weight.* [Clinical staff, practice 2]

Internal Organizational Policies, Procedures, and Culture

Internal processes arose as the most influential factor in the use of the cohort tool.

GPs typically document most patient information in the free-text clinical notes field of the EHR. However, the CDSS algorithm is currently unable to read these notes, relying only on structured text fields such as "reason for visit" or "diagnosis," creating a substantial barrier to its effective use. This mismatch between clinician documentation and the CDSS data input requirements limited the tool's ability to provide comprehensive support.

In 2 practices, the manager asked GPs whether they would be open to a CDSS system that would issue reminder pop-ups at the point of care when a patient presented. They collectively opted not to have it installed as they were concerned about alert fatigue, showing a lack of understanding of how frequently this tool for this algorithm would appear, as UWL is a relatively uncommon presentation.

1. *Dedicated personnel:* Having a dedicated PN or allocated person, such as a pharmacist or practice manager, for QI activities facilitated the use of the cohort tool. In these cases, the integration with the practice workflow was better achieved, ensuring more consistent and effective use of the system. However, designated resources have a cost associated, and many practices do not have sufficient staff available to undertake this type of role.
2. *Existing QI systems:* Some practices already had seemingly effective systems in place for data searches and QI. Two practices, which had the same owner and manager, were very comfortable using the search tool embedded in their own EHR. Although the EHR search tool was simple and would not be able to flag patients with UWL nor identify risk of cancer, the practice manager thought it was good enough for their QI requirements, making the tool seem unnecessary for their needs. They also did not see the use of the point-of-care tool as they had determined it would result in alert fatigue. These established workflows were seen as competing with the new CDSS, depending on how well acquainted the staff were with FHT and whether this UWL algorithm integrated with existing processes.
3. *Varied models of use:* In 3 of the 5 practices, clinical staff conducting audits did not directly engage with the CDSS but instead worked from patient lists provided by practice managers or research coordinators who knew how to use FHT. In these practices, FHT was understood to be mostly a research tool, rather than a tool for everyday QI activities. This indirect use of the system led to a disconnect between the tool's capabilities and the clinical decision-making process in the clinic. The 2 remaining practices were well acquainted with the tool and were already using it for QI and clinical care activities for other conditions on their own account.
4. *Time allocation:* Time allocation for cohort review varied significantly between roles. Nurses and coordinators had allocated time for QI activities and worked within office hours, while the GP (similar to most GPs in Australia), lacking dedicated time, resorted to after-hours completion, impacting workflow integration.
5. *User familiarity and role integration:* In practices where the same person reviewed the cohort and assessed clinical relevance, there tended to be better integration of the module within the practice workflow. These individuals, often practice managers with clinical roles, were typically more familiar with the tool and could more effectively coordinate follow-up actions with GPs. This familiarity and role integration led to more streamlined use of the CDSS.

Quotes exemplifying each one of these factors or processes can be found in Table 3.

XSL•FO
RenderX

Table . Internal factors or processes affecting implementation.

Internal factor or process	Selected quotation
Organization culture and staff practices	• "No one [GP] uses this drop-down system codes [in the reason for visit field in the EHR] that takes one second" (clinical staff, practice 3). • "No, no, I haven't got [FHT] on for the GPs at all. And the reason for that is most of my doctors are anti pop ups" (practice manager, practice 5).
Dedicated personnel	• "So as part of being the nurse coordinator, I do take on a lot of the quality improvement in the practice, so I'm one of the leads on that part" (clinical staff, practice 3).
Existing quality improvement systems	• But most of our data searches in that we actually use [our EHR]. So which is, which is the software that we use. So it's much easier for us to just be working on [our EHR]" (practice manager, practice 4).
Varied models of use	• "In the program, so I went, basically made a list of all the patients and all the like possible causes. I gave it to [PM], and then kind of just assessed, like, what was going on with the patient, whether they needed to come back to see the doctor, or most of them were already, like in discussion with the doctor about that weight loss" (practice manager, practice 5).
Time allocation	• "I think they just did that out of hours, you know, an extra job to do with an evening, because I sent them, I sent everybody, at least a round of their relevant patients, and asked them to have a look" (clinical staff, practice 2).
User familiarity and role integration	• "So the portal, so yep, and like the recalling or deferring patients, so often looking up a lot of if it's the clinical side, I'll tend to look up all the clinical data and stuff like that, and refer and defer and review patient, bring it to the attention of the doctors if they're coming in for another appointment" (clinical staff, practice 3).

Workflow and Communication

Workflow and communication factors were described as crucial determinants in the use of the module. This theme, closely intertwined with internal processes mentioned earlier, stood out for both its prevalence and the depth of insights it provided.

The ease of integrating the tool into existing workflows varied among practices. Those with a special interest in data analysis and QI found it easier to embed the cohort tool into their routines. However, the lack of integration of the system with the practices' message software was a common issue, hindering direct action on strategies such as reminders and recalls.

The process of reviewing the cohort was generally consistent and considered relatively easy across practices.

1. Adaptations and workarounds: Practices developed their own methods for managing the cohort and relaying information, such as using spreadsheets or internal notes. The fact that FHT was adaptable for this was valued and highlights the need for flexible systems that can accommodate various practice styles.
2. Communication challenges: Clinical and administrative staff mentioned that communicating the need for recall to patients could prove challenging, especially when the issue (eg, UWL) was not a priority for patients or when they were unaware of the problem. This indicates that clinician judgment always sits on top of any CDSS and that the provision of a list does not automatically mean it is clinically relevant or preferred to actively recall patients.
3. Training and understanding: Three practices did not fully use all CDSS features (eg, pop-ups) due to lack of experience or training (one practice) or practice decision not to use the pop-up feature to avoid alert fatigue (2 practices). This affected their understanding of the tool's capabilities and its integration into their workflow.
4. Role-specific challenges: The workflow varied by role, with a GP reviewing pop-ups during consultations but not always acting on them due to existing audit processes. Practice nurses sometimes struggled with determining follow-up needs, highlighting the importance of role-specific training and clear guidelines.

Exemplary quotes:

> *[FHT] It doesn't talk back to [our EHR] to add any reminders or add any actions, and knowing how hard [our EHR] is to integrate with...* [Clinical staff, practice 1]

> *I think it's useful, but it's still up to the GPs to sort of use it more? Yeah, I think that's our, that's our biggest struggle is to get more of them using it.* [Clinical staff, practice 3]

XSL•FO
RenderX

> *Or if the patient has had no nurse contact, she may not have seen that patient at all. So it would have been a very awkward conversation for her to actually call the patient and go, Hey, how's it going? Are you trying to lose weight?* [Administrative staff, practice 4]

Clinical Audits

The algorithms identified 62 patients with possible UWL. We extracted data from a total of 60 patients. Two patients were excluded from analysis as records could not be accessed, given (1) no longer an active patient and (2) no weight loss recorded as a reason for the visit or in their past history. The tool correctly identified 36 of 60 (60%) patients with UWL established by clinical assessment of the information in the EHR.

Patients with UWL confirmed in the EHR were older and equally distributed by sex, while patients with intentional weight loss recorded in the EHR were younger and mostly women. Intention to lose weight was recorded as free text in the clinical notes in 98% of cases, with only 20% of patients having intention recorded in a structured field (reason for visit or diagnosis). Sixty-one percent (22/36) of patients with UWL were at moderate to high risk of cancer. Five (8%) of 60 patients were recalled for further investigation. Details of patients' characteristics and CDSS deployment can be found in Table 4.

Table . Patient characteristics and related clinical decision support system recommendations.

	All patients	Patients with UWL[a]	Patients with intended weight loss
Number of patients, n (%)	60 (100)	36 (60)	24 (40)
Age, median (IQR)	57.5 (35.5)	68.5 (130.5)	44 (21.5)
Sex, n (%)			
Males	27 (45)	18 (50)	9 (37.5)
Females	32 (53)	17 (47)	15 (62.5)
Missing	1 (2)	1 (3)	0 (0)
Recommendations[b], n (%)			
Patients at low risk of cancer	20 (33)	9 (25)	11 (46)
Patients at moderate and high risk of cancer	30 (50)	22 (61)	8 (33)
No recommendations	10 (17)	5 (14)	5 (21)
Patient recalled for further follow-up, n (%)			
Yes	5 (8)	5 (14)	0 (0)
No[c]	55 (92)	31 (86)	24 (100)
Where was intention to lose weight recorded[d], n (%)			
Reason for visit	10 (17)	5 (14)	5 (21)
Diagnosis	2 (3)	1 (3)	1 (4)
Clinical notes	59 (98)	35 (97)	24 (100)
Other (inferred from test ordering)	1 (2)	1 (3)	0 (0)

[a]UWL: unintended weight loss.

[b]Refer to Figure 1 for definition of risk and recommendations.

[c]This includes patients for whom the practice determined did not need to be recalled for further follow-up (eg, deferred for review until a later date, already under investigation, known cause of UWL). Reasons for deferral or no action taken are reported in Table 5.

[d]Intention may have been recorded in more than one electronic health record field.

XSL•FO
RenderX

Table . Reasons for deferral in 55 patients identified in unintended weight loss (UWL) cohort (n=55).

Deferral reasons	Values, n (%)
UWL with known cause	16 (30)
Incorrectly identified according to practice	
Intentional weight loss	16 (30)
Data error	1 (1.8)
No evidence of weight loss	6 (11)
UWL with unknown cause and already under investigation	7 (12.7)
No reason for deferral given or missing	9 (14.5)

Ninety-two percent (55/60) of all patients were deferred, as practices deemed no further follow-up was needed. Reasons for deferral were divided into three main groups: (1) unintentional weight loss with known cause, (2) patient incorrectly included in cohort (eg, intentional weight loss, data error, and increased weight), and (3) unintentional weight loss with cause unknown being followed up already. Table 5 presents the frequency of reasons for referral.

A summary of clinical characteristics and follow-up of the 36 patients with UWL is described in Table 6 (more details in Multimedia Appendix 5). Following their initial UWL consultation, the vast majority of patients with UWL (34/36, 94%) received comprehensive follow-up care. This care encompassed subsequent GP appointments, specialist referrals, various investigations, and/or imaging studies as needed. On average, patients attended a median follow-up of 8 (IQR 10) visits after their initial UWL consultation until the audit date.

XSL•FO
RenderX

Table . Summary of clinical characteristics and follow-up of patients with confirmed unintentional weight loss (UWL; n=36 patients).

Follow-up and clinical characteristics	Values
Number of visits after initial UWL visit (6 mo follow-up period), median (IQR)	8 (10)
UWL discussed again after initial UWL visit (6-mo follow-up period), n (%)	
Yes	20 (56)
No	16 (44)
Number of patients followed up after initial UWL presentation?[a], n (%)	34 (94)
Other presenting symptoms at index date[b], n (%)	
None	12 (33)
Nonspecific (eg, appetite loss and fatigue)	17 (47)
Gastrointestinal (eg, abdominal pain and vomiting)	13 (36)
Mental health (stress, anxiety, mood, or insomnia)	7 (19)
Respiratory (asthma exacerbation and cough)	3 (8)
Other (eg, oral and urinary symptoms)	8 (22)
Investigations ordered by GP[c], n (%)	
Yes	31 (86)
No	5 (14)
Investigation ordered[d], n (%)	
Hematology (full blood count and iron studies)	53 (147)
Metabolic (eg, HbA_{1c}[e] and thyroid function)	90 (250)
Cancer markers (eg, Ca12, CEA, and Ca19-9)	8 (22)
Celiac serology (celiac antibodies)	2 (6)
Vitamins (eg, B12 or folate and vitamin D)	25 (70)
Inflammatory markers (CRP[f] and ESR[g])	28 (78)
Other (eg, urine, STIs[h], and virus)	37(101)
Imaging referral, n (%)	
Yes	22 (61)
No	14 (39)
Imaging test ordered[i], n (%)	
CT[j] scan	13 (36)
Colonoscopy	5 (14)
Endoscopy	5 (14)
X-ray	7 (19)
Ultrasound	3 (8)
Other	2 (6)
Referrals, n (%)	
Yes	20 (56)
No	16 (44)
Referrals (detail)[k], n (%)	
General surgeon for colonoscopy	4 (11)
General surgeon for gastroscopy	3 (8)

XSL•FO
RenderX

Follow-up and clinical characteristics	Values
Gastroenterology for endoscopy	2 (6)
Other	16 (44)
Related diagnoses within 6 months of UWL appointment to audit date recorded[l], n (%)	
Yes	20 (56)
No	16 (44)
Comorbidities[m], n (%)	
Mental health conditions	7 (19)
Cancer	4 (11)
Chronic obstructive pulmonary disease	3 (8)
Other	9 (25)

[a]Includes another consultation where UWL was discussed and/or any investigations, imaging, or referral after the index date.

[b]Patients may have had more than one symptom recorded at or after index date.

[c]GP: general practitioner.

[d]Patients may have had more than 1 investigation ordered.

[e]HbA_{1c}: hemoglobin A_{1c}.

[f]CRP: C-reactive protein.

[g]ESR: erythrocyte sedimentation rate.

[h]STI: sexually transmitted infection.

[i]Some patients had more than one imaging request.

[j]CT: computed tomography.

[k]Patients may have been referred to more than 1 specialty.

[l]"Related diagnosis" defined by the practice GP in the electronic health record as a possible reason for UWL recorded within 6 mo of UWL (new or preexisting).

[m]Patients may have had more than 1 diagnosis recorded.

One-third of patients (12/36, 33%) presented to the first UWL visit with UWL as their only symptom. Almost half of patients (17/36, 47%) presented with other nonspecific symptoms such as appetite loss and lethargy, and the most common accompanying specific symptoms were gastrointestinal symptoms (13/36, 36%), followed by mental health symptoms (7/36, 19%).

Eighty-six percentage (n=31) of patients had been investigated with blood tests. The most common tests ordered were full blood count tests (n=30, 83%), liver function tests (n=30, 83%), creatinine testing (n=25, 69%), iron studies (n=23, 64%), calcium testing (n=22, 61%), and thyroid function tests (n=18, 50%).

Diagnostic imaging and endoscopic procedures were frequently used in the investigation of UWL. Nearly two-thirds (22/36, 61%) of patients underwent imaging studies or endoscopic examinations. Of these, more than half (13/22, 59%) underwent CT imaging, which included abdominal, lung, and/or pelvic regions.

While direct colonoscopy referrals were relatively low (5 patients), a more common approach was to refer patients to specialists for consideration of endoscopic procedures. One-quarter (9/36, 25%) of patients were referred to either a gastroenterologist or, in rural practices, a general surgeon for consideration of upper and/or lower endoscopy.

More than half of the patients (20/36, 56%) received a diagnosis related to their UWL, as determined by their GP and documented in the EHR. The most prevalent diagnoses associated with UWL were mental health conditions, followed by cancer.

Four (11%) of 36 patients were diagnosed with cancer, including chronic lymphocytic leukemia, myeloma, breast cancer, and renal cancer. However, only the renal cancer case was diagnosed following the UWL consultation. The other 3 (8%) patients should have been excluded from the study based on the predefined algorithm, as these diagnoses were made prior to the UWL visit and within the preceding 5 years.

Discussion

Principal Findings

Our study aimed to pilot the implementation of a CDSS to identify patients with UWL at risk of cancer in primary care settings. This initiative involved extensive preliminary work, including the use of primary care databases to identify patients with UWL and calculate their cancer risk using a UK-developed risk prediction model [19]. We also co-designed recommendations and resources and tested the CDSS in simulated scenarios before implementation. The initial findings suggested that the CDSS would be both acceptable and feasible, prompting us to proceed with the implementation (unpublished data).

XSL•FO
RenderX

The implementation process involved asking clinical staff to review a list of patients flagged by the tool and make decisions to defer or recall patients based on their clinical expertise. Using 2 relevant frameworks [22,26], our analysis of staff interviews revealed that most found the concept of identifying patients with UWL at risk of cancer acceptable, and the review process was easy to understand. The most influential factors affecting specific use were the practices' internal processes, culture, workflow, and communication practices.

We observed 2 distinct opinions regarding the use and feasibility of the CDSS for implementation. Although FHT had been installed and used by all practices for over a year, staff familiar with QI activities within the FHT program found the tool useful and feasible. They reported that the process was not burdensome and could be integrated into their usual workflow. Conversely, practices unfamiliar with the tool found it less helpful, citing high misclassification of patients with intentional weight loss, lack of integration with their EHR, and inability to perform recalls through their regular workflow. This dichotomy aligns with existing literature, which identifies trust in the tool and seamless integration into workflow as major factors influencing CDSS implementation success [11]. To overcome some of these barriers, a module such as this could be made actionable within the EHR (recalls, tasking, and documentation editable directly from the cohort view) and provide configurable, role-specific triggers to reduce irrelevant pop-ups (which are a known cause of alert fatigue) [29]. Operational measures include role-tailored training, clear follow-up protocols, and decision aids for practice nurses [30]. At the organizational level, provenance and governance processes, local champions, piloted roll-outs across diverse practice types, and explicit evaluation of implementation outcomes (adoption, fidelity, and acceptability) using implementation frameworks (eg, Normalization Process Theory and Proctor's outcomes) will improve representativeness and scalability [31,32]. Together, these levers—designed and evaluated in partnership with frontline staff—should make integration more seamless, reduce cognitive burden, and increase the likelihood that the tool supports safe, equitable care across a wider range of practices.

Another significant barrier was the lack of allocated time for QI activities, particularly for GPs. This finding is consistent with broader literature on the challenges of implementing QI initiatives in primary care settings [33,34].

Clinical EHR audits showed impressively high follow-up care rates, with 94% (34/36) of patients with confirmed UWL receiving appropriate follow-up and a median of 8 follow-up visits per patient. Our study revealed an exceptionally high follow-up rate, with over 50% (20/36) of referrals made to various specialists. Australian and international literature reports delays in referral for unspecific symptoms compared with specific symptoms [35,36]. However, our findings are consistent with results from our previous randomized controlled trial that used FHT to identify patients with abnormal test results at risk of cancer. In that study, even control practices demonstrated follow-up rates exceeding 70% [17].

The high follow-up rates observed in our study may be reflective of the broader patterns of continuity of care in Australian general practice. The vast majority of Australian patients attend only 1 or 2 general practices annually, and older patients and those with chronic conditions demonstrate particularly high continuity of care [37]. Nevertheless, our findings raise important questions about the representativeness of practices participating in research studies overall. It is possible that these practices are inherently high performing compared to the broader landscape of general practice, potentially skewing our understanding of typical follow-up rates and care quality in primary care settings. Future research needs to explore this issue more thoroughly by comparing the performance metrics of research participating practices with those of a random sample of general practices, investigating the characteristics and motivations of practices that regularly engage in research, and developing strategies to involve a more diverse range of practices in primary care research. Understanding these factors is crucial for accurately assessing the generalizability of research findings and for developing interventions that can be effectively implemented across the entire spectrum of primary care practices. Regardless of the latter, the high rates of investigation and imaging in our cohort (86% had blood tests and 61% had imaging) are consistent and may reflect the accessibility of diagnostic tools in the Australian health care system. This accessibility has been credited for Australia's comparatively better cancer outcomes relative to similar high-income countries worldwide [35].

The symptom profile of patients with UWL presents an interesting diagnostic challenge. One-third (12/36, 33%) of patients identified by the UWL module had no other recorded symptoms, and nearly half (17/36, 47%) had nonspecific symptoms such as appetite loss and lethargy recorded in any field in the EHR. The high prevalence of no other symptoms or coexisting nonspecific symptoms aligns with international literature and underscores the difficulty in differentiating between benign causes and more serious underlying conditions [8,38,39]. Nicholson et al [39] highlight the complexity of interpreting weight loss in primary care, especially when it cooccurs with other nonspecific symptoms, emphasizing the need for careful clinical assessment and follow-up.

Mental health symptoms were also common in our cohort. Just more than half of the patients received a UWL-related diagnosis, primarily mental health conditions or cancer.

The most prevalent diagnoses in our cohort were mental health conditions, cancer, and chronic obstructive pulmonary disease, consistent with international literature. A study by Withrow et al [6] described the most prevalent serious conditions for people with UWL, stratified by age and sex. Given our cohort's median age of 68.5 years and equal sex distribution, our findings align with their reported common diagnoses for this age group: cancer, diabetes, and chronic obstructive pulmonary disease in men; and depression, thyroid disorders, and cancer in women.

Limitations

The algorithm identified 36 (60%) of 60 patients with UWL. Identifying UWL in EHRs is clinically and technically challenging due to inconsistent documentation practices and the predominance of free text in Australian primary care records [40,41]. In this study, "intention" of weight loss was recorded in the clinical notes for 98% of cases. Without access to

XSL•FO
RenderX

unstructured data such as the clinical notes, algorithms such as this, or even more sophisticated natural language models and AI applications have very limited potential. Furthermore, if CDSS performance depends wholly on the quality of the recorded symptoms, then if documentation is incomplete, detection will inevitably be affected. Human scribes in primary care have been shown to raise productivity, increase face-to-face time by 39% to 57%, and improve chart quality and physician satisfaction without reducing patient satisfaction [42]. Ambient AI scribes promise to scale this documentation capture, but evaluations show omissions are frequent and occasional high-severity inaccuracies occur, so clinical oversight and careful product selection are essential. Studies in Canada and the United Kingdom reported strong note-generation capabilities across multiple AI scribes but also marked variability in accuracy, usability, and technical performance [43]. Future research should focus on perfecting these technologies. A pragmatic approach would be to layer scribe-captured narratives into CDSS workflows, enabling richer free-text phenotyping (eg, unintentional weight loss with cosymptoms) while preserving clinician control and opportunities for validation, thereby balancing improved capture with safety and trust. The number of false negatives—patients with UWL not identified by the algorithm—also remains unknown, given that we did not review records not flagged by the algorithm due to cost and time constraints. We were therefore only able to identify true positives and false positives identified by the tool.

The algorithm's failure to exclude 3 patients with cancer (2 cases of lymphoma and myeloma, possibly due to information being solely in clinical notes, and 1 case of breast cancer with a coded diagnosis) highlights an unexpected algorithmic error and underscores the challenges in coding and the need to ensure coding is standardized and reviewed over time. Measured weight loss was not incorporated into the algorithm as an inclusion criterion due to a low percentage of 2 consecutive measurements in preliminary analysis; the impact of this omission on the correct identification of patients remains unclear and warrants further investigation. The study's small scale, with a limited number of patients, significantly restricts its generalizability. Although relevant and necessary to ensure patient safety and good allocation of resources, larger, more diverse patient cohorts are necessary for more robust and widely applicable findings.

These limitations highlight areas for improvement in future research, emphasizing the need for enhanced data access, refined algorithmic processes, and larger study populations to increase the reliability and applicability of methods for identifying patients with UWL.

Conclusions

Our study provides valuable insights into the implementation of a CDSS for identifying patients with UWL at risk of cancer in primary care. While we encountered challenges, particularly around tool integration and correct identification of patients with UWL, we also observed high rates of follow-up care and comprehensive investigations in our participating practices. These findings highlight both the potential of CDSS in improving patient care and the complexities involved in their successful implementation. Future research should focus on enhancing CDSS integration with existing workflows, improving access to unstructured data, to ensure clinicians' trust in these tools, specifically to enhance the usability of digital interventions, and leverage other, more sophisticated digital technologies.

Acknowledgments

We would like to thank all participating practice staff for their invaluable contribution to the study and HaBIC team at the University of Melbourne for the development of the UWL cancer module. Generative AI was not used in the study design and analysis or development of the initial draft of the manuscript. The University of Melbourne's Spark AI was used for copyediting of the revised manuscript.

Funding

JMG was funded by a Victorian State Government through a mid-career Cancer Research Fellowship (MCRF 21025). The Victorian State Government played no role in study design, data collection, analysis and interpretation of data. JDE was funded by an NHMRC Investigator grant (APP1195302). NHMRC played no role in study design, data collection, analysis and interpretation of data.

Data Availability

Unidentified transcriptions and qualitative coded data as well as unidentified audited data can be made available upon request secondary to ethics approval. Qualitative codes were developed in an iterative process by the 2 researchers. They were uploaded to NVivo version 14. They can be shared upon request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview schedule.

XSL•FO
RenderX

[DOCX File, 28 KB - cancer_v12i1e90885_app1.docx]

Multimedia Appendix 2
Data extraction template.
[PDF File, 54 KB - cancer_v12i1e90885_app2.pdf]

Multimedia Appendix 3
Uunintended weight loss criteria for clinical audits.
[DOCX File, 16 KB - cancer_v12i1e90885_app3.docx]

Multimedia Appendix 4
Summary of themes.
[DOCX File, 25 KB - cancer_v12i1e90885_app4.docx]

Multimedia Appendix 5
Detailed audit table.
[DOCX File, 24 KB - cancer_v12i1e90885_app5.docx]

Checklist 1
COREQ checklist.
[PDF File, 234 KB - cancer_v12i1e90885_app6.pdf]

References

1. Neal RD, Tharmanathan P, France B, et al. Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes? Systematic review. Br J Cancer 2015 Mar 31;112 Suppl 1(Suppl 1):S92-107. [doi: 10.1038/bjc.2015.48] [Medline: 25734382]
2. Hamilton W, Walter FM, Rubin G, Neal RD. Improving early diagnosis of symptomatic cancer. Nat Rev Clin Oncol 2016 Dec;13(12):740-749. [doi: 10.1038/nrclinonc.2016.109] [Medline: 27458007]
3. Smith TJ, Hillner BE. Bending the cost curve in cancer care. N Engl J Med 2011 May 26;364(21):2060-2065. [doi: 10.1056/NEJMsb1013826] [Medline: 21612477]
4. Britt H, et al. General Practice Activity in Australia 2015-16: Sydney University Press; 2016, Vol. 40. URL: https://ses.library.usyd.edu.au/handle/2123/15514 [accessed 2026-06-19]
5. Primary health care. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/primary-health-care [accessed 2025-05-16]
6. Withrow DR, Oke J, Friedemann Smith C, Hobbs R, Nicholson BD. Serious disease risk among patients with unexpected weight loss: a matched cohort of over 70 000 primary care presentations. J Cachexia Sarcopenia Muscle 2022 Dec;13(6):2661-2668. [doi: 10.1002/jcsm.13056] [Medline: 36056750]
7. Nicholson BD, Hamilton W, Koshiaris C, Oke JL, Hobbs FDR, Aveyard P. The association between unexpected weight loss and cancer diagnosis in primary care: a matched cohort analysis of 65,000 presentations. Br J Cancer 2020 Jun;122(12):1848-1856. [doi: 10.1038/s41416-020-0829-3] [Medline: 32291391]
8. Hernández JL, Riancho JA, Matorras P, González-Macías J. Clinical evaluation for cancer in patients with involuntary weight loss without specific symptoms. Am J Med 2003 Jun 1;114(8):631-637. [doi: 10.1016/s0002-9343(03)00115-3] [Medline: 12798450]
9. Hamilton W. The CAPER studies: five case-control studies aimed at identifying and quantifying the risk of cancer in symptomatic primary care patients. Br J Cancer 2009 Dec 3;101 Suppl 2(Suppl 2):S80-S86. [doi: 10.1038/sj.bjc.6605396] [Medline: 19956169]
10. Shortliffe EH, Cimino JJ. Biomedical Informatics: Computer Applications in Health Care and Biomedicine, 4th edition: Publisher; 2014:1-965. [doi: 10.1007/978-1-4471-4474-8/COVER]
11. Chima S, Reece JC, Milley K, Milton S, McIntosh JG, Emery JD. Decision support tools to improve cancer diagnostic decision making in primary care: a systematic review. Br J Gen Pract 2019 Dec;69(689):e809-e818. [doi: 10.3399/bjgp19X706745] [Medline: 31740460]
12. Price S, Spencer A, Medina-Lara A, Hamilton W. Availability and use of cancer decision-support tools: a cross-sectional survey of UK primary care. Br J Gen Pract 2019 Jul;69(684):e437-e443. [doi: 10.3399/bjgp19X703745] [Medline: 31064743]
13. Chima S, Hunter B, Martinez-Gutierrez J, et al. Implementation of a quality improvement and clinical decision support tool for cancer diagnosis in primary care: process evaluation. JMIR Cancer 2025 Jun 12;11(1):e65461. [doi: 10.2196/65461] [Medline: 40505100]

XSL•FO
RenderX

14. May CR, Mair FS, Dowrick CF, Finch TL. Process evaluation for complex interventions in primary care: understanding trials using the normalization process model. BMC Fam Pract 2007 Jul 24;8:42. [doi: 10.1186/1471-2296-8-42] [Medline: 17650326]
15. Hunter B, Biezen R, Alexander K, et al. Future Health Today: codesign of an electronic chronic disease quality improvement tool for use in general practice using a service design approach. BMJ Open 2020 Dec 18;10(12):e040228. [doi: 10.1136/bmjopen-2020-040228] [Medline: 33371024]
16. Chima S, Martinez-Gutierrez J, Hunter B, Manski-Nankervis JA, Emery J. Optimization of a quality improvement tool for cancer diagnosis in primary care: qualitative study. JMIR Form Res 2022 Aug 4;6(8):e39277. [doi: 10.2196/39277] [Medline: 35925656]
17. Chima S, Martinez-Gutierrez J, Hunter B, et al. Future Health Today and patients at risk of undiagnosed cancer: a pragmatic cluster randomised trial of quality- improvement activities in general practice. Br J Gen Pract 2025 May;75(754):e306-e315. [doi: 10.3399/BJGP.2024.0491] [Medline: 39567181]
18. Hunter B, Alexander K, Biezen R, et al. The development of Future Health Today: piloting a new platform for identification and management of chronic disease in general practice. Aust J Prim Health 2023 Feb;29(1):8-15. [doi: 10.1071/PY22022] [Medline: 36318973]
19. Lee A, de Mendonça L, McCarthy D, et al. Primary care patients presenting with unexpected weight loss in Australian general practices: replication of a diagnostic accuracy study. BMJ Open 2025 Jul 28;15(7):e104690. [doi: 10.1136/bmjopen-2025-104690] [Medline: 40730405]
20. Practice incentives program quality improvement measures: annual data update 2024–25, QIM 3: height and weight recorded and BMI. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/reports/primary-health-care/pipqi-measures-2024-25/contents/pipqi-measures/qim-3-height-and-weight-recorded-and-bmi [accessed 2025-11-06]
21. Martinez-Gutierrez J, De Mendonca L, Ly P, et al. A scoping review of unexpected weight loss and cancer: risk, guidelines, and recommendations for follow-up in primary care. BJGP Open 2024 Dec;8(4):BJGPO. [doi: 10.3399/BJGPO.2024.0025] [Medline: 39054298]
22. Singh H, Sittig DF. A sociotechnical framework for safety-related electronic health record research reporting: the safer reporting framework. Ann Intern Med 2020 Jun 2;172(11 Suppl):S92-S100. [doi: 10.7326/M19-0879] [Medline: 32479184]
23. Van de Velde S, Kunnamo I, Roshanov P, et al. The GUIDES checklist: development of a tool to improve the successful use of guideline-based computerised clinical decision support. Implement Sci 2018 Jun 25;13(1):86. [doi: 10.1186/s13012-018-0772-3] [Medline: 29941007]
24. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
25. Braun V, Clarke V. Research designs: quantitative, qualitative, neuropsychological, and biological. In: APA Handbook of Research Methods in Psychology: American Psychological Association; 2012, Vol. 2:57-71. [doi: 10.1037/13620-004]
26. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
27. Modified monash model. Australian Government Department of Health, Disability and Ageing. 2025. URL: https://www.health.gov.au/topics/rural-health-workforce/classifications/mmm [accessed 2025-09-15]
28. Ancker JS, Edwards A, Nosal S, et al. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. BMC Med Inform Decis Mak 2017 Apr 10;17(1):36. [doi: 10.1186/s12911-017-0430-8] [Medline: 28395667]
29. Carayon P, Schoofs Hundt A, Karsh BT, et al. Work system design for patient safety: the SEIPS model. Qual Saf Health Care 2006 Dec;15 Suppl 1(Suppl 1):i50-i58. [doi: 10.1136/qshc.2005.015842] [Medline: 17142610]
30. May C, Finch T. Implementing, embedding, and integrating practices: an outline of Normalization Process Theory. Sociology 2009 Jun;43(3):535-554. [doi: 10.1177/0038038509103208]
31. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. Implement Sci 2013 Dec 1;8:139. [doi: 10.1186/1748-5908-8-139] [Medline: 24289295]
32. Kortteisto T, Komulainen J, Mäkelä M, Kunnamo I, Kaila M. Clinical decision support must be useful, functional is not enough: a qualitative study of computer-based clinical decision support in primary care. BMC Health Serv Res 2012 Oct 8;12(1):1-9. [doi: 10.1186/1472-6963-12-349] [Medline: 23039113]
33. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999 Oct 20;282(15):1458-1465. [doi: 10.1001/jama.282.15.1458] [Medline: 10535437]
34. Lynch C, Harrison S, Emery JD, et al. Variation in suspected cancer referral pathways in primary care: comparative analysis across the International Benchmarking Cancer Partnership. Br J Gen Pract 2023 Feb;73(727):e88-e94. [doi: 10.3399/BJGP.2022.0110] [Medline: 36127155]
35. Lacey K, Bishop JF, Cross HL, Chondros P, Lyratzopoulos G, Emery JD. Presentations to general practice before a cancer diagnosis in Victoria: a cross-sectional survey. Med J Aust 2016 Jul 18;205(2):66-71. [doi: 10.5694/mja15.01169] [Medline: 27456447]

XSL•FO
RenderX

36. Wright M, Hall J, van Gool K, Haas M. How common is multiple general practice attendance in Australia? Aust J Gen Pract 2018 May;47(5):289-296. [doi: 10.31128/AJGP-11-17-4413] [Medline: 29779298]
37. Nicholson BD, Perera R, Thompson MJ. The elusive diagnosis of cancer: testing times. Br J Gen Pract 2018 Nov;68(676):510-511. [doi: 10.3399/bjgp18X699461]
38. Nicholson BD, Hamilton W, O'Sullivan J, Aveyard P, Hobbs FR. Weight loss as a predictor of cancer in primary care: a systematic review and meta-analysis. Br J Gen Pract 2018 May;68(670):e311-e322. [doi: 10.3399/bjgp18X695801] [Medline: 29632004]
39. Nicholson BD, Aveyard P, Hamilton W, Hobbs FDR. When should unexpected weight loss warrant further investigation to exclude cancer? BMJ 2019 Sep 23;366:l5271. [doi: 10.1136/bmj.l5271] [Medline: 31548272]
40. Canaway R, Boyle D, Manski-Nankervis JA, Gray K. Identifying primary care datasets and perspectives on their secondary use: a survey of Australian data users and custodians. BMC Med Inform Decis Mak 2022 Apr 6;22(1):94. [doi: 10.1186/s12911-022-01830-9] [Medline: 35387634]
41. Zallman L, Finnegan K, Roll D, Todaro M, Oneiz R, Sayah A. Impact of medical scribes in primary care on productivity, face-to-face time, and patient comfort. J Am Board Fam Med 2018;31(4):612-619. [doi: 10.3122/jabfm.2018.04.170325] [Medline: 29986987]
42. Ha E, Choon-Kon-Yune I, Murray L, et al. Evaluating the usability, technical performance, and accuracy of artificial intelligence scribes for primary care: competitive analysis. JMIR Hum Factors 2025 Jul 23;12:e71434. [doi: 10.2196/71434] [Medline: 40700466]
43. Draper TC, Cox T, Lamb-Riddell K, et al. Clinical AI Scribes in primary care: accuracy, error severity and implications for clinical practice. BMJ Digit Health 2025 Sep;1(1):e000092. [doi: 10.1136/bmjdhai-2025-000092]

Abbreviations

CDSS: clinical decision support system
COREQ: Consolidated Criteria for Reporting Qualitative
EHR: electronic health record
FHT: Future Health Today
GP: general practitioner
QI: quality improvement
UWL: unintended weight loss

Edited by M Balcarras; submitted 06.Jan.2026; peer-reviewed by B Senst, K Khashu; revised version received 01.May.2026; accepted 04.May.2026; published 28.Jul.2026.

Please cite as:
Martinez-Gutierrez J, Chima S, Hunter B, Lee A, De Mendonca L, Daly D, Fishman G, Lim FS, Wang B, Nelson C, Nicholson B, Manski-Nankervis JA, Emery J
Piloting a Clinical Decision Support System for Unintended Weight Loss in Primary Care: Mixed Methods Study on Early Cancer Detection
JMIR Cancer 2026;12:e90885
URL: https://cancer.jmir.org/2026/1/e90885
doi:10.2196/90885

© Javiera Martinez-Gutierrez, Sophie Chima, Barbara Hunter, Alex Lee, Lucas De Mendonca, Deborah Daly, George Fishman, Fong Seng Lim, Benny Wang, Craig Nelson, Brian Nicholson, Jo-Anne Manski-Nankervis, Jon Emery. Originally published in JMIR Cancer (https://cancer.jmir.org), 28.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Explainable Machine Learning–Based Prediction of Progression-Free Survival in Prostate Cancer: Retrospective Cohort Study

Hein Minn Tun[1,2], DipFM, MBBS, MPH; Lin Naing[1], MBBS, MPH, MHlthSc, MMedStat; Owais Ahmed Malik[3], MSc, PhD; Muhammad Syafiq Abdullah[4], MBChB, MRCP, MSc; Thu Ta[4], MBBS, MRCP(UK); Hanif Abdul Rahman[1,5], PhD

[1]Pengiran Anak Puteri Rashidah Sa'adatul Bolkiah Institute of Health Sciences, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei-Muara District, Brunei Darussalam

[2]School of Digital Science, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei Darussalam

[3]Faculty of Engineering and Computing, Atlantic Technological University, Donegal, Ireland

[4]The Brunei Cancer Centre, Jerudong Park Medical Centre, Bandar Seri Begawan, Brunei-Muara District, Brunei Darussalam

[5]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

Corresponding Author:
Hein Minn Tun, DipFM, MBBS, MPH
Pengiran Anak Puteri Rashidah Sa'adatul Bolkiah Institute of Health Sciences
Universiti Brunei Darussalam
Jalan Tungku Link
Gadong
Bandar Seri Begawan, Brunei-Muara District, BE1410
Brunei Darussalam
Phone: 673 7428942
Email: 23h8750@ubd.edu.bn

Abstract

Background: Progression-free survival (PFS) is a critical end point in oncology, yet real-world applications of individualized, explainable machine learning (ML) predictions remain limited.

Objective: This study aimed to develop and validate explainable ML models to predict PFS using retrospective data from a national prostate cancer cohort in Brunei Darussalam.

Methods: We analyzed a retrospective cohort of 212 patients (478 longitudinal observations) treated at the Brunei Cancer Centre (January 2018 to December 2024). Clinical, laboratory, and treatment data were harmonized, with missing values imputed via extremely randomized trees. Longitudinal patterns were captured using a recurrent autoencoder to generate latent representations. We compared 4 modeling approaches: Cox proportional hazards, random survival forest (RSF), gradient boosting survival (GBS), and deep neural network survival models. Performance was evaluated using time-dependent area under the receiver operating characteristic curve (AUC), Harrell C-index, and integrated Brier score (IBS), with Shapley additive explanations (SHAP) used for interpretability.

Results: RSF demonstrated improved discriminative performance and balanced calibration, achieving a C-index of 0.906 and AUCs of 0.941 and 0.917 at 4 and 5 years (IBS=0.0698). In contrast, the traditional CPH model performed poorly (C-index 0.531 and AUC 0.706 at 4 years and 0.833 at 5 years). Deep survival (AUCs of 0.941 at 4 years and 0.917 at 5 years, C-index 0.719, and IBS=0.0887) and GBS (AUCs of 0.765 at 4 years and 0.833 at 5 years, C-index 0.844, and IBS=0.0590) models showed moderate performance. SHAP analysis identified sodium, alanine aminotransferase, mean corpuscular hemoglobin, platelet count, and specific treatment categories as key drivers of increased progression risk.

Conclusions: Tree-based ensemble approaches, particularly RSF integrated with SHAP, offer high accuracy for personalized risk stratification in prostate cancer. These findings highlight the potential of explainable ML to enhance clinical decision-making. However, external validation in a larger multi-institutional, multiomics dataset is required before routine clinical implementation.

(JMIR Cancer 2026;12:e91510) doi:10.2196/91510

XSL•FO
RenderX

KEYWORDS

prostate cancer; progression-free survival; explainable AI; machine learning; prognosis

Introduction

Prostate cancer is the fourth most common cancer in men worldwide, and the incidence of prostate cancer is increasing in Africa, Asia, and Latin America and the Caribbean, raising public health concerns worldwide [1,2]. Global and regional analyses demonstrate that the highest incidence rates are reported in northern Europe, North America, and Australia, largely due to the widespread use of prostate-specific antigen (PSA)–based early detection systems, while mortality is higher in Africa, Asia, and Latin America, indicating the global disparity based on screening and diagnostic capacity [2-5]. In Brunei Darussalam specifically, the national and international cancer registry data demonstrate the increasing prostate cancer deaths in recent decades [6,7]. These epidemiological trends in prostate cancer highlight the need for locally adapted prognostic and management tools that can improve the prostate cancer care continuum in Brunei Darussalam.

Progression-free survival (PFS) and disease-free survival are widely used in prostate cancer clinical trials and observational studies to quantify the duration during which a patient remains free from radiographic or clinical disease progression after treatment [8-10]. PFS often serves as a pragmatic outcome in trials of systemic and targeted therapies, as it can be observed earlier than overall survival and may reflect the treatment effect on tumor control. According to PCWG (Prostate Cancer Working Group) criteria, PSA progression is defined as a confirmed ≥25% and ≥2 ng/mL increase from the PSA nadir (or from baseline after ≥12 weeks if no nadir is achieved), with confirmation to exclude PSA flare, while radiographic or clinical progression, if occurring earlier, takes precedence in determining PFS [11]. Additionally, the understanding of PFS extension for prostate cancer is viewed differently between patients and clinicians. Patients generally require a longer PFS benefit to consider treatment worthwhile and tend to prioritize PSA or radiological responses despite potential side effects, whereas clinicians place greater emphasis on balancing treatment-related toxicity with PFS and may recommend therapy even when the expected PFS gain is modest [12].

Over the past decade, machine learning (ML) methods, including gradient-boosted trees, random survival forests, and deep learning architectures, have been increasingly applied to survival prediction and progression modeling across many cancer types, including lung and prostate cancer [13-17]. Furthermore, multimodal data integration of clinical variables, laboratory tests, imaging and radiomics features, and molecular data to produce individualized risk estimates for progression or death often outperforms traditional statistical models in complex, high-dimensional datasets [17,18]. Despite these advances, most ML and deep learning models for prostate cancer survival prediction rely on cross-sectional or clinical trial datasets with short follow-up periods, limiting their applicability in real-world practice [19-21]. Additionally, explainability methods such as Shapley additive explanations (SHAP) play an important role in the clinical applicability of these models by providing attributions for individual predictions [18,22,23].

However, few studies have explored the sequential analysis of progression free survival with explainable methods to provide locally calibrated, interpretable predictions for progression-free outcomes. Additionally, no previous study in Brunei Darussalam has been conducted to develop interpretable ML survival models for prostate cancer progression. To fill these gaps, our study aimed to develop explainable ML models that predict PFS using retrospective clinical, laboratory, and treatment information from a national cohort in Brunei Darussalam. By developing locally tailored, AI-enabled PFS models, this study contributes to the growing body of evidence supporting the use of ML and deep learning in precision oncology, while also advancing AI-based clinical decision support systems to assist oncologists in prostate cancer management and enhance the continuum of care in Brunei Darussalam.

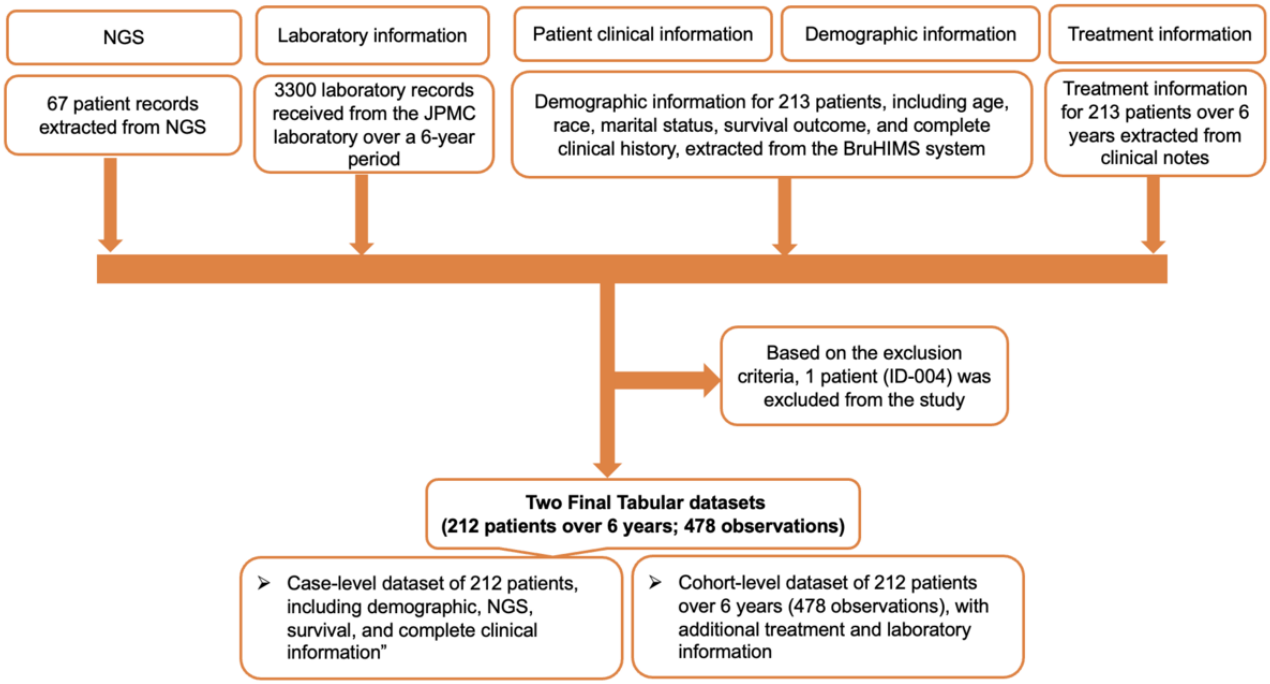
Methods

Study Design and Data Source

This retrospective cohort study included patients with prostate cancer treated at the Brunei Cancer Centre (TBCC), Jerudong Park Medical Centre (JPMC) between January 2018 and December 2024. The study was conducted in accordance with the TRIPOD-AI (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis adapted for Artificial Intelligence) reporting guidelines to enhance transparency, reproducibility, and clinical applicability (Multimedia Appendix 1). Eligible participants were male patients with a histologically confirmed diagnosis of prostate cancer who received at least 1 treatment modality at TBCC. Patients with secondary metastatic prostate cancer or multiple primary malignancies were excluded following clinical review by oncologists. In total, 212 eligible patients contributed 478 time-indexed observations, capturing longitudinal treatment pathways and disease progression, as illustrated in Figure 1. Demographic, clinical, and treatment data were extracted from multiple sources, including the Brunei Darussalam Healthcare Information and Management System, the JPMC laboratory database, and both electronic and manually documented clinical records.

XSL•FO
RenderX

Figure 1. Overview of cohort data processing and source integration for the study. BruHIMS: Brunei Darussalam Healthcare Information and Management System; JPMC: Jerudong Park Medical Centre; NGS: next-generation sequencing.



Data Structure and Outcome Definition

Data were structured across 2 hierarchical levels: the individual case level and the longitudinal cohort level. At the case level, collected variables encompassed demographic characteristics (including age, ethnicity, and marital status), medical history (eg, comorbidities and family history of malignancy), lifestyle factors (particularly smoking status), anthropometric measures (height, weight, and BMI), tumor-related features (Gleason score, tumor stage, and ECOG performance status), and survival outcomes. At the cohort level, comprehensive treatment information and 23 longitudinal laboratory measures were assembled, including baseline and 1-year posttreatment PSA values, full blood count parameters, liver and renal function indicators, erythrocyte sedimentation rate, calcium, and phosphate levels. Absolute and relative changes in PSA were derived, and tumor staging, Gleason grade groups, and National Comprehensive Cancer Network prognostic risk categories were harmonized to address inconsistencies in coding practices [24]. Except for PSA, all laboratory parameters were baseline measurements obtained before each treatment modification, enabling clinicians to assess laboratory results when determining subsequent therapeutic strategies. Data integrity was ensured through systematic cross-validation across multiple data sources and clinical verification by oncologists before analysis. The primary outcome was PFS.

PFS was calculated on a treatment-episode basis from treatment initiation to the earliest occurrence of PSA-defined progression, which was determined using PCWG-consistent criteria based on treatment-specific baseline PSA, posttreatment PSA nadir, and confirmed PSA rises (≥25% and ≥2.0 ng/mL, with timing and confirmation rules applied) or death from any cause, with patients censored at last follow-up in the absence of an event. A PFS event value of 0 indicates no biochemical progression or survival without progression at the last follow-up (censored), whereas a value of 1 indicates the occurrence of biochemical PSA progression or death from any cause. To conduct a survival analysis, PFS events, which denote the occurrence of progression or death, were used rather than PFS, which indicates the time a patient remains progression-free.

Data Preprocessing and Survival Analysis

Statistical analyses were performed using RStudio (version 4.3.3; R Foundation for Statistical Computing), Python (version 3.10; Python Software Foundation) implemented via Google Colab, and Visual Studio Code (version 1.106.0; Microsoft). Two analytic datasets were created: a case-level dataset consisting of 212 patients and a longitudinal cohort-level dataset comprising 478 time-indexed observations. Data preprocessing procedures included data cleaning, harmonization of tumor staging and Gleason grade groups, elimination of duplicate records, and construction of patient-specific longitudinal sequences. At the case level, continuous variables were compared using the Mann-Whitney *U* test, while categorical variables were assessed using the Fisher exact test, with a significance threshold set at $P<.05$. For the cohort-level analysis, mixed-effects linear and logistic regression models were used to evaluate differences between survival groups. The differences between groups, including those stratified by comorbidity status, were assessed using the log-rank test.

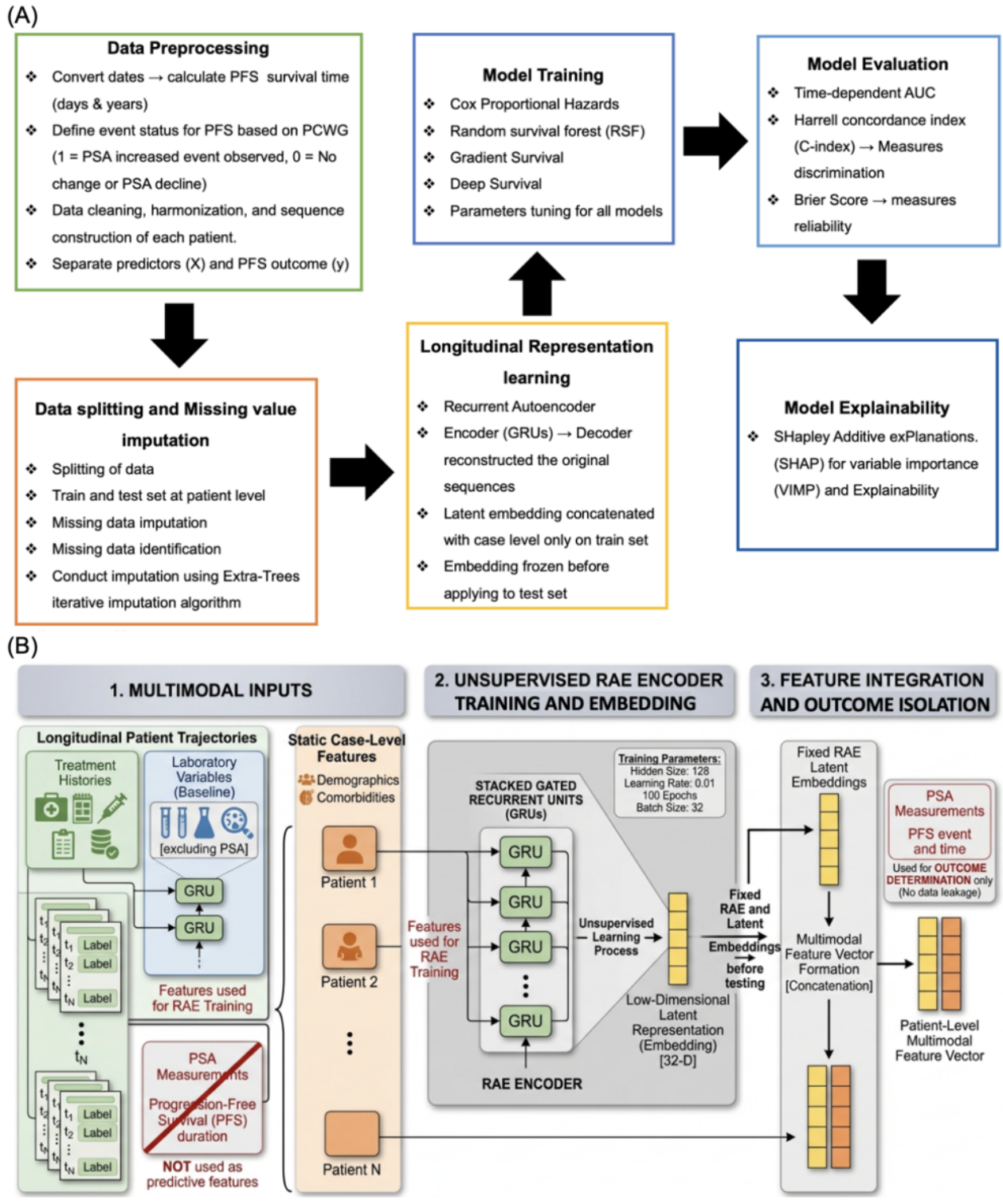
Data Splitting and Missing Data Handling

The overall ML model development pipeline is illustrated in Figure 2A. To minimize information leakage, data were split at the individual case level into training (80%) and testing (20%) sets using a random group-splitting strategy [25]. Missing values in both static and longitudinal variables were addressed using extremely randomized trees–based iterative imputation, a nonparametric approach derived from the random forest algorithm that accommodates mixed data types and captures

XSL•FO
RenderX

nonlinear relationships [26]. Categorical variables were transformed using one-hot encoding, and continuous variables were standardized using *z* score normalization to ensure comparable feature scales. All preprocessing steps, including imputation, scaling, and encoder fitting, were performed solely on the training dataset and subsequently applied to the testing dataset.

Figure 2. (A) Overview of the machine learning model development workflow for progression-free survival (PFS). (B) Recurrent autoencoder (RAE) architecture for longitudinal representation. AUC: area under the curve; GRU: gated recurrent unit; PCWG: Prostate Cancer Working Group; PSA: prostate-specific antigen.



XSL•FO
RenderX

Longitudinal Representation Learning Using a Recurrent Autoencoder

A recurrent autoencoder (RAE) was trained on the training cohort to capture temporal patterns in laboratory and treatment histories (Figure 2B) [20]. PFS duration served as the primary temporal input, with all laboratory variables recorded as baseline measurements before each treatment decision. PSA measurements were used exclusively for PFS event determination according to biochemical progression criteria and were not incorporated as predictive features in model training. The PFS event and time were used only as indexing variables, not as features, to avoid learning in the autoencoder, thereby ensuring that the autoencoder learned purely unsupervised representations of longitudinal patient trajectories without exposure to the outcome. The RAE encoder, composed of stacked gated recurrent units (GRUs), transformed each patient's sequence into a low-dimensional latent embedding, which the decoder then reconstructed to preserve the original temporal patterns. These latent embeddings were concatenated with case-level static features (eg, race and morbidities) to form a multimodal feature vector. To prevent data leakage, the RAE was trained exclusively on the training dataset, and the learned embeddings were fixed before application to the testing set. Hyperparameters included a hidden layer size of 128, an embedding dimension of 32, a learning rate of 0.01, 100 epochs, and a batch size of 32.

Model Development

Traditional Cox Proportional Hazards Model

A Cox proportional hazards model was fitted to the training dataset as a conventional baseline for PFS analysis. The model was adjusted for clinically relevant covariates to quantify the association between predictors and the hazard of disease progression or death, defined as the instantaneous risk of a progression event at a given time, conditional on remaining progression-free up to that point [27,28]. Final covariate selection was guided by model fit, assessed using the log-likelihood ratio test. The proportional hazards assumption was examined using scaled Schoenfeld residuals and corresponding statistical tests to detect systematic departures from time-invariant covariate effects [19-32]. Predicted risk scores were generated for the testing dataset to facilitate direct comparison with ML-based PFS models.

ML Models

Three ML-based survival models were developed using an integrated feature set combining static clinical variables and latent temporal representations derived from the RAE. First, a random survival forest (RSF) model was trained as a nonparametric ensemble method to estimate cumulative hazard functions and rank predictors by importance [33]. Second, a gradient boosting survival (GBS) model (GBT-Survival) was implemented using a Cox partial likelihood loss function, enabling the modeling of nonlinear effects and complex interactions among covariates [34]. Third, a deep neural network–based Cox regression model (DeepSurv) was trained to capture nonlinear relationships between predictors and progression risk by optimizing the Cox partial likelihood [20]. Hyperparameter optimization for all models was performed using grid search cross-validation on the training dataset [35]. For the RSF, the explored parameter grid included the number of estimators (100, 200, 300), maximum depth (5, 7, 10), and minimum samples per split (10, 20, 30). For the GBS (GBT-Survival) model, the grid included learning rates (0.001, 0.01, 0.05, 0.1), the number of estimators (100, 200, 300), and maximum depth (3, 5, 7, 10). Optimal configurations, selected based on the highest concordance index, included RSF (maximum depth=7, minimum samples per split=10, 100 estimators), GBT-Survival (learning rate=0.001, maximum depth=10, 300 estimators), and DeepSurv (hidden layers=[32,35], dropout=0.1, rectified linear unit activation).

Model Evaluation and Explainability

Time-dependent area under the receiver operating characteristic curve (AUC) values could not be reliably estimated at 1, 2, or 3 years for any model because there were too few progression events and insufficient comparable patient pairs within these early periods to assess discrimination accurately. Model performance was assessed using time-dependent AUC values at the 4- and 5-year PFS horizons, together with Harrell concordance index (C-index) to evaluate discriminative ability [36]. Model robustness was examined using bootstrap resampling with 1000 iterations to derive CIs. Calibration was evaluated using the integrated Brier score (IBS) and calibration plots comparing predicted and observed progression-free survival probabilities. To enhance model interpretability, SHAP was applied to the best-performing RSF-Survival model, providing both global and patient-level insights into feature contributions to progression risk. SHAP summary, dependence, and force plots were used to visualize time-dependent predictor effects and identify key drivers of disease progression [19].

Ethical Considerations

Ethics approval for this study was granted by the Medical and Health Research and Ethics Committee, Ministry of Health, Brunei Darussalam (MHREC/MOH/2024/22(1)), on January 25, 2025. All patient data were replaced with pseudonymous research IDs since we received the data from the hospital, and no personally identifiable information was accessed or retained within the analytic or modeling environment.

Software Packages and Analytic Tools

All analyses were performed in Python using established open-source libraries, including *pandas* and *NumPy* for data manipulation; *scikit-survival* for Cox proportional hazards modeling, RSF, GBS analysis, concordance index estimation, and time-dependent AUC; *lifelines* and *pycox* for survival modeling (including DeepSurv); *PyTorch* (torch, torchvision, and torchaudio) for deep learning implementation; *fancyimpute* for iterative imputation; *scikit-learn* for data splitting and preprocessing (GroupShuffleSplit and StandardScaler); and *SHAP* for model interpretability.

XSL•FO
RenderX

Results

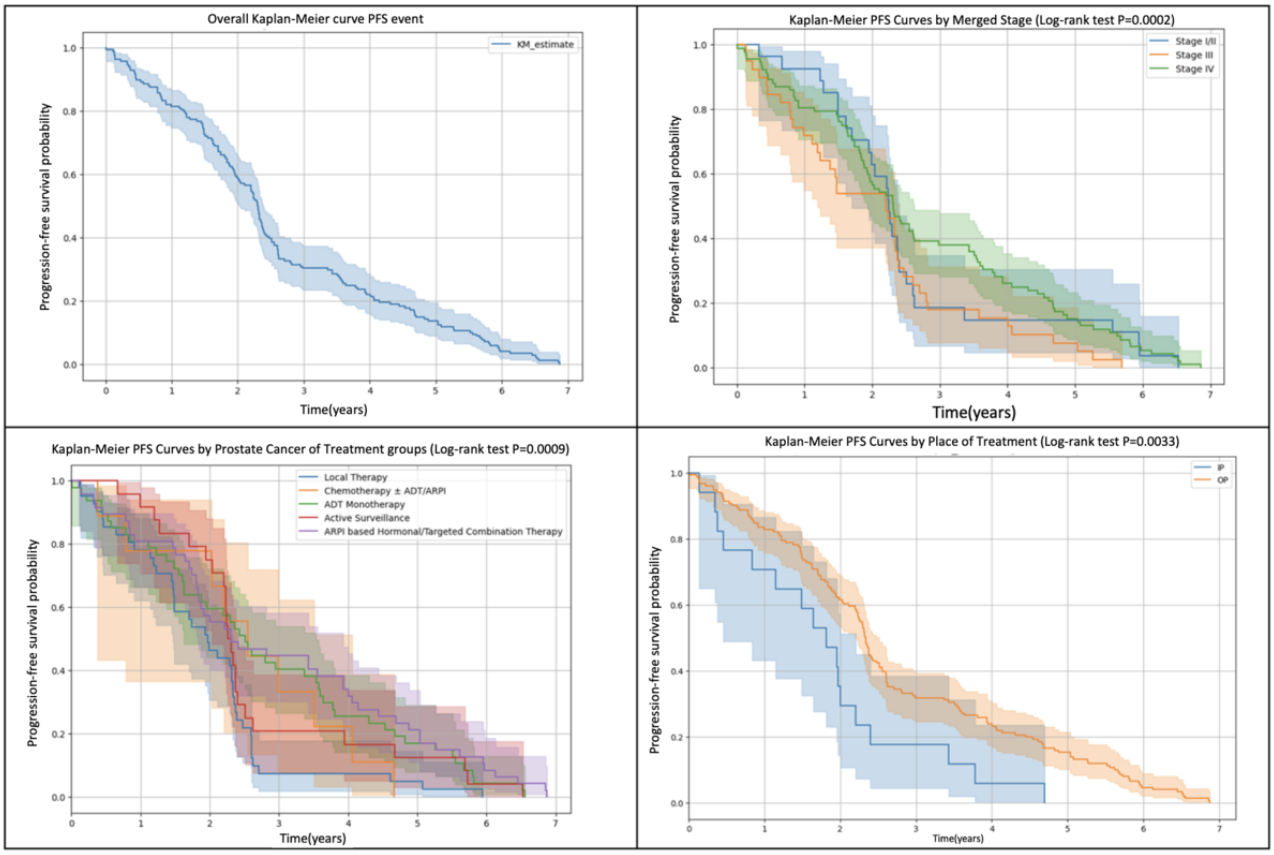
Characteristics and Univariable Analysis of Study Participants

The study included 212 patients at the case level, contributing a total of 478 longitudinal observations over 6 years (detailed participant characteristics are presented in Multimedia Appendix 2). PFS events were reported in 45 observations during follow-up. At the case level (n=212), PFS outcomes differed according to tumor-related factors, including prostate cancer stage and Gleason grade group, whereas other comorbid conditions did not differ significantly between PFS outcome groups. Furthermore, at the cohort level (478 observations), significant differences between observations with and without a PFS event were observed in calcium (P<.001), chloride (P<.001), sodium (P<.001), red blood cell count (P<.01), hemoglobin (P<.01), neutrophils (P<.01), eosinophils (P<.01), baseline PSA (P<.001), and posttreatment PSA (P<.001).

Kaplan-Meier Curve for PFS

The PFS analysis with significant variables is presented using the Kaplan-Meier curves in Figure 3. The overall PFS curve demonstrates a steady decline over time, with the probability of remaining progression-free markedly decreasing after approximately 2 years of follow-up. Stratified analyses revealed significant differences across merged clinical stages (log-rank test, P=.0002), with advanced stages showing poorer PFS than earlier stages. Similarly, significant variation in PFS was observed among different prostate cancer treatment groups (log-rank test, P=.0009), indicating heterogeneity in outcomes depending on the therapeutic approach. Furthermore, PFS differed significantly according to place of treatment (log-rank test, P=.0033), with the inpatient treatment group demonstrating consistently lower survival probability over time.

Figure 3. Kaplan-Meier curves for progression-free survival among prostate cancer patients at the Brunei Cancer Centre (TBCC) from January 2018 to December 2024. ADT: androgen deprivation therapy; ARPI: androgen receptor pathway inhibitor; IP: inpatients; OP: outpatients.



Discriminative Performance of PFS Models

The discriminative performance, as measured by the time-dependent AUC at 1, 3, and 5 years (Figure 4), and the Harrell C-index (Table 1), was used to develop and compare 4 survival models. Evaluation of model performance at 4 and 5 years indicated that ML-based models outperformed the traditional Cox proportional hazards models. The Cox model demonstrated limited discriminative ability, with time-dependent AUCs of 0.706 and 0.833 at 4 and 5 years, respectively, along with a C-index of 0.531. In contrast, the tree-based models, RSF and GBT-Survival, showed markedly superior discrimination. The RSF model demonstrated strong discriminative performance at 4 and 5 years (AUCs of 0.941 and 0.917, respectively; C-index 0.906), while the deep survival model (AUCs of 0.941 at 4 years and 0.917 at 5-year; C-index 0.719) and the GBS model (AUCs of 0.765 at 4 years and 0.833 at 5-year, C-index 0.844) demonstrated moderate discriminative performance.

XSL•FO
RenderX

Figure 4. Heatmap of time-dependent area under the receiver operating characteristic curve (AUC) of progression-free survival (PFS) models, showing the relationship between predicted accuracy in fourth and fifth years. Cox: Cox proportional hazards model; DeepSurv: deep neural network–based Cox regression model; GBS: gradient boosting survival; RSF: random survival forest.

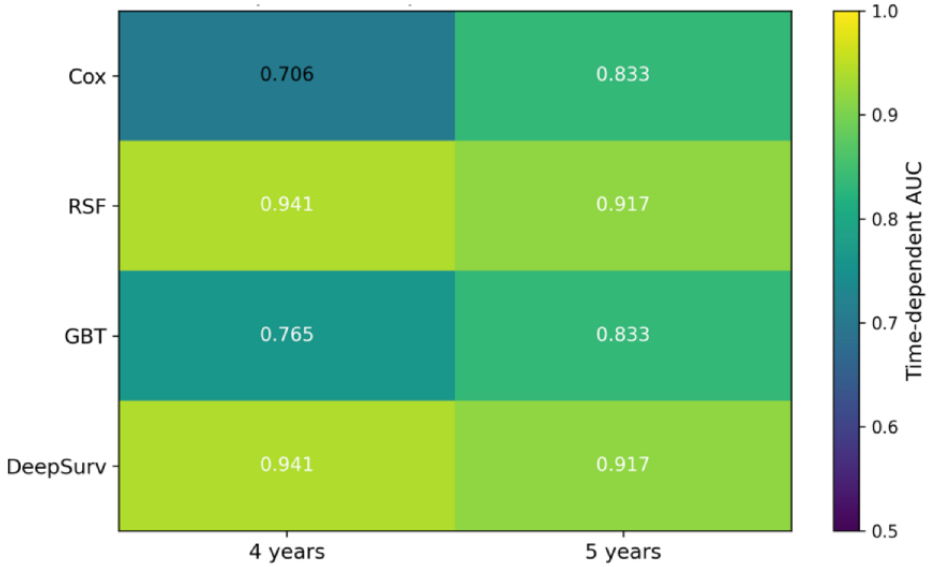


Table 1. Discriminative performance (time-dependent area under the receiver operating characteristic curve [AUC], Harrell concordance index [C-Index]) and calibration accuracy (integrated Brier score) of progression-free survival models.

Models	Time-dependent AUC		Harrell C-index	Integrated Brier score
	4 years	5 years		
Cox proportional hazards	0.706	0.833	0.531	0.0527
Random survival forest	0.941	0.917	0.906	0.0698
Gradient boosting survival	0.765	0.833	0.844	0.0590
Deep neural network survival	0.941	0.917	0.719	0.0887

Calibration Performance Using IBS for PFS Model

Calibration performance of the PFS models was assessed using the IBS, where lower values indicate better agreement between predicted and observed outcomes (Table 1). Among the evaluated models, the Cox proportional hazards model demonstrated the best calibration accuracy with the lowest IBS of 0.0527. The GBS model demonstrated comparable calibration performance (IBS=0.0590), followed by the RSF model (IBS=0.0698). In contrast, the deep neural network survival model exhibited the highest IBS (IBS=0.0887), indicating comparatively poorer calibration despite its better discriminative performance.

Although similar time-dependent AUC values were observed between the RSF and deep neural network survival models at specific predefined time points, the limited number of events in the study may reduce the sensitivity of AUC to detect subtle performance differences between models. Therefore, we placed greater emphasis on the C-index and IBS, as these metrics provide a more comprehensive assessment of survival model performance across the entire follow-up period compared with isolated time-point AUC measurements. According to both discriminative and calibration performance, the RSF model demonstrated stable and well-calibrated risk prediction over time.

Model Explainability Using SHAP for RSF Models for PFS

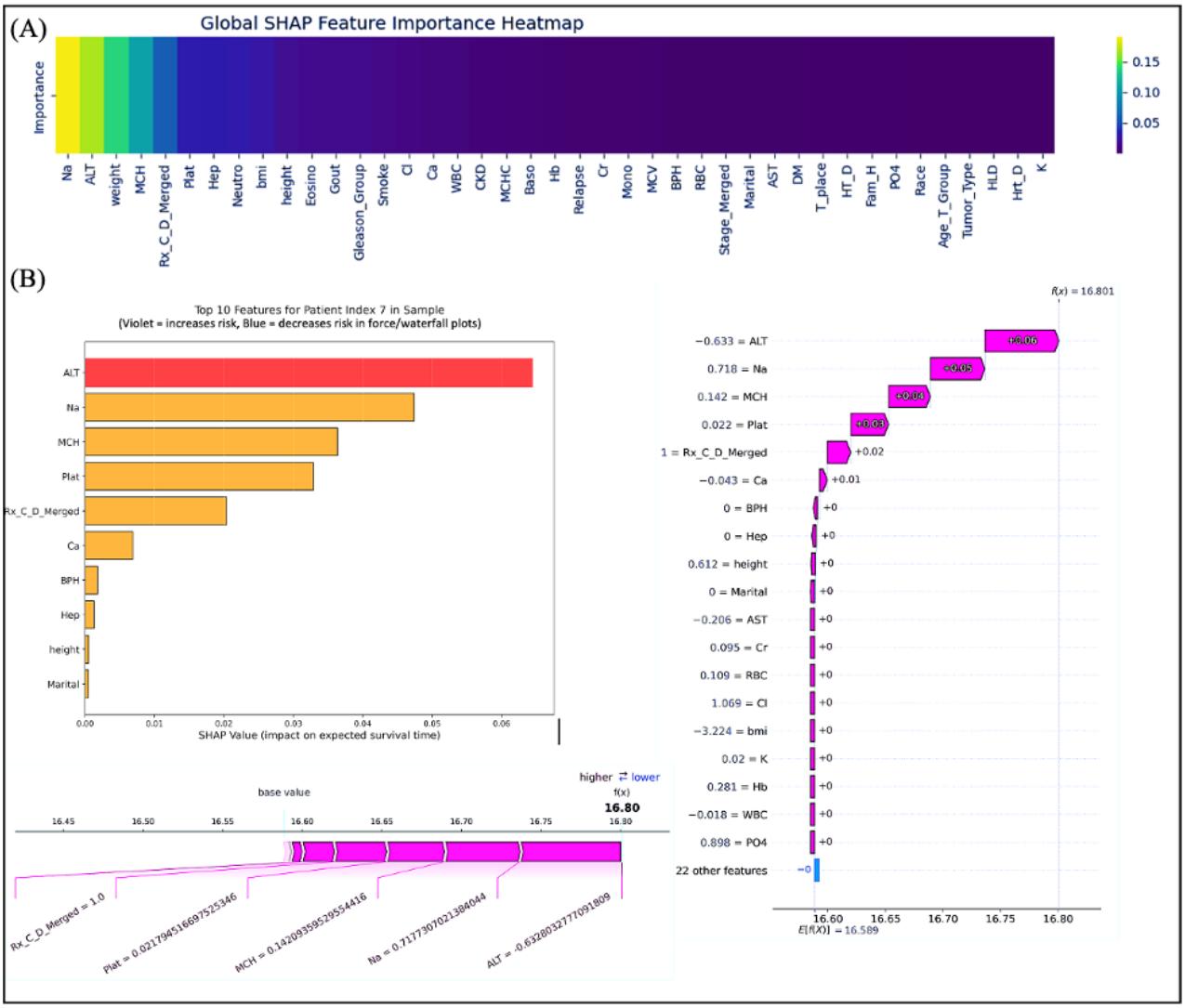
Model explainability for PFS prediction using the RSF was assessed using SHAP, providing both global and individual-level interpretability (Figure 5). At the global level, the SHAP feature importance heatmap (Figure 5A) indicated that biomedical and laboratory parameters, particularly baseline sodium, alanine aminotransferase (ALT), mean corpuscular hemoglobin (MCH) level, weight, and the treatment variable (Rx_C_D_Merged), were the most influential contributors to the model's PFS predictions, while demographic variables showed comparatively lower impact. These features highlighted the dominant role of laboratory biomarkers in capturing progression risk in patients with prostate cancer. SHAP values reflect model attribution, not a causal effect.

At the individual level, with the sample patient, SHAP bar, force and waterfall plots (Figure 5B) illustrated how specific features contributed to the predicted PFS for a representative patient, demonstrating the direction and magnitude of each feature's effect on expected survival time. ALT, sodium, MCH, platelet count, and treatment category increased the predicted risk of progression, whereas other features had minimal or neutral influence. These outputs offer clinical interpretability of how individual patient characteristics contribute to PFS predictions,

XSL•FO
RenderX

thereby enhancing transparency, trust, and potential clinical utility in the prognosis of prostate cancer.

Figure 5. (A) Heatmap of global Shapley additive explanations (SHAP) feature importance for the random survival forest model. (B) Top 10 features by importance for an individual patient, bar, waterfall and force plots. ALT: alanine transaminase; BPH: benign prostatic hyperplasia; Ca: calcium; Cl: chloride; Hb: Hemoglobin; Hep: hepatitis; Na: sodium; Plat: platelet; PSA: prostate-specific antigen; Rx_C_D_Merged: treatment categories.



Discussion

Principal Findings

The study findings demonstrated that ML-based models showed improved internal discrimination compared with the traditional Cox proportional hazards model in predicting PFS among patients with prostate cancer, particularly at long-term horizons. While early PFS prediction at 1 to 3 years was not feasible due to insufficient progression events and limited comparable risk sets, robust discrimination emerged at 4 and 5 years. This reflects the delayed progression patterns and sluggish natural history of prostate cancer that are frequently seen in clinical practice [37-39]. Among the evaluated models, RSF showed the most balanced performance, achieving high time-dependent AUC values and stable calibration as measured by the IBS. These findings highlight the importance of aligning evaluation time points with disease biology and reinforce the limitations of early event prediction risk in prostate cancer cohorts [8,12,40].

Additionally, the study provided an explainability output from the RSF model using SHAP, which allows addressing the "black-box" barrier for clinical adoption of AI in oncology [19]. The SHAP-based global and individual explanations revealed that the laboratory biomarkers and treatment-related variables dominated the prediction of PFS. This aligns with emerging evidence that data-driven survival models capture complex, nonlinear associative predictors of risk that are often missed by conventional regression-based approaches [20,41]. Furthermore, the SHAP force and waterfall plots at the individual patient level enable clinicians to understand not only what the model predicts but how a particular prediction is made, which can further improve trust, auditability, and shared decision-making in clinical oncology [19,42,43].

Moreover, the application of a RAE to model the longitudinal cohort structure enabled effective compression of repeated observations into latent representations, providing a novel and efficient approach for managing complex temporal data while preserving clinically relevant information. The performance of

XSL•FO
RenderX

RSF and GBS models in this cohort was consistent with earlier reports demonstrating the superiority of tree-based methods for cancer survival prediction [19,44]. Additionally, similar to previous studies on prostate cancer, RSF models outperformed Cox models when handling high-dimensional laboratory and longitudinal data, particularly in heterogeneous real-world cohorts [36,45]. The underperformance of Cox regression in this setting supports the rationale for exploring ML survival methods, particularly RSF, which can capture complex nonlinear interactions without proportional hazards assumptions. Although earlier studies reported strong short-term predictive performance, our findings suggest that PFS prediction in prostate cancer is inherently limited at early points [39,40]. The comparatively weaker calibration of the deep neural network survival model highlighted that higher discrimination does not necessarily translate into reliable absolute risk estimation.

According to the SHAP model, sodium, ALT, and hematological parameters such as MCH and platelet count emerged as influential predictors of PFS. Dysnatremia has been associated with cancer progression and systemic inflammation, while elevated ALT may reflect hepatic stress, metabolic dysfunction, or treatment-related toxicity, all of which have been linked to poorer oncological outcomes [46-48]. Moreover, studies have indicated that the aspartate aminotransferase to ALT ratio can be used as a prognostic biomarker for prostate tumor progression [49,50]. These findings support the growing recognition that routine laboratory biomarkers capture systemic physiological states relevant to cancer progression beyond tumor-centric variables alone [51]. However, the prominence of sodium and ALT in our SHAP analysis does not indicate that these markers are biologically superior to other prognostic factors; rather, within this dataset and modeling framework, they emerged as influential predictive contributors that improved the model's ability to estimate risk.

This study has several limitations that should be acknowledged. First, the relatively small sample size, while sufficient for internal model development and evaluation, may restrict the generalizability of the results to other populations, particularly in light of regional differences in clinical practice and disease characteristics. This also limits the model's ability to capture early PFS events. Performance estimates are likely optimistic due to limited event counts. Despite these limitations, our study offers important novel contributions that address key gaps in the literature. No prior study, to our knowledge, has collected the comprehensive range of variables included here or applied a sequential approach to define PFS, whereas most existing studies use a single PFS event per patient, limiting clinical applicability. Second, the retrospective nature of the study may introduce bias due to missing data, unmeasured confounding, and evolving treatment strategies over time. Furthermore, model performance was evaluated using a held-out testing dataset; although cross-validation was applied to reduce overfitting, the absence of external validation remains a limitation.

Third, although the analysis was retrospective and conducted at a single institution, the TBCC functions as the only comprehensive oncology facility in the country and is responsible for the management of nearly all prostate cancer cases nationwide. Nevertheless, the findings highlight the potential utility of ML models incorporating SHAP-based explainability for prognostic risk assessment in prostate cancer. Finally, radiographic progression according to PCWG criteria was not incorporated into the PFS analysis because of the limitations of the imaging data available for this cohort. Future research should include independent external validation and multi-institutional, multiomics datasets with larger and more diverse cohorts to strengthen the generalizability and robustness of these findings.

Conclusions

Our study demonstrates that the ML-based survival models, particularly tree-based ensemble methods, substantially outperform the traditional Cox proportional hazards model in predicting PFS among patients with prostate cancer, especially at longer follow-up horizons. Among the evaluated approaches, the RFS model achieved the most balanced performance, combining strong long-term discrimination with stable calibration, thereby offering a robust framework for prognostic modeling in real-world prostate cancer cohorts. The integration of SHAP-based explainability addressed a key barrier to clinical adoption by providing transparent global and individual-level insights into model predictions highlighting the dominant role of laboratory biomarkers and treatment-related variables in driving PFS risk predictions in this cohort. The application of a RAE to encode longitudinal cohort data further enhanced the model's ability to capture the complex temporal patterns while reducing dimensionality. These findings support the clinical potential of explainable ML models for personalized risk stratification in prostate cancer, while also emphasizing the need for external validation, larger multi-institutional, multiomics datasets, and prospective evaluation prior to routine clinical implementation.

Acknowledgments

The authors would like to thank the Ministry of Health, Brunei Darussalam, and the Brunei Cancer Centre for granting permission to access the required data for this study. We also acknowledge the support of the Medical and Health Research and Ethics Committee for ethics approval and Universiti Brunei Darussalam for guidance and support. The authors are grateful to all individuals and institutions who contributed to the successful completion of this research. During the preparation of this manuscript, the authors used Grammarly (Superhuman Platform Inc) to assist with language refinement. No generative AI tools were used for data analysis, data interpretation, or the generation of scientific conclusions. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

XSL•FO
RenderX

Funding

This work was supported by Universiti Brunei Darussalam through research grant UBD/RSCH/1.6/FICBF/2025/008, awarded to HAR.

Data Availability

The aggregate data and results supporting the findings of this study are included in the published article and its supplementary materials, where applicable. The underlying participant-level dataset is not available because of privacy and confidentiality considerations and to protect the anonymity of study participants. The complete analysis code is publicly available on GitHub [52].

Authors' Contributions

All authors were involved in the conceptualization of this report. HMT was responsible for data curation, investigation, methodology, resources, visualization, and writing of the original draft. HAR was involved in methodology development, supervision, resources, and writing, reviewing, and editing the manuscript. OAM provided supervision and review. LN contributed to supervision, writing, reviewing, and editing the manuscript. MSA and TT contributed to conceptualization, methodology, and resources. All authors approve the content of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Discriminative performance (time-dependent area under the curve, Harrell concordance index) and calibration accuracy (integrated Brier score) of progression-free survival models.
[DOCX File , 13 KB - cancer_v12i1e91510_app1.docx]

Multimedia Appendix 2
Characteristics of participants stratified by progression-free survival.
[DOCX File , 22 KB - cancer_v12i1e91510_app2.docx]

References

1. Prostate. International Agency of Research on Cancer. URL: https://gco.iarc.who.int/media/globocan/factsheets/cancers/27-prostate-fact-sheet.pdf [accessed 2025-12-17]
2. Schafer EJ, Laversanne M, Sung H, Soerjomataram I, Briganti A, Dahut W, et al. Recent patterns and trends in global prostate cancer incidence and mortality: an update. Eur Urol 2025 Mar;87(3):302-313 [FREE Full text] [doi: 10.1016/j.eururo.2024.11.013] [Medline: 39668103]
3. Hassanipour S, Delam H, Arab-Zozani M, Abdzadeh E, Hosseini SA, Nikbakht HA, et al. Survival rate of prostate cancer in Asian countries: a systematic review and meta-analysis. Ann Glob Health 2020 Jan 02;86(1):2 [FREE Full text] [doi: 10.5334/aogh.2607] [Medline: 31976303]
4. Zhou X, Wei Q. Global spatiotemporal trends in prostate cancer burden and its socioeconomic disparities: an observational study from 1990 to 2021. J Clin Oncol 2025 Feb 18;43(5_suppl):325 [FREE Full text] [doi: 10.1200/JCO.2025.43.5_suppl.325]
5. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263 [FREE Full text] [doi: 10.3322/caac.21834] [Medline: 38572751]
6. Brunei Darussalam Cancer Registry Report 2002-2021. Noncommunicable Disease (NCD) Prevention Unit, Ministry of Health Brunei Darussalam. 2022. URL: https://www.researchgate.net/publication/394256941_Brunei_Darussalam_Cancer_Registry_Report_2002-2021 [accessed 2026-09-25]
7. Leong E, Ong SK, Si-Ramlee KA, Naing L. Cancer incidence and mortality in Brunei Darussalam, 2011 to 2020. BMC Cancer 2023 May 22;23(1):466 [FREE Full text] [doi: 10.1186/s12885-023-10962-8] [Medline: 37217954]
8. Halabi S, Vogelzang NJ, Ou SS, Owzar K, Archer L, Small EJ. Progression-free survival as a predictor of overall survival in men with castrate-resistant prostate cancer. J Clin Oncol 2009 Jun 10;27(17):2766-2771 [FREE Full text] [doi: 10.1200/JCO.2008.18.9159] [Medline: 19380448]
9. Halabi S, Roy A, Rydzewska L, Guo S, Godolphin P, Hussain M, et al. Radiographic progression-free survival and clinical progression-free survival as potential surrogates for overall survival in men with metastatic hormone-sensitive prostate cancer. J Clin Oncol 2024 Mar 20;42(9):1044-1054. [doi: 10.1200/JCO.23.01535] [Medline: 38181323]

XSL•FO
RenderX

10. Belin L, Tan A, De Rycke Y, Dechartres A. Progression-free survival as a surrogate for overall survival in oncology trials: a methodological systematic review. Br J Cancer 2020 May;122(11):1707-1714 [FREE Full text] [doi: 10.1038/s41416-020-0805-y] [Medline: 32214230]
11. Scher HI, Morris MJ, Stadler WM, Higano C, Basch E, Fizazi K, et al. Trial design and objectives for castration-resistant prostate cancer: updated recommendations from the Prostate Cancer Clinical Trials Working Group 3. J Clin Oncol 2016 Apr 20;34(12):1402-1418 [FREE Full text] [doi: 10.1200/JCO.2015.64.2702] [Medline: 26903579]
12. Oh EL, Huish W, El-Gamil S, Benson T, Ferguson T. Assessment of patient and clinician perspectives on clinically meaningful extension of progression-free survival in prostate cancer. Eur Urol Open Sci 2024 Nov 12;70:175-182 [FREE Full text] [doi: 10.1016/j.euros.2024.10.018] [Medline: 39611161]
13. Bang S, Ahn YJ, Koo KC. Harnessing machine learning to predict prostate cancer survival: a review. Front Oncol 2025 Jan 10;14:1502629 [FREE Full text] [doi: 10.3389/fonc.2024.1502629] [Medline: 39868377]
14. Bove S, Arezzo F, Cormio G, Silvestris E, Cafforio A, Comes MC, et al. Explainable machine learning for predicting recurrence-free survival in endometrial carcinosarcoma patients. Front Artif Intell 2024 Dec 06;7:1388188 [FREE Full text] [doi: 10.3389/frai.2024.1388188] [Medline: 39712472]
15. Rask Kragh Jørgensen R, Bergström F, Eloranta S, Tang Severinsen M, Bjøro Smeland K, Fosså A, et al. Machine learning-based survival prediction models for progression-free and overall survival in advanced-stage Hodgkin Lymphoma. JCO Clin Cancer Inform 2024 Apr;8:e2300255 [FREE Full text] [doi: 10.1200/CCI.23.00255] [Medline: 38608215]
16. Vershinina O, Sushentsev N, Zaikin A, Blyuss O, Barrett T, Ivanchenko M. Improving the potential for predicting prostate cancer progression in patients on active surveillance using explainable artificial intelligence. Cancers (Basel) 2025 Nov 07;17(22):3598 [FREE Full text] [doi: 10.3390/cancers17223598] [Medline: 41300965]
17. Li Y, Chai X, Yang M, Xiong J, Zeng J, Chen Y, et al. Accurate prediction of disease-free and overall survival in non-small cell lung cancer using patient-level multimodal weakly supervised learning. NPJ Precis Oncol 2025 Jun 19;9(1):197 [FREE Full text] [doi: 10.1038/s41698-025-00981-y] [Medline: 40537513]
18. Chen Z, Ouyang H, Sun B, Ding J, Zhang Y, Li X. Utilizing explainable machine learning for progression-free survival prediction in high-grade serous ovarian cancer: insights from a prospective cohort study. Int J Surg 2025 May 01;111(5):3224-3234 [FREE Full text] [doi: 10.1097/JS9.0000000000002288] [Medline: 39878156]
19. Krzyziński M, Spytek M, Baniecki H, Biecek P. SurvSHAP(t): time-dependent explanations of machine learning survival models. Knowl Based Syst 2023 Feb 28;262:110234 [FREE Full text] [doi: 10.1016/j.knosys.2022.110234]
20. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med Res Methodol 2018 Feb 26;18(1):24 [FREE Full text] [doi: 10.1186/s12874-018-0482-1] [Medline: 29482517]
21. Lee C, Yoon J, van der Schaar M. Dynamic-DeepHit: a deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. IEEE Trans Biomed Eng 2020 Jan;67(1):122-133 [FREE Full text] [doi: 10.1109/TBME.2019.2909027] [Medline: 30951460]
22. Sathekge M, Bruchertseifer F, Vorster M, Lawal IO, Knoesen O, Mahapane J, et al. Predictors of overall and disease-free survival in metastatic castration-resistant prostate cancer patients receiving 225 Ac-PSMA-617 radioligand therapy. J Nucl Med 2020 Jan;61(1):62-69 [FREE Full text] [doi: 10.2967/jnumed.119.229229] [Medline: 31101746]
23. Ponce-Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtmann S. Practical guide to SHAP analysis: explaining supervised machine learning model predictions in drug development. Clin Transl Sci 2024 Nov;17(11):e70056 [FREE Full text] [doi: 10.1111/cts.70056] [Medline: 39463176]
24. Prostate cancer, version 2. National Comprehensive Cancer Network. 2025. URL: https://www.nccn.org/login?ReturnURL=https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf [accessed 2025-05-07]
25. GroupShuffleSplit. Scikit-Learn. URL: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GroupShuffleSplit.html [accessed 2026-09-23]
26. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn 2006 Mar 2;63:3-42 [FREE Full text] [doi: 10.1007/s10994-006-6226-1]
27. Breslow NE. Analysis of survival data under the proportional hazards model. Int Stat Rev 1975 Apr;43(1):45-57 [FREE Full text] [doi: 10.2307/1402659]
28. Kalbfleisch JD, Schaubel DE. Fifty years of the Cox model. Annu Rev Stat Appl 2023;10:1-23 [FREE Full text] [doi: 10.1146/annurev-statistics-033021-014043]
29. In J, Lee DK. Survival analysis: part II - applied clinical data analysis. Korean J Anesthesiol 2019 Oct;72(5):441-457 [FREE Full text] [doi: 10.4097/kja.19183] [Medline: 31096731]
30. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982 Apr;69(1):239-241 [FREE Full text] [doi: 10.1093/biomet/69.1.239]
31. Abeysekera WW, Sooriyarachchi MR. Use of Schoenfeld's global test to test the proportional hazards assumption in the Cox proportional hazards model: an application to a clinical study. J Natn Sci Foundation Sri Lanka 2009 Mar 29;37(1):41-45 [FREE Full text] [doi: 10.4038/jnsfsr.v37i1.456]
32. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994 Sep;81(3):515-526 [FREE Full text] [doi: 10.1093/biomet/81.3.515]

XSL•FO
RenderX

33. Berkowitz M, Altman RM, Loughin TM. Random forests for survival data: which methods work best and under what conditions? Int J Biostat 2024 Apr 24;20(2):315-345 [FREE Full text] [doi: 10.1515/ijb-2023-0056] [Medline: 38656274]
34. Chen Y, Jia Z, Mercola D, Xie X. A gradient boosting algorithm for survival analysis via direct optimization of concordance index. Comput Math Methods Med 2013;2013:873595 [FREE Full text] [doi: 10.1155/2013/873595] [Medline: 24348746]
35. Adnan M, Alarood AA, Uddin MI, Ur Rehman I. Utilizing grid search cross-validation with adaptive boosting for augmenting performance of machine learning models. PeerJ Comput Sci 2022 Feb 21;8:e803 [FREE Full text] [doi: 10.7717/peerj-cs.803] [Medline: 35494796]
36. Mogensen UB, Ishwaran H, Gerds TA. Evaluating random forests for survival analysis using prediction error curves. J Stat Softw 2012 Sep;50(11):1-23 [FREE Full text] [doi: 10.18637/jss.v050.i11] [Medline: 25317082]
37. Pound CR, Partin AW, Eisenberger MA, Chan DW, Pearson JD, Walsh PC. Natural history of progression after PSA elevation following radical prostatectomy. JAMA 1999 May 05;281(17):1591-1597 [FREE Full text] [doi: 10.1001/jama.281.17.1591] [Medline: 10235151]
38. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. 10-year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer. N Engl J Med 2016 Oct 13;375(15):1415-1424 [FREE Full text] [doi: 10.1056/NEJMoa1606220] [Medline: 27626136]
39. Wilt TJ, Jones KM, Barry MJ, Andriole GL, Culkin D, Wheeler T, et al. Follow-up of prostatectomy versus observation for early prostate cancer. N Engl J Med 2017 Jul 13;377(2):132-142 [FREE Full text] [doi: 10.1056/NEJMoa1615869] [Medline: 28700844]
40. Kattan MW, Wheeler TM, Scardino PT. Postoperative nomogram for disease recurrence after radical prostatectomy for prostate cancer. J Clin Oncol 1999 May;17(5):1499-1507 [FREE Full text] [doi: 10.1200/JCO.1999.17.5.1499] [Medline: 10334537]
41. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. New York, NY: Curran Associates Inc; 2017:4768-4777.
42. Tun HM, Rahman HA, Naing L, Malik OA. Trust in artificial intelligence-based clinical decision support systems among health care workers: systematic review. J Med Internet Res 2025 Jul 29;27:e69678 [FREE Full text] [doi: 10.2196/69678] [Medline: 40772775]
43. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. Wiley Interdiscip Rev Data Min Knowl Discov 2019;9(4):e1312 [FREE Full text] [doi: 10.1002/widm.1312] [Medline: 32089788]
44. Adatorwovor R, Ogunsanya ME, Huang B, Charnigo R, Abraham O. Comparison of machine learning models for colon cancer survival: predictive modeling approach. JMIR Cancer 2025 Nov 26;11:e72665 [FREE Full text] [doi: 10.2196/72665] [Medline: 41297023]
45. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface 2018 Apr;15(141):20170387 [FREE Full text] [doi: 10.1098/rsif.2017.0387] [Medline: 29618526]
46. Stangl-Kremser J, Kramer G, Shariat SF. Association of hyponatremia with survival in patients with castration-resistant prostate cancer: a clinical commentary. Clin Genitourin Cancer 2019 Dec;17(6):e1188-e1192 [FREE Full text] [doi: 10.1016/j.clgc.2019.08.001] [Medline: 31623929]
47. Djamgoz MB. Hyponatremia and cancer progression: possible association with sodium-transporting proteins. Bioelectricity 2020 Mar 01;2(1):14-20 [FREE Full text] [doi: 10.1089/bioe.2019.0035] [Medline: 34471833]
48. Bañez LL, Loftis RM, Freedland SJ, Presti JCJ, Aronson WJ, Amling CL, et al. The influence of hepatic function on prostate cancer outcomes after radical prostatectomy. Prostate Cancer Prostatic Dis 2010 Jun;13(2):173-177 [FREE Full text] [doi: 10.1038/pcan.2010.3] [Medline: 20195294]
49. Mitsui Y, Yamabe F, Hori S, Uetani M, Aoki H, Sakurabayashi K, et al. Longitudinal change in castration-resistant prostate cancer biomarker AST/ALT ratio reflects tumor progression. Sci Rep 2023 Sep 15;13(1):15292 [FREE Full text] [doi: 10.1038/s41598-023-42711-z] [Medline: 37714917]
50. Zhou J, He Z, Ma S, Liu R. AST/ALT ratio as a significant predictor of the incidence risk of prostate cancer. Cancer Med 2020 Aug;9(15):5672-5677 [FREE Full text] [doi: 10.1002/cam4.3086] [Medline: 32562455]
51. Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discov 2022 Jan;12(1):31-46 [FREE Full text] [doi: 10.1158/2159-8290.CD-21-1059] [Medline: 35022204]
52. Explainable-machine-learning-PSF-. GitHub. URL: https://github.com/heinminntun/Explainable-Machine-Learning-PSF- [accessed 2026-09-23]

Abbreviations

ALT: alanine aminotransferase
AUC: area under the receiver operating characteristic curve
C-index: concordance index
GBS: gradient boosting survival

XSL•FO
RenderX

GRU: gated recurrent unit
IBS: integrated Brier score
JPMC: Jerudong Park Medical Centre
MCH: mean corpuscular hemoglobin
ML: machine learning
PCWG: Prostate Cancer Working Group
PFS: progression-free survival
PSA: prostate-specific antigen
RAE: recurrent autoencoder
RSF: random survival forests
SHAP: Shapley additive explanations
TBCC: Brunei Cancer Centre

Edited by M Balcarras; submitted 20.Jan.2026; peer-reviewed by R Ismayilov, C Colak, S Banerjee, J Guo, V Palama, A Arora; comments to author 07.Apr.2026; revised version received 08.Apr.2026; accepted 09.Apr.2026; published 05.Oct.2026.

Please cite as:
Tun HM, Naing L, Malik OA, Abdullah MS, Ta T, Rahman HA
Explainable Machine Learning–Based Prediction of Progression-Free Survival in Prostate Cancer: Retrospective Cohort Study
JMIR Cancer 2026;12:e91510
URL: https://cancer.jmir.org/2026/1/e91510
doi:10.2196/91510
PMID:

©Hein Minn Tun, Lin Naing, Owais Ahmed Malik, Muhammad Syafiq Abdullah, Thu Ta, Hanif Abdul Rahman. Originally published in JMIR Cancer (https://cancer.jmir.org), 05.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Scan, Screen, Support—A Digital Pathway for Assessing Supportive Care Needs in Oncology: Implementation Study

Theres Fey[1,2], PhD; Nicole Thurner[1], MSc; Ulrike Haidn[1], MSc; Friederike Mumm[1,3], MD; Birgit Haberland[4], MD; Rachel Wuerstlein[1,5], MD; Sebastian Theurich[3], MD; Georg Wolfrum[6], PhD; Marie-Louise Troidl[1]; Susan Müller[1], MSc; Claudia Bausewein[4], MD; Timo Schinköthe[7,8], MD; Volker Heinemann[1], MD; Nicole Erickson[1], PhD

[1]Comprehensive Cancer Center (CCC Munich LMU), LMU Hospital, Pettenkoferstr. 8a, München, Bavaria, Germany

[2]Bavarian Cancer Research Center (BZKF), Germany

[3]Department of Medicine III, LMU Hospital, München, Bavaria, Germany

[4]Department of Palliative Medicine, LMU Hospital, München, Germany

[5]Department of Obstetrics and Gynecology, LMU Hospital, München, Germany

[6]Department for Medical Technology and IT, LMU Hospital, München, Germany

[7]CANKADO GmbH, München, Germany

[8]Research Center Smart Digital Health, University of the Bundeswehr Munich, München, Germany

Corresponding Author:
Theres Fey, PhD
Comprehensive Cancer Center (CCC Munich LMU), LMU Hospital, Pettenkoferstr. 8a, München, Bavaria, Germany

Abstract

Background: Electronic patient-reported outcome tools have the potential to enhance supportive care in oncology and support the timely and accurate identification of patients' needs.

Objective: This study aimed to develop, implement, and evaluate a user-friendly, web-based digital screening tool at a German Comprehensive Cancer Center that systematically and efficiently assesses the supportive care needs of patients with cancer and enables direct referral to appropriate supportive services through seamless integration with the hospital information system.

Methods: An interdisciplinary team collaborated with the IT department and the company CANKADO, an electronic patient-reported outcome provider, to create a 14-item digital questionnaire. The tool incorporated validated instruments, such as the Distress Thermometer, the Nutrition Risk Screening, and a short form of the Integrated Palliative Care Outcome Scale, aligned with German Cancer Society certification criteria. Patients accessed the questionnaire via QR codes. Screening results were automatically transferred to the hospital information system, where supportive care requests (SCRs) were generated automatically if indicated.

Results: Between June 2024 and May 2025, a total of 8855 QR codes were generated. Of these, 4909 questionnaires were complete and valid for analysis. This information produced 3324 SCRs. Digital screenings resulted in an SCR in 22.4% of cases for psycho-oncology, 18.7% for nutrition, and 27.6% for palliative care. The digital screening maintained or slightly improved screening rates compared to prior methods.

Conclusions: The implementation of a digital supportive care screening was feasible and effective within the Comprehensive Cancer Center setting. Future efforts should focus on overcoming barriers for patients with limited digital access or capabilities to ensure the delivery of equitable supportive care.

(JMIR Cancer 2026;12:e82392) doi:10.2196/82392

KEYWORDS

electronic patient-reported outcomes; ePROs; patient-reported outcomes; PROs; digitalization; routine screening; digital screening; supportive cancer care; psycho-oncology; distress; nutrition screening; palliative care

Introduction

Contemporary oncology acknowledges that supporting patients with cancer through their journey is essential. The diagnosis and treatment of cancer entail physical challenges as well as psychological and social burdens. Consequently, providing care that holistically addresses these multifaceted needs is of precedence [1,2]. Ensuring that individualized, need-based care is integrated into the treatment process requires systematic assessment of these needs [3]. Patient-reported outcome (PRO) assessments in the context of psycho-oncological and palliative care screenings are already established in Comprehensive Cancer

XSL•FO
RenderX

Centers (CCCs) across Germany and are incorporated into clinical guidelines [4-8].

The German Cancer Society recommends the implementation of an interdisciplinary screening procedure to evaluate counseling needs as an integral component of basic oncological assessment [9]. Historically, supportive needs were assessed via interviews or paper-based questionnaires. These methods, however, face limitations with regard to real-time data collection, structured documentation, and accessibility [10]. Thus, digital solutions for the collection of PROs have the potential to optimize support for patients with cancer by facilitating fast and precise identification of support needs while reducing the burden on health care professionals (HCPs) [11-13]. Furthermore, electronic patient-reported outcomes (ePROs) interventions have been shown to have a positive impact on health-related quality of life and symptom burden [13,14].

Notably, digital tools enable continuous and standardized assessment of patient needs across various treatment stages. This consistency is crucial, given that a patient's need for supportive care often evolves throughout the disease trajectory [11].

We therefore aimed to develop and implement a user-friendly, web-based digital screening tool to enable rapid and accurate collection of patient needs within a German CCC. The tool was designed to facilitate direct access to supportive services while ensuring seamless integration with the hospital information system (HIS). In parallel, challenges related to the digital implementation within routine clinical practice were examined.

Methods

Implementation Design

The implementation-oriented design of the project was guided by the EPIS (Exploration, Preparation, Implementation, Sustainment) framework, which provides a structured approach to implementing innovations in health care settings [15,16]. The project comprised 4 key phases corresponding to the EPIS framework.

Exploration

During the exploration phase (2017 - 2020), various platforms for collecting PROs were evaluated within the context of clinical studies at the CCC [17,18], as well as in the Department of Gynecology [19,20] and the Department of Hematology and Oncology [21]. Following this evaluation phase, the CANKADO platform was selected as a digital platform for web-based questionnaire delivery due to its ease of use, sustainability, and affordability [22]. An interdisciplinary working group was established, comprising experts from the fields of medicine and oncology, nutrition science, psycho-oncology, palliative care, and IT. Initial preparations for the implementation of the digital screening system began in 2020 but were delayed as a result of the COVID-19 pandemic. Compliance with information security and data protection regulations was ensured at all times (Textbox 1).

XSL•FO
RenderX

Textbox 1. Implementation steps using the Exploration, Preparation, Implementation, Sustainment (EPIS) framework.

Exploration (2017-2020)

- Formation of an interdisciplinary working group
- Development of a requirements catalog for a browser-based questionnaire platform
- Vendor screening and market research
- Pilot projects with 2 vendors
- Decision to partner with CANKADO's ePRO system
- Feasibility studies with CANKADO

Preparation (2020-2023)

- Questionnaire development
- Establishment of required IT infrastructure (automated patient registration, integration into the HIS, generation of QR codes)
- Programming and technical implementation within CANKADO and the HIS
- Data protection measures

Implementation (2023-2025)

- Pilot implementation
- Adaptation of the questionnaire based on pilot results
- Engagement and endorsement by medical and nursing leadership
- Training of HCPs in oncological departments
- Hospital-wide rollout
- Development of patient information materials

Sustainment (since 2025; ongoing)

- Refresher trainings for HCPs
- Continuous monitoring of screening uptake and effectiveness
- Implementation of support measures for patients with barriers to digital participation
- Evaluation of acceptance and use by HCPs (not part of this manuscript)
- Evaluation of patient acceptance (not part of this manuscript)

Preparation

The preparation phase involved the development of the PRO questionnaire, digital implementation, and ensuring data protection compliance.

Questionnaire Development

A tailored questionnaire, composed of existing validated short-form questionnaires, was developed in close coordination with the relevant disciplines (Department of Psycho-Oncology, Nutrition, and Palliative Care). This method ensured alignment with the certification requirements of the German Cancer Society (see Table 1). The National Comprehensive Cancer Network Distress Thermometer, the short version of the Integrated Palliative Care Outcome Scale (IPOS), and the Nutrition Risk Screening (NRS) 2002 served as the primary sources for item selection [23-26].

XSL•FO
RenderX

Table . Questionnaire.

Category and validated questionnaire		Question
Psycho-oncology (2 questions)[a]		
	NCCN[b] Distress Thermometer	1. Please circle the number (0 - 10) that best describes how much distress you have been experiencing in the past week including today.
	Request	2. Would you like psychological counseling? Yes / No / No answer
Nutrition (5 questions)[c]		
	NRS[d] and PG-SGA[e] SF	3. I currently weigh about _____ kg.
	NRS and PG-SGA SF	4. I am about _____ cm tall.
	NRS and PG-SGA SF	5. Three months ago, I weighed about _____ kg.
	NRS and PG SGA SF	6. How much did you eat last week compared to your usual amount? (Please mark a value between 0% - 150%)
	Request	7. Would you like nutritional counseling? Yes / No / No answer
Palliative care (6 questions)[f]		
	IPOS[g] (introductory text)	For each symptom, please tick the box that best describes how it has affected you over the past 3 days.
	IPOS	8. Pain: Not at all (0), Slightly (1), Moderately (2), Severely (3), Overwhelmingly (4)
	IPOS	9. Shortness of breath: Not at all (0), Slightly (1), Moderately (2), Severely (3), Overwhelmingly (4)
	IPOS	10. Poor mobility: Not at all (0), Slightly (1), Moderately (2), Severely (3), Overwhelmingly (4)
	IPOS	11. Have you been feeling anxious or worried about your illness or treatment? Not at all (0), Occasionally (1), Sometimes (2), Most of the time (3), Always (4)
	IPOS	12. Have any of your family or friends been anxious or worried about you? Not at all (0), Occasionally (1), Sometimes (2), Most of the time (3), Always (4)
	IPOS	13. Have any practical problems resulting from your illness been addressed? (such as financial or personal) Problems addressed/no problems (0), Problems mostly addressed (1), Problems partly addressed (2), Problems hardly addressed (3), Problems not addressed (4)
General (1 question)[h]		
	IPOS	14. How did you complete this questionnaire? On my own / With help from a friend or relative / With help from a member of staff

[a]A supportive care request (SCR) is automatically generated if the Distress score is >4 and psychological counseling is requested or not specified, or if psychological counseling is requested regardless of the Distress score.

[b]NCCN: National Comprehensive Cancer Network.

[c]An SCR is automatically generated if weight loss is >5% within 3 months, or BMI <20.5 kg/m², or food intake <75%, and nutritional counseling is requested or not specified, or if nutritional counseling is requested regardless of nutritional values.

[d]NRS: Nutrition Risk Screening.

XSL•FO
RenderX

[e]PG-SGA: Patient-Generated Subjective Global Assessment.

[f]An SCR is automatically generated if the IPOS-5 score is ≥8 or the IPOS-6 score is ≥10.

[g]IPOS: Integrated Palliative Care Outcome Scale.

[h]This question was taken from the IPOS but applies to the entire questionnaire.

For the identification of malnutrition and the need for nutrition interventions, the NRS was already the established screening tool in the hospital. However, the questions in the NRS were designed and validated to be completed by HCPs. For example, the first question asked whether the body mass index is less than 20.5 kg/m^2, which cannot be answered by a patient using a PRO tool. Therefore, the risk assessment points from the NRS tool were reworded analog to the wording validated in the Patient-Generated Subjective Global Assessment-Short Form, which is a validated PRO tool [23]. The cut-off points for the risk assessment remained in line with the NRS. Based on recent research results, the questions of the IPOS were adapted from IPOS-5 to IPOS-6 in April 2025 [27,28].

Key requirements for the overall questionnaire included brevity, ease of understanding for patients, and applicability across various stages of the disease. The final questionnaire, consisting of 14 items, serves as a prescreening tool. If necessary, a more detailed screening can be conducted by the relevant medical specialty. Regardless of the total score, patients also have the opportunity to actively express a desire for counseling services in the area of psycho-oncology or nutrition (items 2 and 7 of the questionnaire). This question was not included for palliative care.

A screening was considered valid if 1 or more of the following criteria were met: for the distress screening, at least 1 question had to be answered; for the nutrition screening, at least 1 question had to be answered; for the palliative screening using IPOS-5, at least 2 questions had to be answered; and for the palliative screening using IPOS-6, at least 3 questions had to be answered. Screenings that were accessed but had no items completed were not included.

Digital Implementation

The CANKADO ePRO system was initially developed for collecting PRO in clinical trials [20,29-31]. It utilizes a cloud-based electronic health record with a role-based access concept. Development and operation are carried out under ISO 27001 certification. Patients can access their ePRO questionnaires through personalized access via the web or an app. In addition, specific questionnaires can be accessed directly via a QR code. These QR codes are implemented to make access possible regardless of the device and without an additional app. Alternatively, the questionnaire can be completed via a tablet provided by the center and secured by kiosk software.

Based on this existing architecture, the CANKADO ePRO-based questionnaire was integrated into the HIS for use in routine clinical operations in oncology. The aim was to automate the entire process for HCPs and make it fully mappable within the HIS. To enable the integration of CANKADO ePRO into the hospital's existing SAP-based HIS, several components were required: (1) retrieval of a QR code linking to the digital screening questionnaire, (2) receipt of the completed digital screening as both a PDF document and structured data, (3) display of an overview of the most recent survey within a subsystem, and (4) storage of data for subsequent evaluation.

Within the existing HIS, a button was integrated to generate and retrieve the QR code. This request is routed via a Python-based HTTPS service to CANKADO. When a digital screening questionnaire is triggered in the HIS for a specific patient (based on the admission number) for the first time, this initiates the creation of a new, pseudonymized patient file in CANKADO. This file is assigned a generated identifier, which is then sent back to the HIS. At the same time, the corresponding questionnaire is assigned to the newly created patient number and coupled with a QR code, which is transmitted to the HIS. A base64-encoded image of the QR code is securely retrieved and displayed through the HTTPS service of the hospital. The generated QR code can be printed by the nursing staff and then handed to the patient, who completes the questionnaire digitally. The QR code can be regenerated for further assessments of the same patient.

After completion of the questionnaire, the data are automatically transmitted to the HIS. For security reasons, the data are received by a server located in a demilitarized zone (DMZ). The assignment to the real person takes place exclusively within the HIS. For each patient, the PDF document of the completed digital questionnaire is archived with a timestamp in the HIS. The data derived from the digital screening are additionally stored in accordance with the requirements set by the auditing bodies of the certified cancer centers in Germany. An entry in the documentation system is visible to show that a digital screening was performed. Based on the answers, a supportive care request (SCR) is automatically generated via web services (as shown in Table 1). In the final step, a physician reviews the clinical indication and grants approval.

HCPs can additionally be granted permission to access all deidentified patient records via the CANKADO web portal. There, they can view the completed questionnaires and calculated scores and export them if necessary. The structured data from the HIS are provided in JSON format and stored in a MongoDB database to facilitate efficient data evaluation and analysis. The following data can thus be analyzed in a structured manner: number of screenings performed and number of SCRs by the special departments (psycho-oncology, nutrition, palliative care).

Data Protection

The entire process was conducted in accordance with applicable data protection regulations. Approval was granted by the institutional data protection officer prior to implementation. A detailed data processing overview is shown in Table 2.

XSL•FO
RenderX

Table . Data processing overview.

Steps	Collection of personal data	Who accesses the data?	Duration of storage
Initiation			
ePRO creation by CANKADO	No personal data collected	No access	No storage
Creation of user accounts for medical staff	Data collected: 1. Email address 2. Name (optional) 3. First name (optional) 4. Place of employment 5. Address of workplace 6. Employment relationship (eg, study nurse)	Access by: • Trained CANKADO staff	Until the project ends, unless the account is no longer used or the user leaves the organization. Otherwise, until the user deletes the account.
Adding a patient	Data collected: 1. Patient ID (automatically generated by the system) 2. Year of birth (for studies) 3. Other basic data (optional)	Access by: • Trained CANKADO staff (only Patient ID) • Medical staff from the clinic/practice	Until the project ends, unless legal retention periods must be observed. Archiving and storage until the end of legal periods.
Documentation of patient data	Data collected: Project-dependent health data (eg, health status)	Access by: • Medical staff from the clinic/practice	Until the project ends, unless legal retention periods must be observed. Archiving and storage until the end of legal periods.
Project-specific evaluation of patient data	Data collected: Purpose-related analysis of anonymized patient data	Access by: • Selected staff from the clinic/practice	Until the project ends, unless legal retention periods must be observed. Archiving and storage until the end of legal periods.
End			

Implementation

The pilot phase was conducted from April to May 2023 on an outpatient oncology ward. Insights gained during this period regarding the process flow, the stakeholders involved, and the questionnaire itself were used to refine the instrument and address procedural challenges. Starting in June 2024, the interdisciplinary digital screening was rolled out across all oncology inpatient wards and several outpatient units, marking its integration into routine clinical care. To ensure consistent implementation, training sessions were organized with nursing staff and physicians in each department beginning in May 2024. These sessions were systematically conducted by designated, qualified CCC employees. The screening process is shown in Figure 1. The process, along with the designated responsibilities, was outlined in a standard operating procedure made accessible to all HCPs. Barriers and challenges were regularly addressed and evaluated during recurring meetings of the working group, where ongoing developments were also critically discussed.

XSL•FO
RenderX

Figure 1. Screening procedure. HIS: hospital information system.

XSL•FO
RenderX

The implementation of digital screening was first piloted in a single outpatient clinic. This clinic was selected due to its relatively new organizational structure, where routine, comprehensive distress screening had not yet been fully implemented at the start of the project. Distress screening rates were compared between two equivalent six-month periods: July to December 2023 (analog screening only) and July to December 2024 (a combination of analog and digital screening).

Sustainment

The sustainment phase focused on embedding the digital screening into routine clinical practice and ensuring long-term adoption. HCPs' acceptance was assessed through an evaluation in early 2025, and ongoing support structures—including patient navigators for patients facing digital barriers—were introduced to maintain equity and inclusivity. Structured data storage and integration with hospital systems allow for continuous monitoring of screening uptake and effectiveness.

Ethical Considerations

This study used fully anonymized data analyses and involved no intervention. All data were extracted from the hospital patient information system and fully anonymized prior to analysis, precluding any direct or indirect patient reidentification. Thus, the local ethical committee provided written confirmation that no formal ethical consultation was necessary (application number 24 - 0535 KB). An informed consent was not required, and no compensation was provided to participants.

Results

HCP Training

Between May 2024 and May 2025, a total of 48 training sessions were conducted at 32 oncology wards and departments within the LMU Hospital. Thirty-two training sessions were conducted for nursing staff, and 13 were conducted for medical staff. Additionally, 3 training sessions were held for specialized departments involved in providing supplementary counseling services (psycho-oncology, nutrition, and the palliative care team). Overall, 43 of 59 (72.9%) organizational units involved in the care of patients with oncology received training on digital screening procedures. In this phase, the focus was on inpatient units and day clinics; most outpatient areas were not yet included.

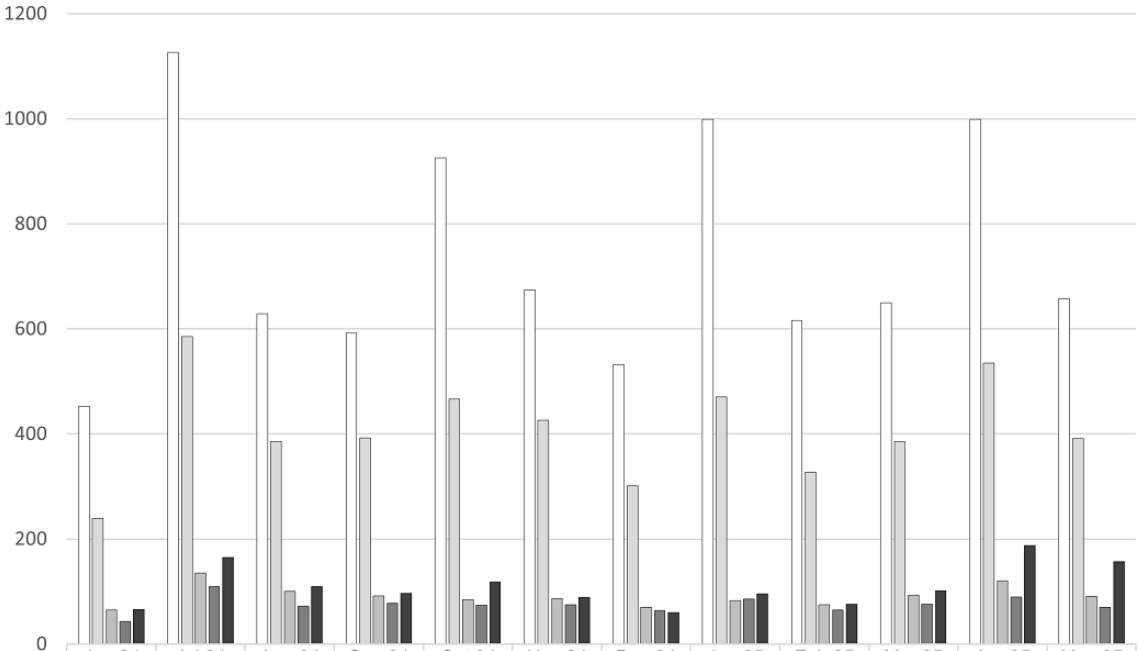
12-Month Implementation Period

After the implementation of the digital screening, a total of 8855 QR codes (738 on average per month) were generated between June 2024 and May 2025. A total of 4909 (55.4%) questionnaires (409 on average per month) were completed and deemed valid. This means that, on average, 329 (44.6%) QR codes per month were either not accessed or accessed without any items being answered. In 946 (10.7%) instances, the questionnaire was accessed via the QR code, but no items were filled out. A total of 3822 (77.9%, 3822/4909) evaluable screenings were performed by the patient only once, while 1087 of 4909 (22.1%) evaluable screenings were performed at least twice.

Among the 4784 responses to the question about how the questionnaire was completed, 80.1% (n=3831) were completed independently by the patients, 15.6% (n=746) with assistance from relatives, and 4.3% (n=207) with support from staff. The screenings resulted in a total of 3324 SCRs, including 1097 (22.4%) for psycho-oncology, 903 (18.7%) for nutritional counseling, and 1324 (27.6%) for palliative care. The analysis per month is shown in Figure 2. The number of QR codes created was highest at the beginning of each quarter.

XSL•FO
RenderX

Figure 2. Development of the generated QR codes, the screenings conducted (at least 1 question answered; screenings that were accessed but had no items completed were not included), and the supportive care requests (SCRs) created in the period from June 2024 to May 2025.

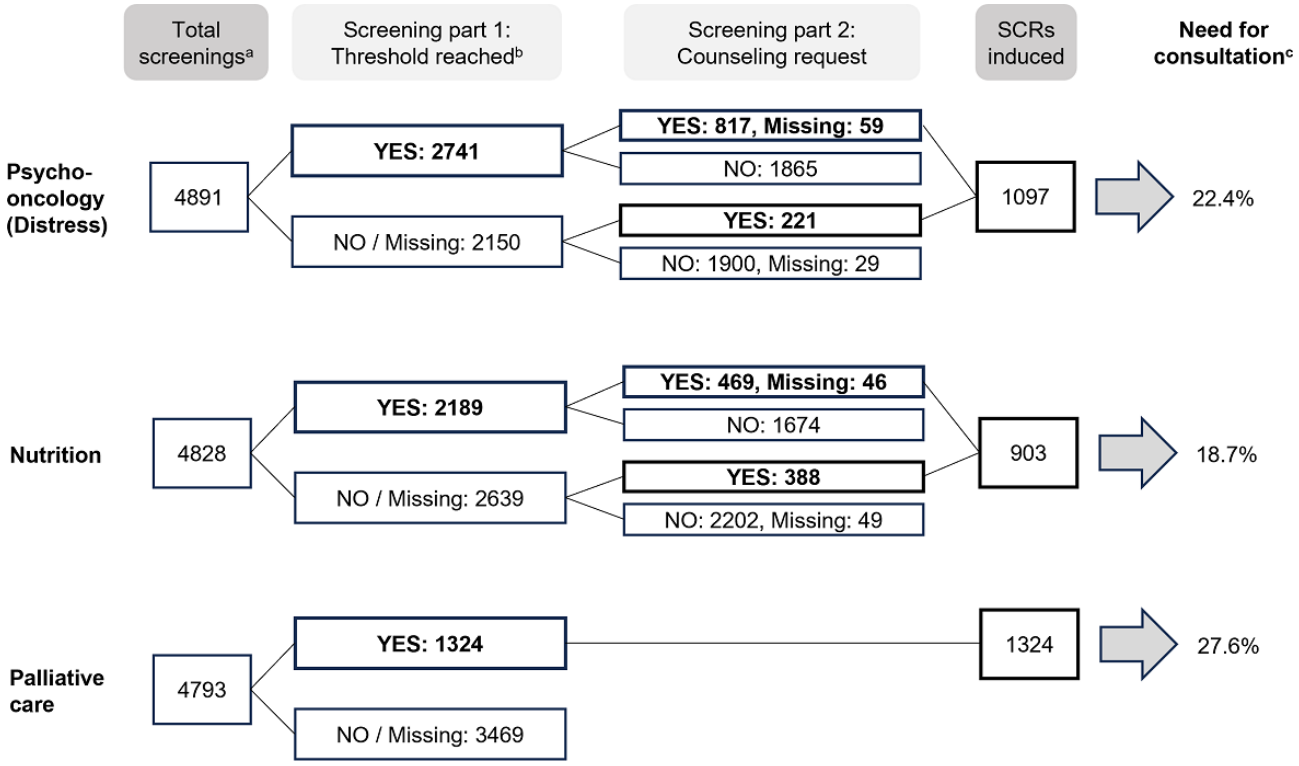


	Jun 24	Jul 24	Aug 24	Sep 24	Oct 24	Nov 24	Dec 24	Jan 25	Feb 25	Mar 25	Apr 25	May 25
QR codes generated	453	1127	629	592	926	674	532	999	616	650	999	658
Digital screenings	239	585	386	393	467	426	302	471	327	386	535	392
SRC psycho-oncology	65	135	101	92	85	87	70	83	75	93	120	91
SCR nutrition	43	110	72	78	74	75	64	86	65	76	90	70
SCR palliative care	66	165	110	97	118	89	60	96	76	102	188	157

In total, 4891 patients' screenings indicated a distress value and/or the patient requested a psychosocial consultation. Of these, 56% (n=2741) had an increased distress value of more than 4 points (see Figure 3). An SCR was triggered for 32% (n=876) of these screenings, which had an increased distress score. Accordingly, no SCR was triggered for 68% (n=1865) of the people with an increased distress score, as they had rejected psycho-oncology counseling. Fifty-nine SCRs were induced due to an increased distress value, and there was no information available on whether the patient had a desire for counseling. The majority of SCRs (n=817, 74.5%) were induced by a combination of an increased distress value and a request for counseling by the patient on the digital screening. Less frequently, the SCR was triggered solely due to the request for counseling (n=221, 20.1%), as the distress value was below the threshold rate and/or not specified. The overall demand for psycho-oncological counseling was 1097 out of 4891 (22.4%).

XSL•FO
RenderX

Figure 3. Evaluation of questionnaires in the areas of psycho-oncology (distress screening), nutrition screening, and palliative care screening, as well as the initiated SCRs (supportive care requests). [a]A screening was considered valid when 1 or more of the following fields were completed: distress screening, at least 1 question answered; nutrition screening, at least 1 question answered; palliative screening (IPOS-5), at least 2 questions answered; palliative screening (IPOS-6), at least 3 questions answered. [b]Screening thresholds are detailed in Table 1. [c]The need for consultation is determined by the number of initiated SCRs in relation to the total number of screenings performed.



A total of 4828 questionnaires were completed with at least 1 item of the nutrition screening and/or the patients provided information about their need for consultation. Of these, 45.3% (n=2189) met the nutrition-specific thresholds. An SCR was triggered for 23.5% (n=515) of these screenings. Accordingly, no SCR was triggered for 76.5% (n=1674) of the screenings with corresponding nutritional values, as the patient had rejected nutrition counseling. A total of 469 of 903 (51.9%) of the SCRs for nutrition were triggered because the threshold value was exceeded and the patient requested nutrition counseling. Additionally, 388 of 903 (43%) SCRs were patient requests for nutrition counseling although the threshold value was not reached or specified. Furthermore, 46 of 903 (5.1%) SCRs were induced due to an increased nutrition threshold value, while there was no information available on whether the patient had a desire for counseling. In summary, the overall need for nutritional counseling was 903 of 4828 (18.7%).

A total of 1324 of 4793 (27.6%) of the patients who completed or partially completed the screening for palliative care (n=4793) also had an elevated IPOS score.

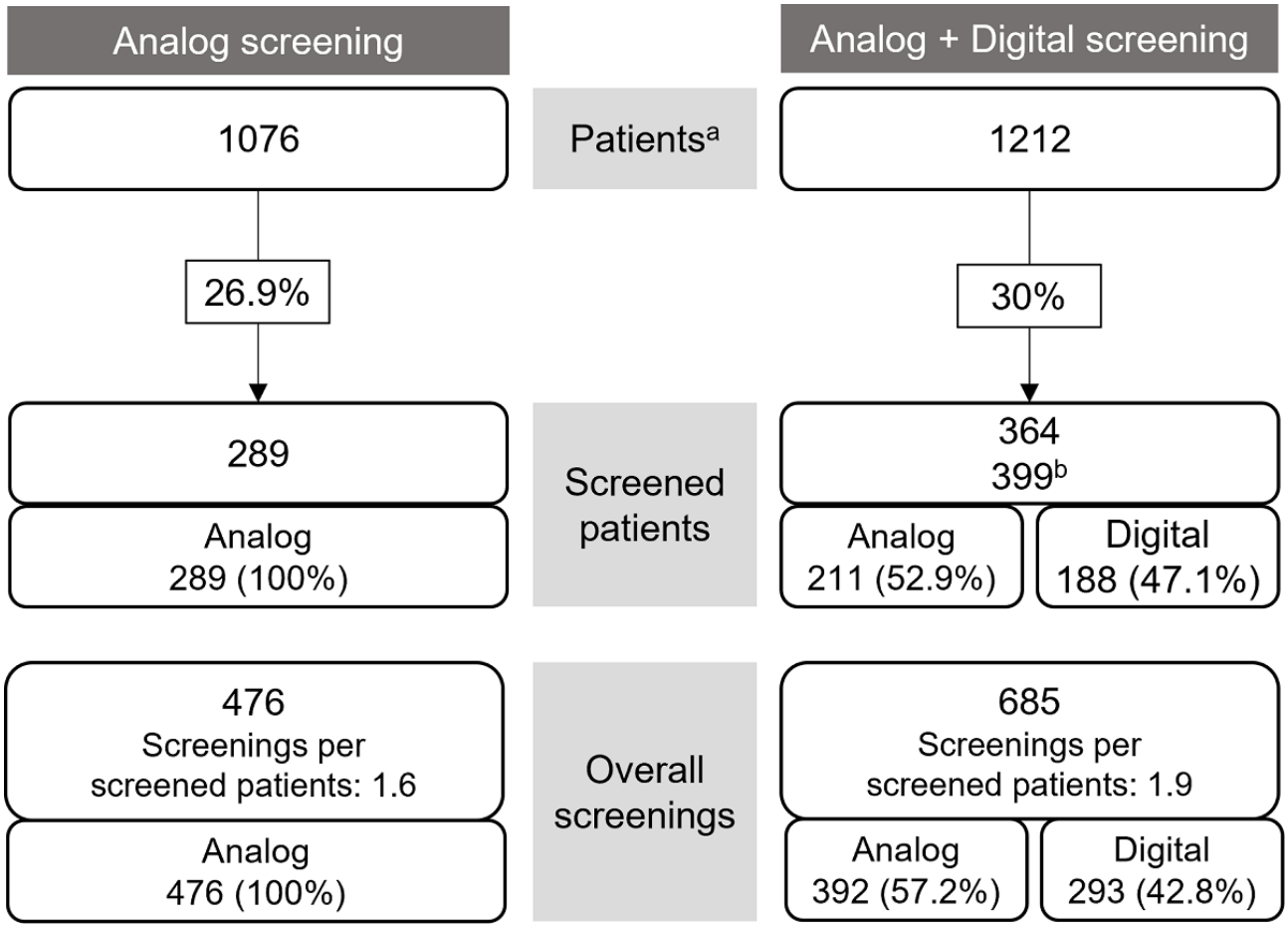
Overall, 1 or more counseling needs were identified and triggered in 3072 of 4909 (62.6%) of the screenings. Accordingly, 1837 of 4909 (37.4%) of the screenings revealed no unmet needs.

Comparison of Development: Analog vs Digital Derived From a Single Outpatient Clinic Within the LMU Hospital

Distress screening rates were compared in 2 comparable 6-month periods: between July and December 2023 (analog screening only) and July to December 2024 (a combination of analog and digital screening). All relevant staff of the outpatient clinic were trained in the use of digital screening in May 2024. Since the transition to digital screening was gradual, analog screenings continued to be conducted alongside the digital ones during this period. Nevertheless, after the introduction of digital screening, the percentage of screened patients increased from 26.9% to 30%. In the 6 months following the implementation, 57.2% of the distress screenings were still conducted using the analog method, while 42.8% were carried out using the digital screening (see Figure 4). The number of screenings per screened patient increased from an average of 1.6 to 1.9 after the implementation.

XSL•FO
RenderX

Figure 4. Comparison of distress screenings conducted over two 6-month periods at the oncological outpatient clinic. From July to December 2023 (6 months), distress screening was still performed using paper forms (analog screening, left side). The analysis on the right includes data collected immediately after the implementation of the digital screening, during a transitional phase in which both paper-based and digital methods were used (digital + analog screening, July-December 2024). [a]Number of patients undergoing treatment in the respective period. [b]Some patients were screened through both analog and digital methods during the specified period and therefore were counted twice.



Discussion

Principal Findings

The study demonstrates the successful implementation of a structured, standardized, and ePRO-driven digital screening process for assessing supportive care needs in oncology at LMU University Hospital. The screening tool enables the systematic collection of data in accordance with the screening requirements set by the German Cancer Society (DKG). It is fully integrated into routine clinical workflows via the HIS. A central design principle was the use of a concise questionnaire limited to 14 items, ensuring feasibility and acceptability in everyday clinical practice. Shorter questionnaires have been associated with higher participation rates, reduced respondent burden, and improved inclusivity across diverse patient populations [32-34]. A comparable implementation approach has been described by Maurer et al [35], who similarly integrated ePRO-based supportive care screening into routine clinical workflows, albeit using a more extensive questionnaire. Although the screening pathway was developed within the framework of German Cancer Society certification, the underlying concept is not restricted to the German health care setting. In fact, the concept can be adapted to other national contexts with comparable digital infrastructures and care pathways. Similar ePRO-driven screening strategies have been reported from other countries, demonstrating the feasibility of embedding digital needs assessments into routine cancer care in diverse health system contexts [36,37].

Feasibility, Adherence, and Screening Outcomes

The questionnaire incorporates short-form PRO versions of the Distress Thermometer, the NRS with wording from the Patient-Generated Subjective Global Assessment, and the IPOS. As a prescreening tool, it supports targeted follow-up assessments by the respective specialties. During the first 12 months, 8855 QR codes were generated, and 4909 (55.4%) screenings were completed. According to the literature, the proportion of patients completing standardized distress or supportive care needs screenings typically ranges from 50% to 70%, depending on the clinical setting and the methodology used [38-40]. The adherence observed in our study is therefore comparable to these published rates. However, the true adherence rate may be underestimated, as some QR codes may not have been distributed for various reasons, such as patient absence. In addition, a number of QR codes were generated for training purposes or as part of system testing during the implementation phase.

Across all completed screenings, the identification of unmet needs resulted in SCRs in 22.4% of screenings for

XSL•FO
RenderX

psycho-oncology, 18.7% for nutrition, and 27.6% for palliative care. These findings are broadly consistent with prior studies reporting substantial unmet needs among patients with oncology, although direct comparisons are limited by differences in screening scope and methodology [41,42]. Previous research reported that 23% of patients with cancer were newly referred to palliative care [43]. Our study observed a higher level of need, which can be attributed to the requirement for physician-led final approval of referrals. This finding also reflects the substantial symptom burden in a patient cohort that was not preselected based on medical status.

Equity, Digital Access, and Patient Support

Despite the advantages of digital screening, patient-related barriers may limit participation. Older age, poor health status, lack of access to digital devices, and limited digital literacy can reduce adherence and equity. To mitigate these barriers, paper-based screenings continue to be offered for selected patients. In addition, targeted measures have been introduced to improve digital inclusivity. Additional informational material has been created to explain the purpose and process of the screening, helping patients understand its importance and encouraging participation. Pilot projects are currently underway in several organizational units, where patients facing barriers to digital participation receive active support from trained patient navigators. These navigators play a central role in enabling patients—regardless of age, health status, or digital literacy—to benefit from the digital screening pathway. Institution-provided tablets are used as supportive tools to facilitate access and completion of screenings. Hybrid models combining digital tools with human support have been shown to enhance participation and effectiveness, highlighting the value of patient navigators [44].

Impact on Screening Rates in Routine Care

Screening rates were evaluated on a single outpatient ward before and after the implementation of digital screening. Following implementation, screening rates were maintained or slightly increased (26.9% vs 30%), largely through a combined use of digital and paper-based methods. Maintaining hybrid workflows ensured continuity of care and inclusivity during the transition phase.

Reported screening rates reflect all patients treated during the respective periods and are therefore lower than DKG benchmarks for newly diagnosed primary patients [45,46]. In 2024, the LMU Oncology Center achieved the required 65% threshold for distress screening, to which the digital implementation may have contributed. Distress assessments documented outside standardized reporting structures could not be captured in the analysis.

Facilitators and Barriers to Implementation

Several facilitators supported successful implementation. The DKG requirement for structured screening provided a strong regulatory framework. In addition, implementation was supported by hospital leadership, the IT department, and the technology provider. The allocation of dedicated personnel enabled systematic training, coordination, and continuous refinement of the system. Digitalization also yielded operational benefits. A unified questionnaire simplified workflow, automated data transfer, eliminated manual data entry and document archiving, and ensured the immediate availability of screening results in the HIS. The automated generation of SCRs based on predefined thresholds further reduced administrative burden and improved transparency.

Nevertheless, several barriers were encountered. Generating QR codes initially appeared more time-consuming than distributing paper questionnaires, and the downstream time savings were not always immediately apparent to staff. Comprehensive and repeated training sessions were, therefore, essential, consistent with prior implementation research [47-49]. Moreover, purely digital approaches cannot fully reach all patient groups. Approximately 20% of patients required assistance to complete the questionnaire, highlighting the need for supportive strategies.

Comparison With International Implementations

International experience with digital distress and supportive care screening reveals similar challenges. In the United States, electronic screening is often integrated into patient portals or electronic health records; however, documentation gaps and limited follow-up remain common [36]. In the United Kingdom, digital Holistic Needs Assessments are available and easy to use, but their application in everyday clinical practice is limited. The main problems include a lack of commitment, time constraints within the care team, and insufficient feedback of ePRO results into concrete care decisions, which reduce motivation to use them among both patients and professionals [50]. Australia has technically mature, digitally integrated screening workflows with high initial participation rates. Nevertheless, there are issues with acceptance in subsequent courses, particularly with repeated screenings, as well as uncertainties among staff regarding how to handle high distress values and their implications. Here, too, usage depends heavily on training, local leadership, and clear responsibilities [37,51].

In Asia, digital ePRO approaches vary considerably. In Japan, distress screening is formally established but is often only partially digitized and has limited structured follow-up [52]. In South Korea and China, there are significant barriers to acceptance: low routine use, cultural reluctance to seek psychosocial help, and low acceptance rates after a positive ePRO screening [53,54]. Singapore is an exception, as digital distress ePROs are closely linked to a multidisciplinary care model and thus achieve a high level of acceptance [55]. In Europe, the Netherlands exemplifies successful implementation when digital screening is mandatory, systematically discussed, and directly linked to care pathways [56]. Overall, international evidence suggests that technical availability alone is insufficient; sustained use depends on clear workflows, automated follow-up, leadership support, and perceived clinical relevance [57].

Generalizability and Transferability

Although conducted at a single CCC, several features support transferability. The screening content is based on internationally validated instruments and guideline-aligned supportive care domains. Technically, the CANKADO platform is HIS-agnostic. Interfaces are implemented via standardized web-based

XSL•FO
RenderX

application programming interfaces and enable integration with heterogeneous hospital IT environments beyond the SAP-based infrastructure used at LMU Hospital. Structured data exchange enables flexible mapping to local documentation and reporting requirements.

The system complies with the General Data Protection Regulation and the more stringent French Hébergeurs de Données de Santé framework, supporting international applicability. While the depth of integration achieved at LMU Hospital has not yet been replicated elsewhere, the platform is already used in multiple international settings [19,30,58,59]. Future multicenter studies are needed to evaluate scalability, resource requirements, and contextual adaptations across different health care systems.

Limitations

Despite the demonstrated feasibility and advantages of digital supportive care screening in oncology, several limitations should be acknowledged. Technical challenges, such as temporary connectivity issues, system maintenance, or software-related disruptions, may intermittently affect access to the digital platform and result in delayed or incomplete data capture. In addition, the implementation and maintenance of a secure, interoperable digital screening infrastructure require substantial technical, administrative, and human resources. Ensuring ongoing compliance with data protection regulations and IT security standards represents a continuous organizational effort.

A key limitation relates to equity and accessibility. Digital screening inherently excludes certain patient groups, particularly individuals without access to suitable digital devices or with limited digital literacy. Older patients, those with cognitive or physical impairments, and patients with advanced disease or severe symptom burdens may face additional barriers to independent participation. Although supportive measures, such as paper-based alternatives and assisted completion, were available, these barriers may have reduced the inclusivity and representativeness of the screening results.

Human resource constraints also represent a relevant limitation. High workload among HCPs, especially in nursing and medical staff, may delay the initiation of screening or the review and approval of triggered SCRs. In particular, delays in physician review of SCRs may hinder timely referral to counseling services and continuity of care. At present, systematic evaluation of approved SCRs, completed counseling sessions, or canceled referrals is not yet possible within the existing data structure; however, this functionality is currently under development and will be addressed in future analyses.

Furthermore, the digital questionnaire currently does not adequately address linguistic and cultural diversity. The lack of multilingual versions may limit accessibility for patients with limited proficiency in the primary language of care and may reduce the effectiveness of the screening in culturally diverse populations. While digital implementation allows relatively straightforward adaptation to additional languages, this feature has not yet been fully implemented at the time of evaluation.

Finally, the study was conducted at a single academic CCC with a high degree of digital maturity and IT integration. Although the technical architecture and conceptual approach are transferable, implementation depth, available resources, and organizational workflows may differ substantially across institutions. These contextual factors should be considered when interpreting the findings and underscore the need for future multicenter evaluations.

Implications for Clinical Practice

The findings of this study provide several practical implications for the implementation of digital supportive care screening in routine oncology practice. First, successful implementation requires careful preparation, clear governance structures, and the allocation of sufficient technical and human resources. Defining the scope of supportive care domains and selecting a concise questionnaire that balances clinical relevance with feasibility are essential to facilitate integration into daily workflows.

Comprehensive and role-specific training programs for HCPs are critical. Initial training should be complemented by regular refresher sessions to maintain engagement, address uncertainties, and integrate system updates. Assigning clear responsibilities to specific professional groups—for example, nursing staff for screening initiation and physicians for SCR approval—supports accountability and process reliability. Introducing the screening initially in motivated units or pilot wards may foster early adoption and generate positive momentum for broader rollout.

Maintaining flexibility in screening modalities is particularly important during implementation phases. Hybrid workflows that combine digital and paper-based screening can ensure continuity of care and inclusivity, especially for patients facing access barriers. From a technical perspective, screening data should be available both as archivable PDF documents and as structured data within the HIS to support documentation, quality assurance, and secondary analyses. Systems should also be adaptable to future modifications of questionnaire content or screening frequency.

Equally important is the proactive consideration of equity and accessibility. Implementation strategies should explicitly address patients with limited digital skills, physical impairments, or language barriers. Support structures such as patient navigators, assisted completion, and the provision of institution-owned devices can substantially enhance digital participation. Finally, digital screening will only be sustainably adopted if both patients and HCPs perceive clear clinical benefits. Automated follow-up processes, timely referrals, and visible improvements in patient care are, therefore, essential to reinforce acceptance and long-term use.

Conclusion

A digital oncological screening in the supportive care setting was successfully implemented at a German CCC. Automated integration into the HIS enabled efficient identification of supportive care needs, reduced administrative burden, and supported patient-centered care. Continuous training and standardized procedures were critical to success. Future initiatives will focus on improving digital equity and accessibility to ensure that all patient groups benefit from supportive oncology care.

XSL•FO
RenderX

Acknowledgments

We thank Gabriele Kraus-Pfeiffer and Oliver Guldner for their valuable support during the execution of the pilot phase. AI-assisted tools were utilized to support the literature search process and to assist with language editing, translation, and manuscript reformatting. Tools used included DeepL and ChatGPT. All AI-generated content was critically reviewed, revised, and approved by the authors. All authors take full responsibility for the accuracy and integrity of the final manuscript.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Authors' Contributions

TF and NE were primarily responsible for the study design, including conceptualization, methodology, and supervision. VH, FM, BH, RW, and ST supervised the clinical components and contributed to the overall study concept. NE, FM, and BH developed the questionnaire. TS managed the required modifications from CANKADO's side. GW implemented the necessary Information Technology programming at the hospital. UH, SM, and NT conducted the training sessions and collected the data. TF and NT performed data analysis and interpretation. TF drafted the original manuscript, which was then edited by NE. All authors reviewed and approved the final version.

Conflicts of Interest

NE has received speaker honoraria from Audi BKK, Baxter, Beiersdorf, CSL Behring, Fresenius, Freiraum, Fraunhofer Institut, Janssen-Cilag GmbH, Cognitando GmbH, and GHD. She served on the expert advisory board at Baxter and received consulting fees from Fresenius. None of these activities were related to the content of this paper. RW received honoraria for consultancy, advisory boards, and travel support from Agendia, Amgen, APOGHEVA, Aristo, Astra Zeneca, Bayer, Celgene, Clinsol, Daiichi-Sankyo, Eisai, Esteve, Exact Sciences, Gilead, Hexal/Sandoz, Lilly, MSD, Mundipharma, Mylan, Nanostring, Novartis, Onkowissen, Paxman, Palleos, Pfizer, Pierre Fabre, PINK, Riemser, Roche, Seagen, Sidekick, Stemline, Tesaro Bio, Teva, Veracyte, Viatris, Wiley, Thieme, Digimed, Eickeler, FOMF, Art Tempi, Aurikamed, Clinsol, Pomme Med, iMEDInstitut, medconcept, MCI, MediSeminar, Medicultus, H+O Communications, New Oncology Concepts, SPCC, streamed up, medupdate GmbH, and institutions: SPCC, WSG, AGO, and DKG.

References

1. Jordan K, Aapro M, Kaasa S, et al. European Society for Medical Oncology (ESMO) position paper on supportive and palliative care. Ann Oncol 2018 Jan 1;29(1):36-43. [doi: 10.1093/annonc/mdx757] [Medline: 29253069]
2. Berman R, Davies A, Cooksley T, et al. Supportive care: an indispensable component of modern oncology. Clin Oncol (R Coll Radiol) 2020 Nov;32(11):781-788. [doi: 10.1016/j.clon.2020.07.020] [Medline: 32814649]
3. Richardson A, Medina J, Brown V, Sitzia J. Patients' needs assessment in cancer care: a review of assessment tools. Support Care Cancer 2007 Oct;15(10):1125-1144. [doi: 10.1007/s00520-006-0205-8] [Medline: 17235503]
4. Stengel A, Dinkel A, Karger A, et al. Best practice: psychoonkologisches Screening an Comprehensive Cancer Centers [Article in German]. Forum 2021 Sep;36(4):278-283. [doi: 10.1007/s12312-021-00944-x]
5. Braulke F, Para S, Alt-Epping B, et al. Systematic symptom screening in patients with advanced cancer treated in certified oncology centers: results of the prospective multicenter German KeSBa project. J Cancer Res Clin Oncol 2023 Sep;149(11):8829-8842. [doi: 10.1007/s00432-023-04818-8] [Medline: 37145199]
6. van Oorschot B. Einführung zur Best-Practice-Reihe: multiprofessionelles supportivscreening in der routine mittels PROM [Article in German]. Die Onkologie 2024 Aug;30(8):720-720. [doi: 10.1007/s00761-024-01558-3]
7. Leitlinienprogramm onkologie: psychoonkologische diagnostik, beratung und behandlung von erwachsenen krebspatient*innen [Report in German]. : Deutsche Krebsgesellschaft, Deutsche Krebshilfe, (AWMF); 2023 URL: https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/Psychoonkologie/Version_2/LL_Psychoonkologie_Langversion_2.1.pdf [accessed 2026-03-06]
8. Erweiterte S3-Leitlinie Palliativmedizin für Patienten mit einer nicht-heilbaren Krebserkrankung [Website in German]. : Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF; 2020 URL: https://register.awmf.org/assets/guidelines/128-001OLl_S3_Palliativmedizin_2025-03-abgelaufen.pdf [accessed 2026-02-26]

XSL•FO
RenderX

9. SOP Onkologisches Basisscreening [Website in German]. : Deutsche Krebsgesellschaft (DKG); 2024 URL: https://www.onkozert.de/wordpress/wp-content/uploads/2024/03/SOP%20Onkologisches%20Basisscreening_240312.pdf [accessed 2026-02-26]
10. Salmani H, Nasiri S, Alemrajabi M, Ahmadi M. Advancing patient-centered cancer care: a systematic review of electronic patient-reported outcome measures. Front Rehabil Sci 2024;5:1427712. [doi: 10.3389/fresc.2024.1427712] [Medline: 39386354]
11. Elkefi S, Trapani D, Ryan S. The role of digital health in supporting cancer patients' mental health and psychological well-being for a better quality of life: a systematic literature review. Int J Med Inform 2023 Aug;176:105065. [doi: 10.1016/j.ijmedinf.2023.105065] [Medline: 37224644]
12. Hopstaken JS, Verweij L, van Laarhoven CJHM, Blijlevens NMA, Stommel MWJ, Hermens RPMG. Effect of digital care platforms on quality of care for oncological patients and barriers and facilitators for their implementation: systematic review. J Med Internet Res 2021 Sep 24;23(9):e28869. [doi: 10.2196/28869] [Medline: 34559057]
13. Marthick M, McGregor D, Alison J, Cheema B, Dhillon H, Shaw T. Supportive care interventions for people with cancer assisted by digital technology: systematic review. J Med Internet Res 2021 Oct 29;23(10):e24722. [doi: 10.2196/24722] [Medline: 34714246]
14. Husain N, Ansari Z, Shamim MA, et al. Electronic Patient Reported Outcome Measures and quality of life in cancer (E-PROMISE): systematic review of the evidence and meta-analysis. BMJ Open Qual 2025 Apr 28;14(2):e003209. [doi: 10.1136/bmjoq-2024-003209] [Medline: 40294959]
15. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Adm Policy Ment Health 2011 Jan;38(1):4-23. [doi: 10.1007/s10488-010-0327-7] [Medline: 21197565]
16. Moullin JC, Dickson KS, Stadnick NA, Rabin B, Aarons GA. Systematic review of the Exploration, Preparation, Implementation, Sustainment (EPIS) framework. Implement Sci 2019 Jan 5;14(1):1. [doi: 10.1186/s13012-018-0842-6] [Medline: 30611302]
17. Erickson N, Schinkoethe T, Eckhardt C, et al. Patient-reported outcome measures obtained via e-health tools ease the assessment burden and encourage patient participation in cancer care (PaCC Study). Support Care Cancer 2021 Dec;29(12):7715-7724. [doi: 10.1007/s00520-021-06351-1] [Medline: 34159428]
18. Fey T, Erickson N, Stahler A, et al. Should routine risk reduction procedures for the prevention and control of pandemics become a standard in all oncological outpatient clinics? The prospective COVID-19 cohort study: protect-CoV. Med Oncol 2022 Apr 10;39(6):104. [doi: 10.1007/s12032-022-01700-4] [Medline: 35397689]
19. Harbeck N, Fasching PA, Wuerstlein R, et al. Significantly longer time to deterioration of quality of life due to CANKADO PRO-React eHealth support in HR+ HER2- metastatic breast cancer patients receiving palbociclib and endocrine therapy: primary outcome analysis of the multicenter randomized AGO-B WSG PreCycle trial. Ann Oncol 2023 Aug;34(8):660-669. [doi: 10.1016/j.annonc.2023.05.003] [Medline: 37201751]
20. Degenhardt T, Fasching PA, Lüftner D, et al. PRECYCLE: multicenter, randomized phase IV intergroup trial to evaluate the impact of eHealth-based patient-reported outcome (PRO) assessment on quality of life in patients with hormone receptor positive, HER2 negative locally advanced or metastatic breast cancer treated with palbociclib and an aromatase inhibitor or palbociclib and fulvestrant. Trials 2023 May 17;24(1):338. [doi: 10.1186/s13063-023-07306-z] [Medline: 37198674]
21. Schinköthe T, Gabri MR, Mitterer M, et al. A web- and app-based connected care solution for COVID-19 in- and outpatient care: qualitative study and application development. JMIR Public Health Surveill 2020 Jun 1;6(2):e19033. [doi: 10.2196/19033] [Medline: 32406855]
22. About us. Cankado.ai. URL: https://cankado.ai/about [accessed 2026-02-26]
23. Erickson N, Storck LJ, Kolm A, et al. Tri-country translation, cultural adaptation, and validity confirmation of the Scored Patient-Generated Subjective Global Assessment. Support Care Cancer 2019 Sep;27(9):3499-3507. [doi: 10.1007/s00520-019-4637-3] [Medline: 30684046]
24. Riba MB, Donovan KA, Andersen B, et al. Distress management, version 3.2019, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw 2019 Oct 1;17(10):1229-1249. [doi: 10.6004/jnccn.2019.0048] [Medline: 31590149]
25. IPOS-5 in English. Palliative care Outcome Scale (POS). URL: https://pos-pal.org/maix/ipos-5-in-english.php [accessed 2026-02-26]
26. Murtagh FE, Ramsenthaler C, Firth A, et al. A brief, patient- and proxy-reported outcome measure in advanced illness: validity, reliability and responsiveness of the Integrated Palliative care Outcome Scale (IPOS). Palliat Med 2019 Sep;33(8):1045-1057. [doi: 10.1177/0269216319854264] [Medline: 31185804]
27. Müller E, Gahr S, Schnabel A, et al. Screening auf Bedarf an spezialisierter Palliativversorgung [Article in German]. Die Onkologie 2024 Sep;30(9):850-856. [doi: 10.1007/s00761-024-01574-3]
28. Müller E, Müller MJ, Sölder P, Rautenberg B, En-Nosse M, Becker G. Development and validation of Integrated Palliative Care Outcome Scale Short Form for efficient specialist palliative care screening in oncology (ScreeningPALL): a cross-sectional study. J Palliat Med 2025 Apr;28(4):453-460. [doi: 10.1089/jpm.2024.0343] [Medline: 39929478]
29. Gluz O, Nitz UA, Christgen M, et al. Efficacy of endocrine therapy plus trastuzumab and pertuzumab vs de-escalated chemotherapy in patients with hormone receptor-positive/ERBB2-positive early breast cancer: the neoadjuvant WSG-TP-II randomized clinical trial. JAMA Oncol 2023 Jul 1;9(7):946-954. [doi: 10.1001/jamaoncol.2023.0646] [Medline: 37166817]

XSL•FO
RenderX

30. Michel LL, Schwarz D, Romar P, et al. Efficacy of hand cooling and compression in preventing taxane-induced neuropathy: the POLAR randomized clinical trial. JAMA Oncol 2025 Apr 1;11(4):408-415. [doi: 10.1001/jamaoncol.2025.0001] [Medline: 40048176]
31. Harbeck N, Wrobel D, Zaiss M, et al. Neratinib as extended adjuvant treatment of HER2-positive/HR-positive early breast cancer patients in Germany, Austria, and Switzerland: interim results of the prospective, observational ELEANOR study. Breast Care (Basel) 2024 Feb;19(1):1-9. [doi: 10.1159/000533657] [Medline: 38384488]
32. Kost RG, de Rosa JC. Impact of survey length and compensation on validity, reliability, and sample characteristics for ultrashort-, short-, and long-research participant perception surveys. J Clin Transl Sci 2018 Feb;2(1):31-37. [doi: 10.1017/cts.2018.18] [Medline: 30393572]
33. Sharma H. How short or long should be a questionnaire for any research? Researchers dilemma in deciding the appropriate questionnaire length. Saudi J Anaesth 2022;16(1):65-68. [doi: 10.4103/sja.sja_163_21] [Medline: 35261591]
34. Hamilton CB, Hoens AM, McKinnon AM, et al. Shortening and validation of the Patient Engagement In Research Scale (PEIRS) for measuring meaningful patient and family caregiver engagement. Health Expect 2021 Jun;24(3):863-879. [doi: 10.1111/hex.13227] [Medline: 33729634]
35. Maurer J, Saibold A, Gerl K, et al. Systematic development of a patient-reported ONCOlogical-ROUTinE-Screening (ONCO-ROUTES) procedure at the University Cancer Center Regensburg. J Cancer Res Clin Oncol 2024 Sep 28;150(9):435. [doi: 10.1007/s00432-024-05955-4] [Medline: 39340547]
36. Garcia SF, Wortman K, Cella D, et al. Implementing electronic health record-integrated screening of patient-reported symptoms and supportive care needs in a comprehensive cancer center. Cancer 2019 Nov 15;125(22):4059-4068. [doi: 10.1002/cncr.32172] [Medline: 31373682]
37. Mazariego C, Williamson K, Leon S, et al. Statewide integration of electronic patient-reported outcome measures into routine oncology care: a mixed-methods implementation study. EClinicalMedicine 2025 Nov;89:103547. [doi: 10.1016/j.eclinm.2025.103547] [Medline: 41140450]
38. Riedl D, Gastl R, Gamper E, et al. Cancer patients' wish for psychological support during outpatient radiation therapy: findings from a psychooncological monitoring program in clinical routine. Strahlenther Onkol 2018 Jul;194(7):655-663. [doi: 10.1007/s00066-018-1288-0] [Medline: 29532098]
39. Carlson LE, Waller A, Groff SL, Zhong L, Bultz BD. Online screening for distress, the 6th vital sign, in newly diagnosed oncology outpatients: randomised controlled trial of computerised vs personalised triage. Br J Cancer 2012 Aug 7;107(4):617-625. [doi: 10.1038/bjc.2012.309] [Medline: 22828610]
40. van Oorschot B, Zechel S, Roch C, Jentschke E. Erste Erfahrungen mit einem Multisymptomscreening auf Unterstützungsbedarf bei Patienten mit metastasierten Krebserkrankungen [Article in German]. Die Onkologie 2023 Jun;29(6):539-547. [doi: 10.1007/s00761-023-01358-1]
41. Papageorgiou L, Le Provost JB, Di Palma M, et al. Supportive care needs of newly diagnosed cancer patients in a comprehensive cancer center: identifying care profiles and future perspectives. Cancers (Basel) 2024 Feb 29;16(5):1017. [doi: 10.3390/cancers16051017] [Medline: 38473376]
42. Springer F, Mehnert-Theuerkauf A, Gebhardt C, Stolzenburg JU, Briest S. Unmet supportive care needs among cancer patients: exploring cancer entity-specific needs and associated factors. J Cancer Res Clin Oncol 2024 Apr 12;150(4):190. [doi: 10.1007/s00432-024-05715-4] [Medline: 38607426]
43. Ostgathe C, Wendt KN, Heckel M, et al. Identifying the need for specialized palliative care in adult cancer patients - development and validation of a screening procedure based on proxy assessment by physicians and filter questions. BMC Cancer 2019 Jul 1;19(1):646. [doi: 10.1186/s12885-019-5809-8] [Medline: 31262285]
44. Kim S, Whipkey AJ, Bae J, Bhakta A. Digital health interventions for colorectal cancer screening uptake: a scoping review. PLoS Digit Health 2025 Sep;4(9):e0001028. [doi: 10.1371/journal.pdig.0001028] [Medline: 40996970]
45. Hintergrund und FAQ zur Kennzahl "Psychoonkologisches Distress-Screening" [Report in German]. : Deutsche Krebsgesellschaft (DKG) / OnkoZert; 2023 URL: https://www.onkozert.de/wordpress/wp-content/uploads/2024/01/FAQ%20Distress%20Screening%20231208.pdf?v=15617818 [accessed 2026-02-26]
46. Erhebungsbögen und dokumente [Article in German]. German Cancer Society (Deutsche Krebsgesellschaft). URL: https://www.krebsgesellschaft.de/deutsche-krebsgesellschaft/zertifizierung/erhebungsboegen.html [accessed 2026-03-06]
47. Ahmed J, Mehmood S, Rehman S, Ilyas C, Khan LUR. Impact of a structured template and staff training on compliance and quality of clinical handover. Int J Surg 2012;10(9):571-574. [doi: 10.1016/j.ijsu.2012.09.001] [Medline: 22983018]
48. Damery S, Flanagan S, Jones J, Jolly K. The effect of providing staff training and enhanced support to care homes on care processes, safety climate and avoidable harms: evaluation of a care home quality improvement programme in England. Int J Environ Res Public Health 2021 Jul 16;18(14):7581. [doi: 10.3390/ijerph18147581] [Medline: 34300034]
49. Cummings BM, Carballo V, Kalibatas O, et al. Impact of clinical process improvement training in an integrated delivery system. Am J Med Qual 2021;36(3):156-162. [doi: 10.1177/1062860620943960] [Medline: 32734765]
50. Islam MS. Treat patient, not just the disease: holistic needs assessment for haematological cancer patients. Oncol Rev 2018 Jul 4;12(2):374. [doi: 10.4081/oncol.2018.374] [Medline: 30283608]

XSL•FO
RenderX

51. Koczwara B, Knowles R, Beatty L, et al. Implementing a web-based system of screening for symptoms and needs using patient-reported outcomes in people with cancer. Support Care Cancer 2022 Dec 21;31(1):69. [doi: 10.1007/s00520-022-07547-9] [Medline: 36542190]
52. Okuyama T, Kizawa Y, Morita T, et al. Current status of distress screening in designated cancer hospitals: a cross-sectional nationwide survey in Japan. J Natl Compr Canc Netw 2016 Sep;14(9):1098-1104. [doi: 10.6004/jnccn.2016.0119] [Medline: 27587622]
53. Hahn C, Joo SH, Chae JH, Lee CU, Kim TS. Feasibility of psychosocial distress screening and management program for hospitalized cancer patients. Psychiatry Investig 2017 Nov;14(6):734-745. [doi: 10.4306/pi.2017.14.6.734] [Medline: 29209376]
54. Tang L, Zhang Y, Pang Y. Patient-reported outcomes from the distress assessment and response tool program in Chinese cancer inpatients. Psychooncology 2020 May;29(5):869-877. [doi: 10.1002/pon.5358] [Medline: 32040238]
55. Ke Y, Neo PSH, Yang GM, et al. Impact of a multidisciplinary supportive care model using distress screening at an Asian ambulatory cancer center: a cluster randomized controlled trial. JCO Oncol Pract 2024 Sep;20(9):1207-1218. [doi: 10.1200/OP.23.00505] [Medline: 38838278]
56. van Nuenen FM, Donofrio SM, van de Wiel HBM, Hoekstra-Weebers JEHM. Cancer patients' experiences with and opinions on the process "Screening of Distress and Referral Need" (SDRN) in clinical practice: a quantitative observational clinical study. PLoS One 2018;13(6):e0198722. [doi: 10.1371/journal.pone.0198722] [Medline: 29902202]
57. Aanes SG, Wiig S, Nieder C, Haukland EC. Implementing digital patient-reported outcomes in routine cancer care: barriers and facilitators. ESMO Real World Data Digit Oncol 2024 Dec;6:100088. [doi: 10.1016/j.esmorw.2024.100088] [Medline: 41646091]
58. Hirschberg L, Henze F, Paradies K, et al. Evaluation of therapy support through a standardized nursing consultation for patients undergoing oral tumor therapy in gynecological oncology within the prospective CAMPA initiative. Eur J Oncol Nurs 2025 Feb;74:102770. [doi: 10.1016/j.ejon.2024.102770] [Medline: 39799641]
59. Kaidar-Person O, Antunes M, Cardoso JS, et al. Evaluating the ability of an artificial-intelligence cloud-based platform designed to provide information prior to locoregional therapy for breast cancer in improving patient's satisfaction with therapy: the CINDERELLA trial. PLoS One 2023;18(8):e0289365. [doi: 10.1371/journal.pone.0289365] [Medline: 37535564]

Abbreviations

CCC: Comprehensive Cancer Center
EPIS: Exploration, Preparation, Implementation, Sustainment
ePRO: electronic patient-reported outcome
HCP: health care professional
HIS: hospital information system
IPOS: Integrated Palliative Care Outcome Scale
NRS: Nutrition Risk Screening
PRO : patient-reported outcome
SCR: supportive care request

Edited by M Balcarras; submitted 14.Aug.2025; peer-reviewed by GLP Bodagala, H Anandan, K Yoshiuchi; revised version received 15.Jan.2026; accepted 19.Jan.2026; published 19.Mar.2026.

Please cite as:
Fey T, Thurner N, Haidn U, Mumm F, Haberland B, Wuerstlein R, Theurich S, Wolfrum G, Troidl ML, Müller S, Bausewein C, Schinköthe T, Heinemann V, Erickson N
Scan, Screen, Support—A Digital Pathway for Assessing Supportive Care Needs in Oncology: Implementation Study
JMIR Cancer 2026;12:e82392
URL: https://cancer.jmir.org/2026/1/e82392
doi:10.2196/82392

© Theres Fey, Nicole Thurner, Ulrike Haidn, Friederike Mumm, Birgit Haberland, Rachel Wuerstlein, Sebastian Theurich, Georg Wolfrum, Marie-Louise Troidl, Susan Müller, Claudia Bausewein, Timo Schinköthe, Volker Heinemann, Nicole Erickson. Originally published in JMIR Cancer (https://cancer.jmir.org), 19.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient Perspectives on AI-Powered Medical Robots in Breast and Prostate Cancer Care: Qualitative Study

Mahiya Habib[1*], BA; Janet Ellis[2,3*], MBBCHIR, MD, FRCPC; Aaron Palachi[1,4], BA, MA; Melissa B Korman[1,3], BSc, MSc, PhD; Tatjana Kay[4], BA, MA; Karen Barlow[5,6], BSc; Jordana DeSouza[1], BA; Rosanna Macri[7,8,9], MHSc, Q Med; Abdullah Alabousi[10], MD; Mehran Anvari[5,6,11], MBBS, FACS, FRCSC, PhD

[1]Department of Clinical Evaluative Sciences, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

[10]Department of Radiology, McMaster University, St. Joseph's Healthcare Hamilton, Hamilton, ON, Canada

[11]Department of Surgery, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

[2]Department of Psychiatry, Sunnybrook Health Science Centre, 2075 Bayview Ave, Toronto, ON, Canada

[3]Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[4]Department of Psychology, Toronto Metropolitan University, Toronto, ON, Canada

[5]Centre for Surgical Invention and Innovation, Hamilton, ON, Canada

[6]Centre for Minimal Access Surgery, St. Joseph's Healthcare, McMaster University, Hamilton, ON, Canada

[7]Department of Bioethics, Sinai Health, Toronto, ON, Canada

[8]Joint Centre for Bioethics, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[9]Department of Radiation Oncology, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Janet Ellis, MBBCHIR, MD, FRCPC
Department of Psychiatry, Sunnybrook Health Science Centre, 2075 Bayview Ave, Toronto, ON, Canada

Abstract

Background: Nearly 250,000 cancer cases are diagnosed annually in Canada, with breast and prostate cancer representing 25% and 22% of new cases, respectively. Artificial intelligence (AI) applications can potentially improve the accuracy, efficiency, and timeliness of cancer care, including screening, diagnostic imaging, and early treatment. However, patient acceptability of AI in cancer care remains underexplored.

Objective: This study aimed to understand the feelings and perceptions of and acceptability to patients with breast and prostate cancer regarding the inclusion of AI-powered medical robots for cancer screening, diagnosis, and early treatment and to identify barriers and facilitators to implementation.

Methods: In this qualitative study, semistructured interviews were conducted with 15 patients with breast (n=6) or prostate (n=9) cancer. Participants (mean [SD] age 67 [12] y; range 41 - 88 y) were recruited from Sunnybrook Health Sciences Centre between May and November 2022. Each participant completed 2 semistructured interviews, each approximately 1 hour in length, conducted by telephone or Zoom by a research assistant. The first interview explored initial reactions and expectations regarding AI-assisted medical technologies, allowing us to tailor questions for the second interview to better understand practical means of introducing AI into care, while further exploring salient concepts. Data were analyzed using a conventional content analysis approach. Two research assistants independently and inductively coded transcripts, iteratively generating and refining a codebook. Data saturation was assessed after 10 interviews and confirmed through analysis of all 15 participants.

Results: Three overarching categories were developed: (1) individual beliefs, understanding, and attitudes; (2) integration of AI into care; and (3) health structure, systems, and processes. Participants reported openness toward AI-assisted medical robots but emphasized the importance of reduced wait times, demonstrated safety and reliability, and patient-centered care. Patients indicated that with appropriate education and transparent communication, they would be willing to accept AI-assisted care due to its enhanced accuracy and efficiency. Key barriers included concerns about reliability, potential loss of human interaction, and inadequate mitigation strategies for technical failures. Facilitators included timely results, improved efficiency, accessible educational resources, and assurance that AI would complement rather than replace human expertise.

Conclusions: Participants demonstrated cautious willingness to accept AI-powered medical robots in cancer care if positioned as complementary to, rather than substitutes for, human-provided care. These findings underscore the need for maintaining the presence of health care providers during AI-assisted procedures, providing clear and accessible education, and ensuring transparent communication about safety and reliability. Personalizing patient education and offering multiple modes of information delivery may foster confidence and improve acceptability. While findings are exploratory and reflect the perspectives of a small,

XSL•FO
RenderX

predominantly urban sample, they provide actionable insights into patient concerns and priorities that may inform future research and guide early implementation strategies in integrating AI into cancer care pathways.

(JMIR Cancer 2026;12:e69710) doi:10.2196/69710

KEYWORDS

artificial intelligence; breast and prostate cancer; cancer care; patient perspectives; patient-centered care; qualitative research; acceptability

Introduction

By the end of 2025, an estimated 248,700 new cancer cases will be diagnosed in Canada. Breast and prostate cancers are expected to account for approximately 23% of all new cancer cases [1]. For patients, receiving cancer care can be a prolonged and complicated process from screening to diagnosis and treatment [2,3]. In this article, the term *screening* is used broadly, consistent with the terminology adopted in North American cancer care research, to encompass both asymptomatic individuals undergoing high-risk surveillance (eg, breast magnetic resonance imaging (MRI) or elevated prostate-specific antigen testing) and those in the early diagnostic phase following symptom presentation or initial cancer detection. This broader operational definition reflects the continuum of early detection and diagnostic processes experienced by patients with breast or prostate cancer in the Canadian health care context. Despite the high prevalence of breast and prostate cancers, patients continue to face long diagnostic and treatment intervals, limited access to specialists, and systemic delays that negatively impact outcomes and quality of life [2-7]. Weller et al [4] refer to 3 stages of cancer care: the patient interval (symptom presentation to seeking medical attention), the diagnostic interval (seeking medical attention to obtaining a diagnosis), and the treatment interval (diagnosis to treatment) [5]. The diagnostic and treatment intervals usually include medical tests and multiple appointments with health care providers across several medical offices or hospitals [5]. Typical care pathways for prostate cancer (spanning the patient, diagnostic, and treatment intervals described by Weller et al [4]) are estimated to take between 163 and 367 days [3]. For breast cancer, the continuum from the patient interval to the completion of the treatment interval can take approximately 60 days [2]. The series of multiple diagnostic and treatment procedures contributes to long wait times and fragmented care. Such delays negatively affect quality of life, increase uncertainty, and may lead to worse health outcomes [6,7].

Timely medical attention is critical for optimal patient care, but there is a shortage of health care providers in the current Canadian health care system (ie, imaging specialists, specialized oncologists, and surgeons), which adds difficulties in meeting increasing patient demand [8,9]. Advanced-stage breast and prostate cancers may require more invasive interventions, including mastectomies and prostatectomies [10,11]. Early detection and diagnosis may allow for more targeted and less invasive treatment regimens [10,12], whereas when cancer is detected too late, less invasive options may not be possible. Across oncology and other medical specialties, health care providers experience growing workload pressures due to increasing patient volumes, staffing shortages, and other systemic inefficiencies. This has contributed to nationally elevated burnout rates and compassion fatigue among health care providers [8,9]. The limited availability and overworking of health care providers contribute to accessibility issues for patients, especially in rural and remote regions, where accessibility and availability of specialized health resources are often even more sparse [13,14].

In a systematic review, Reece et al [7] identified several key issues regarding delayed or failed follow-up for breast cancer screening in primary care settings: physician-patient miscommunication, automated alert systems creating an overinflux of information, reduced coordination of patient medical health records, and inconvenient clinic hours and lack of availability of primary care. Urquhart et al [15] interviewed survivors of prostate cancer about their experiences with follow-up and posttreatment cancer care. Participants described various issues, including lack of information, resources, and psychosocial support, which left them feeling unprepared in their recovery process. These perceived issues in their care were related to lack of availability of health care providers and not being able to see the same health care provider throughout their care due to a staffing shortage within oncology [8,15].

Recent advances in technology, such as the use of artificial intelligence (AI) in cancer screening, represent a potential mechanism for increasing efficiency in wait times in cancer pathways, reducing the demand on health care providers, and improving accessibility within the health care system [16,17]. AI has been shown to be an effective tool for breast and prostate cancer diagnosis and in prostate cancer Gleason scoring [18]. Successful AI applications in diagnostic imaging have enabled the detection of early cancerous lesions with greater accuracy and efficiency, with trialed technology often outperforming experts [19,20]. A 2017 study showed that automated deep learning algorithms identified nodal metastases in histopathological analyses of breast tissue more accurately than an expert panel of pathologists [21]. AI deep learning algorithms have also assessed mammograms with proficiency equal to expert radiologists, with 5.7% and 9.4% reduction in false positive and false negative rates, respectively [22]. In prostate cancer screening, deep learning algorithms have demonstrated the potential to automate Gleason scoring of histopathologic images of adenocarcinomas, achieving 75% agreement with expert pathologists, demonstrating great model performance [18]. Similar AI-powered applications have been used with diffusion-weighted MRI to delineate cancerous and noncancerous prostate tissue [23]. As AI becomes more sophisticated, new opportunities in both diagnostic imaging and treatments are emerging with potential to transcend the accuracy

XSL•FO
RenderX

and time limitations of traditional methods. Specifically, with the advent of AI and robotics, telemedicine applications are feasible and have the potential to mobilize high-quality health care, allowing interventions to be carried out by medical robots, without a highly trained specialist onsite [18,23].

While AI-powered robotic systems have the potential to expedite procedures and improve patient care, the success of these technologies is contingent on patient acceptance. The needs and preferences of patients should influence both the development and the process of implementation of such technology into the health care system. This requires a comprehensive understanding of the patients' perceptions toward the use of such technology in their care. A greater understanding of what may be deemed acceptable is necessary before conducting large-scale research evaluating the use of AI in breast and prostate cancer screening and treatment [24,25]. The present study sought to elucidate patient perceptions, feelings, and acceptability regarding AI-powered medical robots in breast and prostate cancer care and to identify barriers and facilitators to implementation using a person-centered approach.

Methods

Study Design

This study employed a qualitative descriptive approach to answer the research question: How do patients feel about the use of AI-powered medical robots for cancer screening, diagnosis, and early treatment? Study objectives were to (1) illuminate patient understandings and feelings regarding the inclusion of AI in their care and (2) identify barriers and facilitators of implementing AI-powered robotic systems into cancer screening from patients' perspectives. This study is reported in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines (Checklist 1) [26]. A postpositivist stance [27] was taken, which assumes that while absolute truth cannot be fully captured, systematic qualitative methods can approximate patient perspectives and generate transferable insights. This orientation informed our design by emphasizing structured data collection (semistructured interviews), analytic rigor (systematic coding by multiple analysts), and reflexivity to minimize bias. This study was part of a larger program of work initiated by a multidisciplinary research team developing and testing AI technology for cancer care. Although not embedded within a larger trial, it was designed to inform future research protocols and ethical considerations for implementing AI-powered medical robots. The overarching program sought to explore the feasibility of systems capable of scanning, diagnosing, and treating patients within a single appointment. This qualitative study, therefore, focused on understanding patient comfort and acceptability of such technologies across the cancer care continuum, including perspectives on initiating treatment immediately following diagnosis.

Participants

Participants were recruited from breast and prostate cancer clinics at the Odette Cancer Centre at Sunnybrook Health Sciences Centre between May 2022 and November 2022. Purposive sampling was used to identify individuals at elevated risk of breast or prostate cancer, or with a recent diagnosis, to ensure that perspectives from both patient groups were represented. Potential participants were first identified by their treating physicians, who requested permission for research staff to contact them. Research staff approached participants by telephone following referral, explained the purpose of the study, and assessed their eligibility and interest. Inclusion criteria were (1) English speaking, (2) aged 18+ years, and (3) patients deemed *high risk* by their physician and in need of MRI as part of their screening process for breast cancer or diagnosed with breast cancer *or* male patients with elevated prostate-specific antigen levels and considered *high risk* for prostate cancer or diagnosed with prostate cancer.

Ethical Considerations

This study was reviewed and approved by the Research Ethics Board at Sunnybrook Health Sciences Centre (Research Ethics Board 5361). All participants provided verbal informed consent prior to participation, which was documented by the study team in accordance with institutional guidelines. To protect privacy and confidentiality, interview transcripts were deidentified during the transcription. All data were stored securely on password-protected institutional servers with access restricted to the study team. Participants were compensated with CAD $50 (US $36.03) per interview session, in recognition of their time and contribution.

Data Collection

Demographic information including sex and gender, ethnic identity, socioeconomic status, and geographic data (rural or urban) was collected prior to the first interview. While participants did report gender, they were referred based on a biological basis (ie, risk for or current diagnosis of prostate or breast cancer). Two semistructured interview sessions, each approximately 1 hour in length, were conducted over telephone or video (Zoom) by a graduate-level research assistant (TK) who had been trained in qualitative interviewing, with a background in clinical psychology, psychology, global health, and sociology. The 2 interviews were conducted approximately 2 weeks apart, allowing time for reflection and deeper meaning-making while maintaining recall continuity [28,29].

An interview guide was used, with flexibility for the interviewer to probe participants' experiences and perspectives. The interview guide included open-ended questions about participants' general perceptions of AI in health care; specific feelings about AI-powered robots in cancer screening, diagnosis, and treatment; perceived barriers and facilitators to AI adoption; and preferred communication/education strategies. The full interview guide is provided in Multimedia Appendix 1. During interviews, the interviewer clarified that the term *AI-powered medical robots* referred broadly to robotic systems enhanced with AI that could be applied to different parts of the cancer pathway, including imaging, biopsy guidance, and treatment support. This clarification was intended to provide participants with a shared reference point, while still allowing their natural perceptions and assumptions to guide their responses. Participants were not given a detailed technical briefing, as the intent was to elicit their existing awareness and spontaneous perceptions of AI technologies in health care. This approach

XSL•FO
RenderX

ensured that participant reflections represented authentic, experience-based understandings rather than those influenced by study-provided information.

All participants were informed during recruitment and reminded prior to each interview that the study focused on their cancer care experiences and their perspectives regarding AI-powered medical robots. The first interview emphasized participants' initial reactions, beliefs, and expectations about AI-powered medical robots in cancer care. The second interview revisited these themes in more depth, provided an opportunity to clarify or expand on responses, and explored additional reflections participants had after the first discussion. This consistent framing ensured that discussions remained grounded in participants' personal experiences of cancer care while allowing flexibility for individual interpretation and reflection.

Data Analysis

Interview recordings were transcribed verbatim by 2 research assistants (MH and JD). A conventional content analysis approach, which relies on inductive coding [30], was followed. Coders (1) immersed themselves in the transcripts, (2) performed line-by-line coding, (3) grouped codes into subcategories, and (4) clustered subcategories into broader categories. The codebook (Multimedia Appendix 2) was developed inductively from the data, with new codes added when concepts not captured by existing codes were identified. Iterative updates to the codebook were justified when coders encountered new or refined meanings across interviews, and changes were documented in the audit trail. To ensure rigor, MH and JD independently coded 30% of the transcripts, meeting to compare and reconcile code discrepancies. Disagreements were resolved through consultation with a third research team member (MBK), a graduate-level research coordinator with training in qualitative research methodology. Once agreement on the initial codebook was achieved, the coders each analyzed half of the remaining transcripts, routinely consulting with one another and MBK to discuss emerging ideas and refine codes as needed. The final coding tree was established by the 3 reviewers with input from a fourth member of the research team (AP) and the site principal investigator (JE). Including multiple perspectives helped limit the potential impact of biases of any given researcher.

Though formal coding began after all interviews were completed, the interviewer made informal notes on emerging themes during the interviews to assess data saturation [31] and guide iterative refinement of the interview guide. Reflexive memos were maintained to document thoughts on the saturation process. Reviewers maintained a thorough audit trail outlining the identification and adaptation of new and existing codes. Saturation was then formally assessed retrospectively, over the course of data analysis. Coders reviewed transcripts sequentially and confirmed that after the tenth interview, no new substantive codes were identified. The remaining interviews contributed additional depth and nuance but did not generate new categories, confirming that saturation had been reached. Given that participants were recruited from a single large urban tertiary cancer center and the sample was predominantly White and urban-residing, it is possible that this relative demographic homogeneity contributed to the point at which saturation was reached. This demographic profile reflects the typical patient population served within the breast and prostate cancer clinics at Sunnybrook Health Sciences Centre.

Positionality Statement

The research team represented a range of disciplinary, cultural, and experiential backgrounds. Interviews were conducted by a female master's student and research assistant with training in clinical psychology, global health, psychology, and sociology. Coding and analysis were carried out by 2 female research assistants (BAs in psychology), a female research coordinator with a master's degree in medical sciences and formal qualitative research training, and a male research coordinator with a BA in psychology and training in clinical psychology. Team members reflected diversity in gender and cultural backgrounds, which contributed to varied perspectives in the analytic process. The site principal investigator, a female psychiatrist specializing in psychosocial oncology and trauma, and the coprincipal investigator, a male professor of surgery with expertise in intelligent robotics, provided oversight for the study.

Trustworthiness

To mitigate potential bias and increase credibility of the findings generated in this study, we used reflexive practices, including regular peer debriefing between the coders and the third reviewer, to identify, acknowledge, and reconcile assumptions and interpretations. Engaging in reflexivity as we continuously rereviewed and discussed the transcripts allowed us to ensure that patient voices remained central in the findings, which supported confirmability of the findings. We also transcribed audio recordings verbatim for analysis, kept a thorough audit trail, and have thoroughly contextualized our work and described our methods as means of increasing dependability of this work. To address transferability, we have provided a thorough description of the study methods (including the interview questions) and the participant demographics.

Results

Participant Demographics

In total, 15 participants (6 female and 9 male participants) were recruited, all of whom completed both interviews. Of the 14 participants who reported demographic data (Table 1), 80% (n=12) were White. The mean age of participants was 67 (SD 12; range 41 - 88) years. 80% (n=12, 7 male and 5 female participants) of reporting participants lived in urban settings, 13% (n=2, 1 male and 1 female participant) lived in rural settings. Of the 10 participants who reported their annual household income, 7 (70%) reported an annual household income over US $60,000.

XSL•FO
RenderX

Table . Participant demographics[a].

Demographic category	Count, n (%)
Sex	
Male	9 (60.0)
Female	6 (40.0)
Ethnic identity	
White/Caucasian	12 (80.0)
Asian	1 (6.7)
Middle Eastern	1 (6.7)
Age range (y)	
40 - 49	1 (6.7)
50 - 59	2 (13.3)
60 - 69	5 (33.3)
70 - 79	5 (33.3)
80 - 89	1 (6.7)
Household income[b] (US $)	
Below 14,630 (CAD 20,000)	1 (6.7)
14,630 (CAD 20,000) - 29,258 - (CAD 39,999)	2 (13.3)
29,259 (CAD 40,000) - 43,888 - (CAD 59,999)	—[c]
43,889 (CAD 60,000) - 58,517 - (CAD 79,999)	4 (26.7)
58,518 (CAD 80,000) - 72,416 - (CAD 99,000)	—
73,148+ (CAD 100,000+)	3 (20.0)

[a]One prostate cancer participant did not report any demographic information.

[b]Household income was originally collected in Canadian dollars and converted to US dollars using a currency exchange rate of CAD $1 = US $0.73

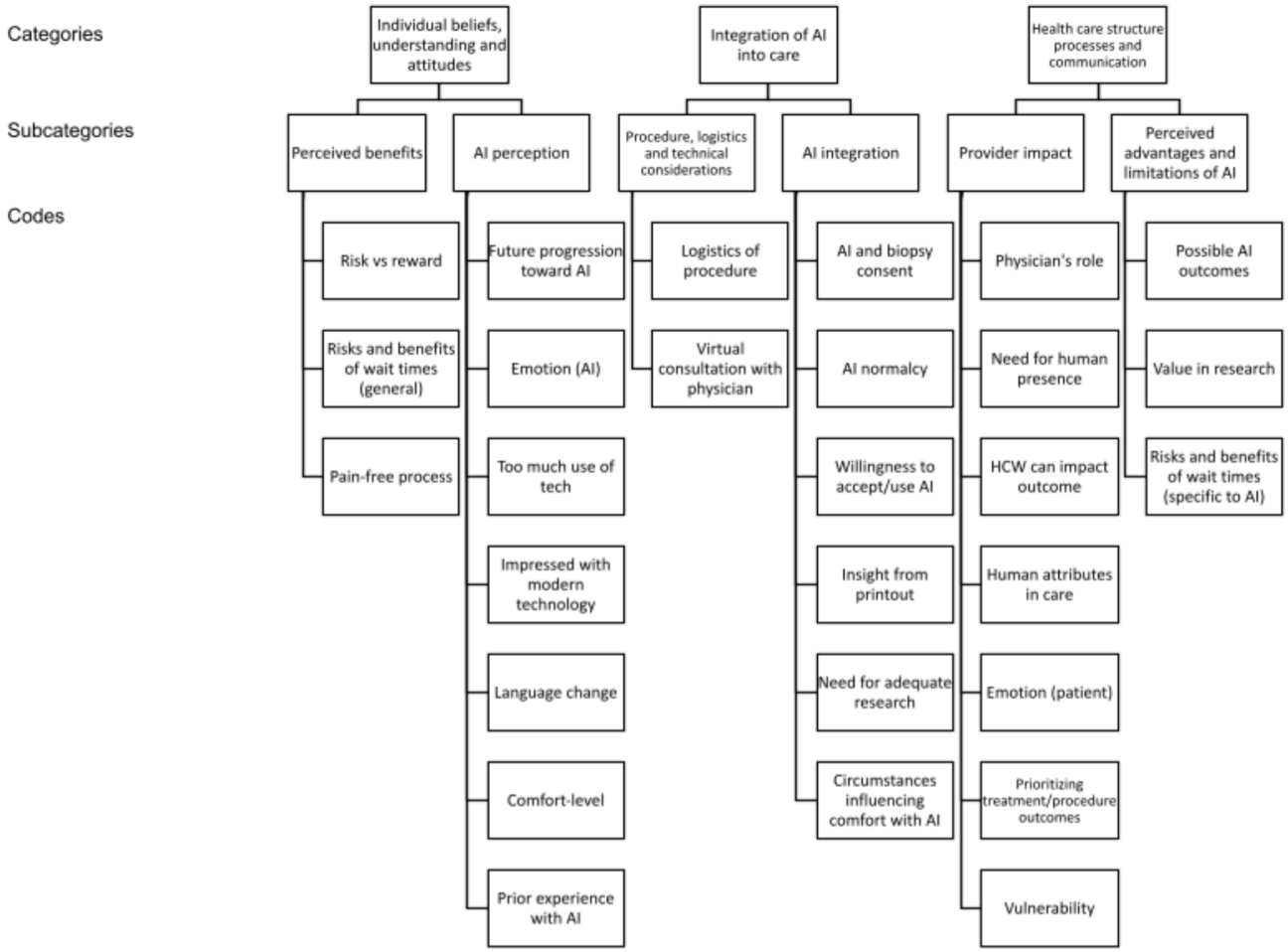
[c]Not available.

Thematic Findings

The analysis yielded 28 codes, which were then organized into 6 subcategories and further grouped into 3 overarching categories to best understand participant feelings of integrating AI-assisted medical technologies into their cancer care. The 3 overarching categories were (1) individual beliefs, (2) understanding, and (3) attitudes, integration of AI into care, and health care structure, process, and communication (Figure 1).

XSL•FO
RenderX

Figure 1. Final coding tree: organization of codes>subcategories>categories. AI: artificial intelligence; HCW: health care worker.



In the *individual beliefs, understanding, and attitudes* category, participants described a mix of openness and hesitancy toward AI-assisted technologies in their cancer care. Acceptability often depended on whether AI could meaningfully improve the experience of care, such as offering greater accuracy, reducing discomfort, or improving efficiency. Ongoing concerns centered on overreliance on technology, safety, and the need for clear backup plans in case of system failure. While some participants negatively described a general overreliance on technology, others noted that acceptance would increase as evidence accumulated:

> *[T]hey [are] as good as a human.* [ID 15, F, 53]

Concerns about dependability were especially salient in the context of unexpected events:

> *[C]ertain things can happen, who would be available, in an electrical shutdown and so on? [...] I don't know the capability of these machines, how reliable they are [...] it's reassuring to have help nearby to jump in.* [ID 50, M, 73]

Across participants, trust in AI was described as something that would grow through experience and familiarity.

The *Integration of AI into Care* category focused on the procedural, technical, and logistical considerations of integrating AI-robotic medical technology into cancer care, as well as general perceptions on AI integration into care. Participants distinguished between different stages and levels of invasiveness in the cancer pathway. Many were comfortable with AI in roles such as imaging or screening but more cautious regarding AI involvement in biopsies or surgery.

As one participant summarized:

> *I think I am okay with [the machine doing] the screening, if you could speed up the wait time. For the actual surgery I think I need more proven evidence, that it is 100% positive and reliable.* [ID 30, F, 62]

Human presence, whether in-person or virtual, was consistently described as important for reassurance, oversight, and troubleshooting during procedures:

> *[K]nowing that they were there and that there was some sort of troubleshooting capacity, would be helpful.* [ID 80, M, 65]

Participants also highlighted the need for accessible explanations of how AI works, its benefits, and its limitations. They emphasized that information should be made available proactively and through multiple accessible formats, given that patients may be unfamiliar or uncomfortable with digital resources:

> *I had to take the initiative to search through the website to get [information] and knowing so many people, they are not familiar with using technology, or some seniors, I think that was, you know, it limited*

XSL•FO
RenderX

> *the opportunity for people to get [access to] support resources.* [ID10, M, 60]

The category *health care structure, processes, and communication* described challenges within the existing health care system, particularly long wait times, fragmented communication, and limited opportunities for personalization. These issues shaped patients' broader care experiences and influenced how they envisioned AI fitting into their care. Some felt "lost in the system" [ID 45, F, 69], due to the perception that the current health care system lacks personalization.

Participants also associated long delays and fragmented care with distress and uncertainty:

> *I couldn't get any surgery done immediately. [...] It was very difficult. And I couldn't get care in a timely way.* [ID 40, M, 78]

Several participants expressed hope that AI-assisted technologies might address some of these system-level pressures by supporting more timely and/or coordinated care, while emphasizing that technological integration should not reduce human connection in the care process.

Barriers and Facilitators

Across interviews, participants described specific conditions that shaped how acceptable AI-assisted medical technologies felt to them in their care. Consistent with qualitative studies examining patient adoption of new health technologies and with implementation science frameworks (eg, studies that explore contextual factors influencing uptake) [31], these perceived conditions are referred to as *barriers* (factors that reduced comfort or confidence) and *facilitators* (factors that increased openness or ease of acceptance). Table 2 provides an overview of the barriers to integrating AI-assisted medical robots into cancer screening, diagnosis, and treatment.

Table . Factors influencing patients' decision to accept AI[a] in cancer care (barriers and facilitators).

Factor	Barrier	Facilitator
Information	Lack of information regarding the technology left patients unable to answer whether they were comfortable using it.	Providing adequate patient education and related resources would facilitate the uptake of AI-assisted technology.
Vulnerability	These patient populations are particularly in need of human care due to the vulnerable nature of the procedures.	—[b]
Reliability of technology	Fear of technical malfunctions or similar issues.	Potential for increased precision and efficiency with the use of technology; some patients noted that the most important thing is a positive outcome, regardless of whether a doctor or robotic device is completing the procedure.
Human element	Machine lacking in compassion/bedside manner in critical thinking/problem-solving skills.	No risk of human error when using AI technology.
Reduced wait times	—	The potential for reduced time between initial appointment and treatment can improve patient outcomes and eliminate distress associated with waiting periods.
Prior use of AI-assisted technology in health care	—	Patients express prior knowledge or direct experience with some form of AI technology in health care, creating a greater sense of comfort with using AI.
Pain	—	Some participants stated that they would be more inclined to consent to the procedures if they were less painful than the standard of care.
Options available	Patients might prefer standard of care when they have the option (given that the risks/rewards are similar between these options).	If there were no other options, or standard of care came with increased risk (due to waiting times, etc), patients would be more likely to consent to use of AI-assisted technology.

[a]AI: artificial intelligence.

[b]Not applicable.

Barriers most often reflected concerns about the interpersonal and safety-related dimensions of AI-assisted care. Participants frequently worried about the absence of human qualities such as compassion, empathy, and emotional support, particularly during invasive or high-stakes procedures. Many also questioned the reliability and safety of AI systems, describing uncertainty about how unexpected complications would be managed without a health care provider physically present. Some participants expressed apprehension that technological integration could reduce opportunities for communication with clinicians, thereby amplifying existing feelings of depersonalization within the health care system.

XSL•FO
RenderX

Facilitators described by patients centered on ways AI-assisted technologies could meaningfully improve their care experience. The potential to reduce wait times, streamline care, and deliver more timely results was consistently valued and often linked to reductions in distress, uncertainty, and logistical burden. Participants also described their confidence in AI's technical capabilities, such as accuracy and precision, as an important factor that could support acceptance. Several participants noted that when outcomes were perceived as likely to improve, the mode of delivery (human vs robot) became less central. Prior familiarity with AI in health care and the possibility of less painful or less invasive procedures further contributed to openness toward AI integration.

Discussion

Principal Findings

This study found that patients with breast and prostate cancer expressed cautious openness toward the use of AI-powered medical robots in their care, despite initial skepticism. Participants recognized potential benefits such as reduced wait times, improved accuracy, and less invasive procedures and also voiced concerns about safety, reliability, and the potential loss of human connection in care. Participants expressed enthusiasm for AI-assisted medical robots when given the choice, suggesting openness to innovation despite some skepticism. They often prioritized successful treatment outcomes over concerns of who or what delivered care, indicating that demonstrated efficacy could outweigh reservations. Trust in AI was described as something that would build gradually over time as experience and familiarity increased. Their accounts also underscored the critical need for systemic improvements in health care processes, particularly in reducing delays and streamlining care. These findings highlight the perceived potential of AI technologies to address longstanding system challenges, while also reflecting ambivalence regarding the risks of technological reliance. Notably, participants were generally open to the use of AI-powered medical robots in initial screening and imaging but were more hesitant when considering invasive steps such as biopsy or treatment, where human presence and oversight were prioritized.

As an exploratory qualitative study, the aim was to elucidate patients' feelings, understandings, and perceptions about the use of AI-powered medical robots in cancer screening, diagnosis, and early treatment and to identify barriers and facilitators to patient acceptance. Participant responses illustrated a mix of hesitancy and hope around embracing this technology for their own care, which is a typical response to technological advances or change [30,31]. Importantly, participants' reflections consistently emphasized that the human presence in care—particularly the compassion, reassurance, and emotional support provided by health care providers—is viewed as irreplaceable [32]. This emphasis on the importance and value of human presence in care (particularly that health care providers should remain central in a patient's care processes) underscores the importance of emotional support and other traits that are unique to humans. Without deliberate safeguards to preserve these aspects, the integration of AI technologies risks exacerbating feelings of depersonalization in care. These insights, while not generalizable due to the small purposive sample, provide valuable direction for future research and early implementation strategies.

Comparison to Prior Work

Overall, participants' reflections were broadly consistent with existing research examining patient perspectives on AI and technology-assisted care. Similar to prior work, participants expressed interest in potential benefits such as improved efficiency and reduced wait times. These factors are repeatedly identified in the literature as persistent challenges in cancer care pathways [7,16,17]. Our findings align with studies showing that patients often view AI as promising when it addresses established system-level pain points, particularly diagnostic delays and fragmented processes.

Concerns regarding the loss of human connection were also consistent with previous research demonstrating resistance to medical AI when patients perceive threats to empathy, communication, or relational aspects of care [33]. Participants emphasized the irreplaceable value of compassion and reassurance, echoing patient-centered care literature, highlighting communication, emotional support, and trust as core needs during cancer care [34-36]. Prior studies similarly note that while AI may assist with technical tasks, it cannot fully replicate relational qualities valued in clinician—patient interactions [32,33].

Consistent with earlier work, participants indicated that evidence of safety and effectiveness would be central to their acceptance of AI-assisted technologies. This is in line with studies showing that patient trust in AI increases when technologies are validated, transparent, and clearly integrated within human-led care models [30,31]. The desire for clear education about AI, expressed by many participants, is also supported by previous findings emphasizing the importance of accessible, multimodal patient education to support uptake of new technologies [35,37,38]. Altogether, the results reinforce themes seen across the broader literature: patients may be cautiously optimistic about AI in cancer care when it enhances, rather than replaces, human-delivered care; when it improves timeliness and coordination; and when adequate evidence and communication support informed decision-making. These consistencies suggest that early implementation strategies should prioritize transparency, patient education, and explicit preservation of human presence, particularly during invasive or emotionally charged stages of cancer care.

Strengths and Limitations

There are several limitations to this study. First, recruitment occurred at a single large urban hospital, which may not be representative of all patient populations. Participants were drawn from a tertiary cancer center in Toronto, potentially excluding individuals from rural or underserved regions and therefore excluding perspectives from those with limited access to tertiary care. This is related to a previously established limitation regarding the disparity in who has access to high-quality health care [39]. This may have influenced findings by overrepresenting patients with greater access to resources and

XSL•FO
RenderX

technology. Future research should intentionally recruit participants from rural and underserved regions to increase transferability.

Second, the sample was predominantly White (80%), urban-dwelling, and higher-income. Urban and higher-income populations may have more exposure to and comfort with advanced technologies compared to their rural and lower-income counterparts [40]. Additionally, the perspectives of those who live in rural settings may be impacted by experiences with resource limitations and accessibility of health care [41,42]. As a result, the findings may not reflect the views of patients from more diverse racial, ethnic, or socioeconomic groups. Within the constraints of this study setting and recruitment pathways, we were unable to broaden the sample to include more diverse participants. Although participants were provided with a shared reference point for what was meant by "AI-powered medical robots" (ie, robotic systems enhanced with AI used in imaging, biopsy guidance, and treatment support), they were not given a detailed technical briefing. As such, their responses may have reflected variability in baseline knowledge. Some participant responses also appeared general in tone (eg, referring broadly to *AI* or *technology*), but these were expressed in relation to their cancer care experiences. This likely reflects natural variability in familiarity with AI concepts among participants rather than deviation from the study focus. This could be seen as a limitation in terms of consistency, but it also provided valuable insights into organic perceptions, assumptions, and information needs regarding AI in cancer care. Third, data collection took place during the COVID-19 pandemic, which may have impacted attitudes toward access to health care, or experiences with health care delays may have been impacted by the health care climate at the time [43,44]. This context likely heightened sensitivity to wait times and system inefficiencies. While this provided timely insights, future studies should examine whether similar themes emerge in nonpandemic contexts. Finally, the sample size was relatively small, though appropriate for a qualitative descriptive study. Data saturation was achieved and confirmed during analysis; however, larger samples in future research could support broader transferability of findings and capture a greater diversity of perspectives. Despite its size, the study's use of 2 interviews per participant and a multianalyst coding process strengthened the dependability and credibility of the findings. Additionally, collecting perspectives from both breast and prostate cancer populations provided a richer understanding of how AI-assisted technologies may be perceived across different cancer care experiences.

Future Directions

Future research should ensure to include perspectives from patients who represent more diverse lenses as a means of increasing transferability of findings and understanding specific needs for integrating AI-assisted medical technology into cancer care in varying regions and cultural contexts. Future studies should also aim to include perspectives from all involved in the health care process (such as physicians and oncology assistants) to better understand how to pragmatically incorporate AI-assisted medical robots into cancer care screening, diagnosis, and treatment. More work is needed to understand which implementation models work best for both health care providers and patients, including how to provide education on the technology to both, what level of human presence is deemed acceptable and appropriate, and safe mitigation strategies in the event of technical issues. Future trials should also examine patient acceptability separately across different stages of the cancer pathway (eg, screening, diagnosis, and treatment) to determine where AI integration is most feasible and where stronger human involvement is needed.

Conclusions

In conclusion, this work explored patient perspectives on the use of AI-assisted medical robots in cancer screening, diagnosis, and early treatment, highlighting barriers and facilitators to its acceptance. The findings reveal a cautious optimism among patients, tempered by concerns about the reliability of AI and the potential for reduced human contact. Patients expressed a desire for AI to complement, rather than replace, human judgment, ensuring that final decisions remain in the hands of experienced health care providers. Alongside emphasizing the need for human presence and emotional support within the health care experience, patients described a need for education and clear communication regarding the use of new technology and suggested potential methods of making these resources more accessible (eg, through media and information sessions). As a qualitative exploratory study, these conclusions should be understood as reflecting patient perceptions within the sampled group, rather than definitive statements about broader populations. Future research is required to test these perspectives in larger and more diverse samples and to evaluate whether these patient-identified barriers and facilitators translate into measurable outcomes during actual implementation.

Acknowledgments

This study was supported by the Center for Surgical Invention and Innovation at McMaster University.

Generative AI was not used in the preparation or writing of this article.

Data Availability

The qualitative interview data generated and analyzed during this study are not publicly available due to privacy and confidentiality restrictions, as transcripts contain potentially identifiable information from patient participants. Deidentified excerpts supporting the study findings are included within the article. Researchers who wish to access the full deidentified data set may contact the corresponding author, and requests will be considered on a case-by-case basis in accordance with institutional Research Ethics Board approval requirements.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: JE, MBK, MA, KB
Data curation: MBK, TK, JE, MH
Formal analysis: MH, JD, MBK, JE
Funding acquisition: JE, MA, KB, RM, AA
Investigation: JE, MA, KB
Methodology: JE, MBK, MA, KB
Project administration: MBK, MH, TK
Software: JE
Supervision: JE
Visualization: MH, JD, MBK, JE
Writing - original draft: MBK, MH, AP, JE
Writing - review & editing: JE, MBK, MA, MH, KB, TK, RM, AP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 22 KB - cancer_v12i1e69710_app1.docx]

Multimedia Appendix 2
Codes and definitions.
[DOCX File, 11 KB - cancer_v12i1e69710_app2.docx]

Checklist 1
Completed COREQ checklist.
[PDF File, 464 KB - cancer_v12i1e69710_app3.pdf]

References

1. Kaviani E, Passi K. Forecasting cancer incidence in Canada by age, sex, and region until 2026 using machine learning techniques. Algorithms 2025;18(5):265. [doi: 10.3390/a18050265]
2. Richardson LC, Royalty J, Howe W, Helsel W, Kammerer W, Benard VB. Timeliness of breast cancer diagnosis and initiation of treatment in the National Breast and Cervical Cancer Early Detection Program, 1996-2005. Am J Public Health 2010 Sep;100(9):1769-1776. [doi: 10.2105/AJPH.2009.160184] [Medline: 20019308]
3. Stevens C, Bondy SJ, Loblaw DA. Wait times in prostate cancer diagnosis and radiation treatment. Can Urol Assoc J 2010 Aug;4(4):243-248. [doi: 10.5489/cuaj.09122] [Medline: 20694099]
4. Weller D, Vedsted P, Rubin G, et al. The Aarhus statement: improving design and reporting of studies on early cancer diagnosis. Br J Cancer 2012 Mar 27;106(7):1262-1267. [doi: 10.1038/bjc.2012.68] [Medline: 22415239]
5. Petrova D, Špacírová Z, Fernández-Martínez NF, et al. The patient, diagnostic, and treatment intervals in adult patients with cancer from high- and lower-income countries: a systematic review and meta-analysis. PLoS Med 2022 Oct;19(10):e1004110. [doi: 10.1371/journal.pmed.1004110] [Medline: 36264841]
6. Chan RJ, Milch VE, Crawford-Williams F, et al. Patient navigation across the cancer care continuum: an overview of systematic reviews and emerging literature. CA Cancer J Clin 2023;73(6):565-589. [doi: 10.3322/caac.21788] [Medline: 37358040]
7. Reece JC, Neal EFG, Nguyen P, McIntosh JG, Emery JD. Delayed or failure to follow-up abnormal breast cancer screening mammograms in primary care: a systematic review. BMC Cancer 2021 Apr 7;21(1):373. [doi: 10.1186/s12885-021-08100-3] [Medline: 33827476]
8. Chaput G. Oncology workforce crisis: let's work smarter, not harder. Curr Oncol 2023 Dec 11;30(12):10408-10409. [doi: 10.3390/curroncol30120758] [Medline: 38132392]
9. Cao DJ, Hurrell C, Patlas MN. Current status of burnout in Canadian radiology. Can Assoc Radiol J 2023 Feb;74(1):37-43. [doi: 10.1177/08465371221117282] [Medline: 35938488]
10. Wever EM, Draisma G, Heijnsdijk EAM, de Koning HJ. How does early detection by screening affect disease progression? Modeling estimated benefits in prostate cancer screening. Med Decis Making 2011;31(4):550-558. [doi: 10.1177/0272989X10396717] [Medline: 21406620]
11. Ferrell BR, Grant M, Funk B, Otis-Green S, Garcia N. Quality of life in breast cancer. Part I: physical and social well-being. Cancer Nurs 1997 Dec;20(6):398-408. [doi: 10.1097/00002820-199712000-00003] [Medline: 9409061]

XSL•FO
RenderX

12. Crosby D, Bhatia S, Brindle KM, et al. Early detection of cancer. Science 2022 Mar 18;375(6586):eaay9040. [doi: 10.1126/science.aay9040] [Medline: 35298272]
13. Bergin RJ, Emery JD, Bollard R, White V. Comparing pathways to diagnosis and treatment for rural and urban patients with colorectal or breast cancer: a qualitative study. J Rural Health 2020 Sep;36(4):517-535. [doi: 10.1111/jrh.12437] [Medline: 32485017]
14. Dobson C, Rubin G, Murchie P, Macdonald S, Sharp L. Reconceptualising rural cancer inequalities: time for a new research agenda. Int J Environ Res Public Health 2020 Feb 24;17(4):1455. [doi: 10.3390/ijerph17041455] [Medline: 32102462]
15. Urquhart R, Scruton S, Radford S, Kendell C, Hirsch E. Exploring men's experiences with follow-up care following primary treatment for prostate cancer in Atlantic Canada: a qualitative study. Curr Oncol 2023 Nov 25;30(12):10111-10123. [doi: 10.3390/curroncol30120735] [Medline: 38132369]
16. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019 Jun;6(2):94-98. [doi: 10.7861/futurehosp.6-2-94] [Medline: 31363513]
17. Farrokhi M, Moeini A, Taheri F, et al. Artificial Intelligence in Cancer Care: From Diagnosis to Prevention and Beyond 2023. [doi: 10.5281/zenodo.10040422]
18. Arvaniti E, Fricker KS, Moret M, et al. Automated Gleason grading of prostate cancer tissue microarrays via deep learning. Sci Rep 2018 Aug 13;8(1):12054. [doi: 10.1038/s41598-018-30535-1] [Medline: 30104757]
19. Herent P, Schmauch B, Jehanno P, et al. Detection and characterization of MRI breast lesions using deep learning. Diagn Interv Imaging 2019 Apr;100(4):219-225. [doi: 10.1016/j.diii.2019.02.008] [Medline: 30926444]
20. Kulkarni S, Seneviratne N, Baig MS, Khan AHA. Artificial intelligence in medicine: where are we now? Acad Radiol 2020 Jan;27(1):62-70. [doi: 10.1016/j.acra.2019.10.001] [Medline: 31636002]
21. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 2017 Dec 12;318(22):2199-2210. [doi: 10.1001/jama.2017.14585] [Medline: 29234806]
22. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. Nature New Biol 2020 Jan;577(7788):89-94. [doi: 10.1038/s41586-019-1799-6] [Medline: 31894144]
23. Song Y, Zhang YD, Yan X, et al. Computer‐aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. J Magn Reson Imaging 2018 Dec;48(6):1570-1577. [doi: 10.1002/jmri.26047] [Medline: 29659067]
24. Yin J, Ngiam KY, Teo HH. Role of artificial intelligence applications in real-life clinical practice: systematic review. J Med Internet Res 2021 Apr 22;23(4):e25759. [doi: 10.2196/25759] [Medline: 33885365]
25. Pesapane F, Giambersio E, Capetti B, et al. Patients' perceptions and attitudes to the use of artificial intelligence in breast cancer diagnosis: a narrative review. Life (Basel) 2024 Mar 29;14(4):454. [doi: 10.3390/life14040454] [Medline: 38672725]
26. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
27. Fox NJ, Ramazanoglu C. Postpositivism. In: Given LM, editor. The SAGE Encyclopedia of Qualitative Research Methods: SAGE Publications; 2008:659-664 URL: https://methods.sagepub.com/ency/edvol/sage-encyc-qualitative-research-methods/chpt/postpositivism [accessed 2026-01-14] [doi: 10.4135/9781412963909.n332]
28. Charmaz K. Constructing Grounded Theory, 2nd edition: SAGE Publications; 2014. URL: https://www.kobo.com/ww/en/ebook/constructing-grounded-theory-1 [accessed 2026-01-14]
29. Wilhelmy A, Köhler T. Qualitative research in work and organizational psychology journals: practices and future opportunities. Eur J Work Organ Psychol 2022;31(2):161-185. [doi: 10.1080/1359432X.2021.2009457]
30. Esmaeilzadeh P, Mirzaei T, Dharanikota S. Patients' perceptions toward human–artificial intelligence interaction in health care: experimental study. J Med Internet Res 2021 Nov 25;23(11):e25856. [doi: 10.2196/25856] [Medline: 34842535]
31. Madanian S, Nakarada-Kordic I, Reay S, Chetty T. Patients' perspectives on digital health tools. PEC Innov 2023 May;2:100171. [doi: 10.1016/j.pecinn.2023.100171] [Medline: 37384154]
32. Habib M, Korman MB, Aliasi-Sinai L, et al. Understanding compassionate care from the patient perspective: highlighting the experience of head and neck cancer care. Can Oncol Nurs J 2023;33(1):74-86. [doi: 10.5737/2368807633174] [Medline: 36789223]
33. Longoni C, Bonezzi A, Morewedge CK. Resistance to medical artificial intelligence. J Consum Res 2019 Dec;46(4):629-650. [doi: 10.1093/jcr/ucz013]
34. Elkefi S, Asan O. The impact of patient-centered care on cancer patients' QOC, self-efficacy, and trust towards doctors: analysis of a national survey. J Patient Exp 2023;10:23743735231151533. [doi: 10.1177/23743735231151533] [Medline: 36698621]
35. Giuliani M, Papadakos T, Papadakos J. Propelling a new era of patient education into practice—cancer care post-COVID-19. Int J Radiat Oncol Biol Phys 2020 Oct 1;108(2):404-406. [doi: 10.1016/j.ijrobp.2020.05.036] [Medline: 32890521]
36. Mitchell KAR, Brassil KJ, Rodriguez SA, et al. Operationalizing patient‐centered cancer care: a systematic review and synthesis of the qualitative literature on cancer patients' needs, values, and preferences. Psychooncology 2020 Nov;29(11):1723-1733. [doi: 10.1002/pon.5500] [Medline: 32715542]

XSL•FO
RenderX

37. Shamszare H, Choudhury A. Clinicians' perceptions of artificial intelligence: focus on workload, risk, trust, clinical decision making, and clinical integration. Healthcare (Basel) 2023 Aug 16;11(16):2308. [doi: 10.3390/healthcare11162308] [Medline: 37628506]
38. Tawfik DS, Sinha A, Bayati M, et al. Frustration with technology and its relation to emotional exhaustion among health care workers: cross-sectional observational study. J Med Internet Res 2021 Jul 6;23(7):e26817. [doi: 10.2196/26817] [Medline: 34255674]
39. Maddison AR, Asada Y, Urquhart R. Inequity in access to cancer care: a review of the Canadian literature. Cancer Causes Control 2011 Mar;22(3):359-366. [doi: 10.1007/s10552-010-9722-3] [Medline: 21221758]
40. Babbage DR, van Kessel K, Terraschke A, Drown J, Elder H. Attitudes of rural communities towards the use of technology for health purposes in New Zealand: a focus group study. BMJ Open 2020 Jun 1;10(6):e037892. [doi: 10.1136/bmjopen-2020-037892] [Medline: 32487583]
41. Coombs NC, Campbell DG, Caringi J. A qualitative study of rural healthcare providers' views of social, cultural, and programmatic barriers to healthcare access. BMC Health Serv Res 2022 Apr 2;22(1):438. [doi: 10.1186/s12913-022-07829-2] [Medline: 35366860]
42. Brems C, Johnson ME, Warner TD, Roberts LW. Barriers to healthcare as reported by rural and urban interprofessional providers. J Interprof Care 2006 Mar;20(2):105-118. [doi: 10.1080/13561820600622208] [Medline: 16608714]
43. de Joode K, Dumoulin DW, Engelen V, et al. Impact of the coronavirus disease 2019 pandemic on cancer treatment: the patients' perspective. Eur J Cancer 2020 Sep;136:132-139. [doi: 10.1016/j.ejca.2020.06.019] [Medline: 32683273]
44. Walker MJ, Wang J, Mazuryk J, et al. Delivery of cancer care in Ontario, Canada, during the first year of the COVID-19 pandemic. JAMA Netw Open 2022 Apr 1;5(4):e228855. [doi: 10.1001/jamanetworkopen.2022.8855] [Medline: 35467731]

Abbreviations

AI: artificial intelligence
COREQ: Consolidated Criteria for Reporting Qualitative Research
MRI: magnetic resonance imaging

Edited by M Balcarras, N Cahill; submitted 05.Dec.2024; peer-reviewed by CK Tan, N Calanzani, S Feder; revised version received 18.Nov.2025; accepted 29.Dec.2025; published 29.Jan.2026.

Please cite as:
Habib M, Ellis J, Palachi A, Korman MB, Kay T, Barlow K, DeSouza J, Macri R, Alabousi A, Anvari M
Patient Perspectives on AI-Powered Medical Robots in Breast and Prostate Cancer Care: Qualitative Study
JMIR Cancer 2026;12:e69710
URL: https://cancer.jmir.org/2026/1/e69710
doi:10.2196/69710

© Mahiya Habib, Janet Ellis, Aaron Palachi, Melissa B Korman, Tatjana Kay, Karen Barlow, Jordana DeSouza, Rosanna Macri, Abdullah Alabousi, Mehran Anvari. Originally published in JMIR Cancer (https://cancer.jmir.org), 29.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Mediastinal Germ Cell Tumors: Up Close and Personal

B Wayne Bequette[1,2], PhD

[1]Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, NY, United States

[2]Center for Engineering and Precision Medicine, New York, NY, United States

Corresponding Author:
B Wayne Bequette, PhD
Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, NY, United States

Abstract

Mediastinal germ cell tumors and how they differ from those associated with testicular cancer are reviewed through the experience with my son, Brendan Fahy Bequette. We reflect on others who have been diagnosed with rare tumors and discuss our need to honor Brendan's memory by "paying it forward."

(JMIR Cancer 2026;12:e93678) doi:10.2196/93678

KEYWORDS

germ cell tumor; mediastinal tumors; rare cancers; cancer treatment; acute myeloid leukemia

Background

Testicular cancer treatment is a great success story, largely due to the efforts of Dr Lawrence Einhorn at Indiana University to develop cisplatin-based chemotherapy protocols, with initial clinical trial success in 1974 [1,2]. Currently, someone diagnosed with testicular cancer has a 5-year expected survival rate of over 95%. A high-profile example is the cyclist Lance Armstrong, who was treated by Dr Einhorn in 1996. Germ cell tumors form the bulk of testicular cancers (95%). The TP53 gene is known as a tumor suppressor gene, and 50% of cancers have a TP53 gene mutation. Testicular cancers, however, have a low rate of TP53 mutations, contributing to the success of platinum-based chemotherapy treatment.

Mediastinal (in the chest between the lungs) germ cell tumors are very rare (2% of germ cell tumors) with a much lower treatment success than those that present in the testicles [3,4]. Unlike testicular cancers, mediastinal germ cell tumors tend to have a high rate of TP53 mutations. Nonseminoma mediastinal tumors treated with standard BEP (bleomycin, etoposide, cisplatin)–VIP (etoposide, ifosfamide, cisplatin) protocols have an approximately 5-year survival rate of 50%, which includes second-line treatment involving high dose chemotherapy with autologous stem cell transplant, where a patient's stem cells are separated from their blood and reinfused as part of the treatment [5]. Unfortunately, a relapse after high-dose chemotherapy results in a median survival time of 5 months. There are ongoing efforts to develop immunotherapy [6] and cell and gene (chimeric antigen receptor [CAR-T]) treatments for patients who have relapsed (often called platinum-resistant).

Personal Background

My son, Brendan Fahy Bequette, a filmmaker in New York City, was diagnosed with a mediastinal germ cell tumor in 2020 during the peak of the COVID-19 pandemic. While preparing for his initial treatment, we had a personal family contact with Dr Einhorn, who asked "how did Brendan know?" because a germ cell tumor normally progresses much further than Brendan's before detection. Brendan was a runner and had noted a pain in his chest before going to an outpatient clinic and receiving a chest X-ray 2 days before his 24th birthday. Dr Einhorn immediately contacted Dr David Shaffer at Albany Medical College (AMC), who was providing the treatment. Dr Shaffer also began correspondence with Dr Darren Feldman at Memorial Sloan Kettering Cancer Center (MSKCC) in New York. Clearly in this field of rare cancers, the specialists know each other quite well.

Brendan received his first line of treatment, BEP-VIP, at AMC; unfortunately, a brain scan indicated a metastasis, and his second line of treatment involved high-dose chemotherapy with an autologous stem cell transplant at MSKCC. While in remission, he began training for a marathon as a way of celebrating his conquering of cancer; unfortunately, the cancer returned after 2 months.

His third line of treatment was in an immunotherapy clinical trial using durvalumab and tremelimumab. Unfortunately, the cancer progressed, and he was removed from the trial after 1 month. As often happens with immunotherapy, his immune system had responded by attacking his thyroid and affecting his heart rate, causing him to need thyroid medication for the rest of his life. Overall, the clinical trial failed to meet the protocol-defined success for the 29 patients treated [7]. For the next 6 months, Brendan received several other lines of treatment, including 4 months of cisplatin plus epirubicin.

Brendan's last hope was a clinical trial for a novel CAR-T treatment for solid tumors [8]. Unfortunately, the trial was delayed by over 6 months while the sponsor automated the

XSL•FO
RenderX

process, and Brendan died on February 28, 2022, two weeks after the last enrollment.

Grief and Future Steps

My wife, Pat Fahy, and I have channeled our grief by publishing a book (The Photographs of Brendan Fahy Bequette) of black-and-white photos, many of which Brendan had taken during his cancer treatment [9,10]. The proceeds of this book support art and film awards in his name, thus "paying it forward" for young artists. When included with college scholarships in his memory, we honor approximately 8 young artists per year.

Further Reflections

The New Yorker article by Tatiana Schlossberg (John F Kennedy's granddaughter), "A Battle With My Blood," is subtitled "When I was diagnosed with leukemia, my first thought was that this couldn't be happening to me, to my family" [11].

Similarly, when we learned that our son had a tumor in his chest, Pat and I experienced the "this is not happening to us" state while simultaneously realizing that it was and that we needed to help Brendan start making a lot of important decisions about his health care. Tatiana goes on:

> *I did not—could not—believe that they were talking about me. I had swum a mile in the pool the day before, nine months pregnant. I wasn't sick. I didn't feel sick. I was actually one of the healthiest people I knew. I regularly ran five to ten miles in Central Park.*

Brendan was also fit, having been a long-distance runner since his freshman year of high school. Indeed, the reason that he went to an outpatient clinic was that he could feel a pain near his heart when sleeping the previous night—he was seriously in touch with his body. As a filmmaker in New York, he also often ran in Central Park.

When Pat and I hear that someone has been diagnosed with cancer, we immediately wonder or ask, "what type of cancer?" Tatiana was diagnosed with acute myeloid leukemia (AML) with chromosome 3 inversion, which is rare (roughly 1.5% of AML cases). Similarly, Brendan was diagnosed with a mediastinal germ cell tumor, which is approximately 2% of all germ cell tumors.

Roughly 200 or so people are diagnosed with AML with chromosome 3 inversion in the United States each year. This is almost the same as the roughly 200 or so people diagnosed with a mediastinal germ cell tumor each year in the United States. Thus, both Tatiana and Brendan had rare cancers. Notably, both were treated at MSKCC.

Role of Chemical Engineers in Therapeutics Manufacturing

Chemical engineers are largely responsible for the design and operation of the manufacturing plants that produce pharmaceuticals. Although the chemical engineering curriculum focuses on mathematical models and "algorithms," I encourage my students to think about the "human in the loop," which includes the safety of the personnel operating the plants and the consumer using the final product. I provide personal examples, such as my sister who was diagnosed with type 1 diabetes in 1977. This motivated me to collaborate on the development of an automated insulin delivery system to improve the lives of people living with type 1 diabetes. Similarly, I use the example of my son to motivate the ongoing development of immunotherapies for the treatment of cancers and other diseases [12]. In this way, I hope that my son continues to "pay it forward," as we educate the next generation of scientists and engineers.

Funding

The author declared that no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Vastag B. Testicular cancer seen as "Stalking Horse" for other cancers. J Natl Cancer Inst 1999 Nov 3;91(21):1801-1802. [doi: 10.1093/jnci/91.21.1801] [Medline: 10547383]
2. Vastag B. High-profile cancer cases prompt awareness efforts. J Natl Cancer Inst 1999 Nov 3;91(21):1802-1803. [doi: 10.1093/jnci/91.21.1802] [Medline: 10610405]
3. Rosti G, Secondino S, Necchi A, Fornarini G, Pedrazzoli P. Primary mediastinal germ cell tumors. Semin Oncol 2019 Apr;46(2):107-111. [doi: 10.1053/j.seminoncol.2019.04.001] [Medline: 31076171]
4. Ozgun G, Nappi L. Primary mediastinal germ cell tumors: a thorough literature review. Biomedicines 2023 Feb 8;11(2):487. [doi: 10.3390/biomedicines11020487] [Medline: 36831022]
5. Richardson NH, Taza F, Abonour R, et al. High-dose chemotherapy and peripheral blood stem cell transplantation as salvage therapy in primary mediastinal nonseminomatous germ cell tumors: the Indiana University experience. Cancer 2024 Sep 15;130(18):3115-3122. [doi: 10.1002/cncr.35375] [Medline: 38768296]

XSL•FO
RenderX

6. Adra N, Vaughn DJ, Einhorn LH, et al. A phase II study assessing the safety and efficacy of ASP1650 in male patients with relapsed refractory germ cell tumors. Invest New Drugs 2022 Oct;40(5):1087-1094. [doi: 10.1007/s10637-022-01276-w] [Medline: 35759134]
7. Funt SA, Knezevic A, Freeman BA, et al. A single-arm, phase II study of durvalumab (D) and tremelimumab (T) for relapsed/refractory germ cell tumors (GCT). J Clin Oncol 2023 May 31;41(16_suppl):5040-5040. [doi: 10.1200/JCO.2023.41.16_suppl.5040]
8. Mackensen A, Haanen J, Koenecke C, et al. CLDN6-specific CAR-T cells plus amplifying RNA vaccine in relapsed or refractory solid tumors: the phase 1 BNT211-01 trial. Nat Med 2023 Nov;29(11):2844-2853. [doi: 10.1038/s41591-023-02612-0] [Medline: 37872225]
9. Grondahl P. Cancer couldn't steal Brendan Fahy Bequette's creativity. Time Union. 2022 Nov 30. URL: https://www.timesunion.com/news/article/Posthumous-book-of-photographs-by-Brendan-Fahy-17617643.php [accessed 2026-03-18]
10. Brendan Fahy Bequette Fund. URL: https://bfbfund.com [accessed 2026-03-18]
11. Schlossberg T. A battle with my blood. The New Yorker. 2025 Nov 22. URL: https://www.newyorker.com/culture/the-weekend-essay/a-battle-with-my-blood [accessed 2026-03-18]
12. Bequette BW. Process Systems engineering in precision medicine: opportunities in autologous CAR-T therapy. Eng Life Sci 2026 Feb;26(2):e70067. [doi: 10.1002/elsc.70067] [Medline: 41647312]

Abbreviations

AMC: Albany Medical College
AML: acute myeloid leukemia
BEP: bleomycin, etoposide, cisplatin
CAR-T: chimeric antigen receptor
MSKCC: Memorial Sloan Kettering Cancer Center
VIP: etoposide, ifosfamide, cisplatin

Edited by M Balcarras; submitted 17.Feb.2026; this is a non–peer-reviewed article;accepted 09.Mar.2026; published 20.Mar.2026.

Please cite as:
Bequette BW
Mediastinal Germ Cell Tumors: Up Close and Personal
JMIR Cancer 2026;12:e93678
URL: https://cancer.jmir.org/2026/1/e93678
doi:10.2196/93678

© B Wayne Bequette. Originally published in JMIR Cancer (https://cancer.jmir.org), 20.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Monitoring Steps and Heart Rate Using a Withings Smartwatch in Children and Adolescents With Cancer: Validation Study

Emma den Hartog[1], PhD; Marta Fiocco[1,2], PhD; Patrick van der Torre[1], MSc; Wim J E Tissing[1,3], MD, PhD; Emma J Verwaaijen[1], PhD

[1]Princess Máxima Center for Pediatric Oncology, Heidelberglaan 25, Utrecht, The Netherlands

[2]Mathematical Institute, Leiden University, Leiden, The Netherlands

[3]Department of Pediatric Oncology and Hematology, University of Groningen, Beatrix Children's Hospital, Groningen, The Netherlands

Corresponding Author:
Wim J E Tissing, MD, PhD
Princess Máxima Center for Pediatric Oncology, Heidelberglaan 25, Utrecht, The Netherlands

Abstract

Background: Consumer smartwatches offer potential for continuous monitoring of physical activity during childhood cancer treatment. However, their validity for measuring step count and heart rate in this population has not been established.

Objective: This study examined the validity of a Withings smartwatch for monitoring step count and heart rate in children and adolescents undergoing cancer treatment.

Methods: Children and adolescents aged 8 to 18 years under active cancer treatment wore a Withings smartwatch for 12 weeks. For validation purposes, children and adolescents simultaneously wore an ActiGraph wGT3X-BT accelerometer for 1 week under free-living conditions to assess step count agreement and were connected to a clinical heart rate monitor under resting conditions to assess heart rate agreement. Linear mixed models with participant ID as a random effect were used to estimate mean differences and derive 95% limits of agreement for repeated measurements.

Results: A total of 14 children and adolescents participated in each of the validation analyses. The smartwatch recorded significantly fewer steps than the ActiGraph accelerometer for both steps per hour (mean difference 174 steps per hour; 95% limits of agreement=−283 to 633; $P<.001$) and steps per day (mean difference 3154 steps per day; 95% limits of agreement=−394 to 6702; $P<.001$), with increasing discrepancy at higher activity levels, suggesting proportional bias. No statistically significant difference in heart rate was observed between the smartwatch and the clinical heart rate monitor under resting conditions (mean difference −1.07 beats per minute; 95% limits of agreement=−17.55 to 15.41; $P=.24$).

Conclusions: In children and adolescents undergoing cancer treatment, the Withings smartwatch showed minimal systematic bias for resting heart rate assessment, although individual variability was considerable. In contrast, step counts demonstrated systematic underestimation and wide limits of agreement compared with a research-grade accelerometer. The use of smartwatch-derived step counts for clinical benchmarking or goal setting in this population cannot currently be recommended.

(JMIR Cancer 2026;12:e77766) doi:10.2196/77766

KEYWORDS

validity; smartwatch; steps; heart rate; pediatric oncology

Introduction

Children undergoing treatment for cancer demonstrate substantially reduced levels of physical activity compared to healthy peers [1-3]. Disease- and treatment-related symptoms such as fatigue, musculoskeletal pain, and general malaise frequently limit mobility and contribute to muscle wasting, reduced functional capacity, and decreased participation in daily life [1-3]. Persistently low physical activity may contribute to a cycle of physical deconditioning, which can further impair functional independence, social participation, and quality of life [4,5]. In addition to physical consequences, reduced activity levels have been associated with adverse psychosocial outcomes, including reduced emotional well-being and diminished cognitive and social functioning [6-10].

Therefore, promoting and monitoring physical activity during treatment may not only address physical deconditioning but also support psychosocial functioning and quality of life. Given these multidimensional consequences, insight into physical activity behavior during treatment is clinically relevant for children with cancer, their families, and health care professionals. Monitoring physical activity may support early identification of functional decline, enable timely and individualized guidance, and inform shared decision-making regarding physical activity and rehabilitation strategies.

XSL•FO
RenderX

However, such applications require accurate and feasible methods to assess physical activity under real-world conditions.

Objective measurement of physical activity in pediatric oncology relies primarily on research-grade accelerometers such as the ActiGraph, which are widely used as reference devices for quantifying physical activity, such as step count [11]. However, these devices do not provide information on physiological responses to activity, such as heart rate, and their use in routine clinical care is limited by practical considerations [12]. Data collection is usually restricted to predefined assessment periods; data are not directly accessible for children, parents, or health care professionals; and data extraction requires specialized software and expertise. Moreover, these devices may be experienced as inconvenient or uncomfortable by children. For these reasons, they are considered less suitable for continuous, long-term monitoring under free-living conditions [11].

Consumer-level wearable devices such as smartwatches offer the possibility to integrate measurement of physical activity (eg, step count) and physiological parameters such as heart rate within a single, low-burden device that can be worn continuously in daily life. This integration may provide a comprehensive picture of physical activity behavior during childhood cancer treatment while enabling real-time access to data for children, families, and health care professionals [11]. In this context, wearable devices have potential utility for monitoring activity patterns over time, supporting engagement in physical activity, and informing individualized guidance and rehabilitation strategies [13].

Several validation studies have evaluated consumer-grade wearable devices predominantly in adult populations, demonstrating promising accuracy for step count and heart rate under free-living conditions [14,15]. In pediatric populations, validation studies remain limited, and findings are heterogeneous [16]. Importantly, validation results obtained in adult or healthy pediatric populations cannot be readily generalized to children undergoing cancer treatment. In children with cancer, treatment-related symptoms, altered movement patterns, reduced gait speed, intermittent activity bouts, and physiological changes may influence sensor performance and algorithm accuracy. Therefore, device performance should be evaluated specifically within this population.

Among available consumer-grade devices, Withings smartwatches have frequently been used in research settings due to their suitability for research applications, including the availability of dedicated research infrastructure for data access and export and relatively stable software environments. Withings devices have been applied in pediatric studies to monitor physical activity behavior, primarily as tools to support physical activity interventions [17-19]. Although these devices have been validated in adult populations [17,20-25], validation has primarily focused on heart rate–related outcomes in pediatric populations, demonstrating acceptable agreement with reference methods [26]. Nonetheless, validation of Withings smartwatches in children undergoing cancer treatment in a wider context of physical activity has not yet been performed, and evidence regarding the validity of step count in any pediatric population remains limited.

Therefore, the aim of this study was to examine the validity of a Withings smartwatch for monitoring step count and heart rate in children and adolescents undergoing treatment for cancer.

Methods

Participants and Study Design

Participants were recruited within the framework of the Watch Us Move study, a prospective qualitative study examining the feasibility, barriers to, and facilitators of real-time and long-term physical activity tracking using a smartwatch in children and adolescents during treatment for cancer [18]. Eligible participants were aged 8 to 18 years; diagnosed with a malignancy; and receiving active treatment at the Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands. Additional inclusion criteria were the ability to walk at least 50 meters without the use of a walking aid, adequate comprehension of the Dutch language, and access to a device capable of running the app. Exclusion criteria included children and adolescents who were scheduled to undergo stem cell transplantation during the study period or who had intellectual disabilities that precluded participation in long-term smartwatch use.

Eligible patients were identified during routine clinical visits between May 2023 and June 2024 and were approached by the research team in consultation with their treating pediatric oncologist or pediatric physiotherapist.

According to the design of the Watch Us Move study, children and adolescents wore a smartwatch for 12 weeks consecutively. For this substudy, children and adolescents were asked to wear the smartwatch and an ActiGraph (Ametris) accelerometer simultaneously for 1 week and keep a wear log and report nonwear periods throughout the week. Moreover, children and adolescents were asked to wear the smartwatch while connected to a heart rate monitor under resting conditions while being admitted to the day care treatment ward.

Ethical Considerations

In accordance with Dutch legislation, for children aged 8 to 11 years, written informed consent was obtained from parents or legal guardians. For children and adolescents aged 12 to 15 years, written informed consent was obtained from both the participant and their parents or legal guardians. Adolescents aged 16 to 18 years provided written informed consent themselves. The study protocol was reviewed by the Medical Research Ethics Committee NedMec (reference number 23-003).

Devices

Withings Smartwatch (Device Under Evaluation)

The Withings Steel HR and Pulse HR smartwatches were used as the device under evaluation in this study. These devices were selected for several methodological and practical reasons. First, Withings provides access to structured research data through a dedicated application programming interface, allowing for the secure retrieval of step count and heart rate data for analysis. In addition, the software environment of these devices is relatively stable, reducing the likelihood of major algorithm changes during the study period. Furthermore, Withings uses

XSL•FO
RenderX

European data storage centers, which was considered important in light of the General Data Protection Regulation. Finally, the relatively small size, moderate cost, and design resembling a conventional wristwatch may enhance acceptability and feasibility for use in children and adolescents undergoing cancer treatment.

The Steel HR (model: HWA03b-36black-Inter or HWA03b-36white-Inter; 36.3 mm) is a hybrid smartwatch, and the Pulse HR (model: WAM03-Black Mirror-All-Inter; 18 × 44 mm) is a health and fitness tracker. Both devices use identical software for monitoring step count and heart rate.

Children and adolescents were allowed to choose the wrist on which they wore the Withings smartwatch to optimize comfort and adherence. When no preference was expressed, they were instructed to wear the device on the nondominant wrist in accordance with manufacturer recommendations. Wrist side was not systematically recorded.

Both devices were worn on the wrist and were capable of measuring step count, heart rate, distance covered, and estimated energy expenditure (calories). Data were synchronized with the Health Mate mobile app. Withings smartwatch data were manually retrieved from the Withings data server via the research data interface.

A day was considered valid when heart rate was recorded at least once per hour for a minimum of 6 hours between 6 AM and midnight. Nonwear time was defined as a period of at least 60 consecutive minutes without heart rate and step measurements. This definition was applied to ensure consistency with the nonwear definition used for the reference ActiGraph device (Troiano algorithm) [27]. For step count validation, only valid wear days were taken into account. For heart rate validation, the complete period during which children and adolescents were connected to the clinical heart rate monitor was considered.

ActiGraph (Reference Device for Step Count)

The ActiGraph wGT3X-BT accelerometer was used as the reference device for step count measurement under free-living conditions. The ActiGraph is widely used in pediatric research and, although it does not represent a criterion gold standard under free-living conditions, it is frequently applied as a research-grade comparator device in physical activity validation studies [28-30]. The ActiGraph was worn on the same wrist as the Withings smartwatch whenever possible. If this caused irritation or burden, it was worn on the opposite wrist. Wrist side (dominant vs nondominant) was not systematically recorded.

ActiGraph data were obtained using the ActiLife software (Ametris). A day was considered valid when steps were measured for at least 6 hours between 6 AM and midnight, with a minimum of 4 valid wear days per participant. Nonwear time was defined using the Troiano algorithm [27], in which a period of at least 60 consecutive minutes of zero activity counts was classified as nonwear, allowing for 1 to 2 minutes of activity counts between 0 and 100 [27]. Only valid wear days were included in step count analysis.

Philips IntelliVue X3 (Reference Device for Heart Rate)

The Philips IntelliVue X3 clinical monitor was used as the reference device for heart rate measurements. This hospital-grade monitoring system provides continuous electrocardiography-based heart rate measurements and is routinely used in clinical care. Therefore, it was considered an appropriate reference under resting clinical conditions. Children and adolescents were connected to the monitor under resting conditions while admitted to the day care treatment ward. Heart rate data were retrieved from the electronic patient records.

Data Analysis

Descriptive statistics were used to summarize patient characteristics. Categorical variables were presented as numbers with percentages. Normally distributed continuous variables were reported as means with SDs, and nonnormally distributed data were presented as medians with IQRs or ranges.

Outcome measures included steps per hour (total number of steps), steps per day and heart rate per minute (beats per minute [bpm]). Measurements obtained using the smartwatch were compared with those from the reference devices (ActiGraph and heart rate monitor). Only days considered valid for both the smartwatch and ActiGraph were included in the step count analyses. All available heart rate measurements obtained during connection to the clinical monitor were included in the heart rate analysis.

Linear mixed-effects models were estimated to assess the mean difference between devices for each outcome measure. These models account for the presence of repeated measurements within participants. Participant ID was included as a random intercept. The fixed intercept of each model represents the estimated mean difference (bias) between devices.

For agreement analyses, 95% limits of agreement (LoA) were derived from the mixed-effects models. In accordance with recommendations for repeated-measure Bland-Altman analyses, the SD of the differences was calculated as the square root of the sum of the between- and within-subject variance components (ie, √[variance_participant + residual variance]) [31]. The LoA were then calculated as mean difference ± 1.96 × SD_total. Agreement was visually assessed using scatterplots and Bland-Altman plots. All analyses were performed in SPSS Statistics (version 29.0.1; IBM Corp) or R (version 4.4.1; R Foundation for Statistical Computing).

Results

Step Validation

During the study period, 14 children and adolescents wore the smartwatch simultaneously while wearing an ActiGraph in free-living conditions for 1 week (Figure 1). Of these 14 children and adolescents, 8 (57.1%) were female, and they had a median age of 13.5 (IQR 10.5-15.3) years and weight-to-height *Z* score of 0 (IQR −1.5 to 1.6; Table 1).

XSL•FO
RenderX

Figure 1. Flow diagram of the inclusion of patients.

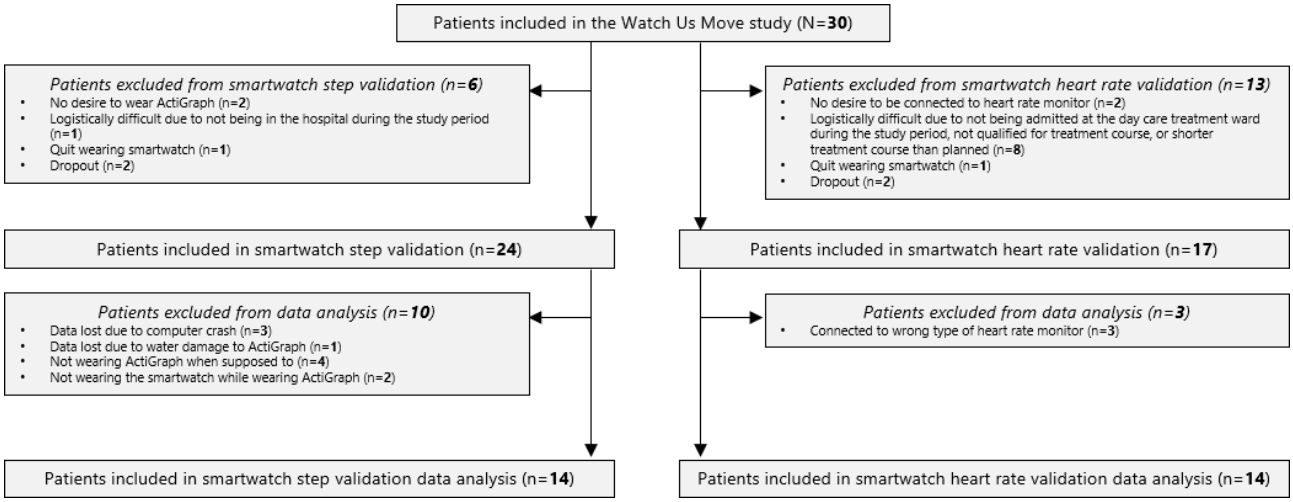


Table . Patient characteristics and free-living step count data during the step count validation study (n=14).

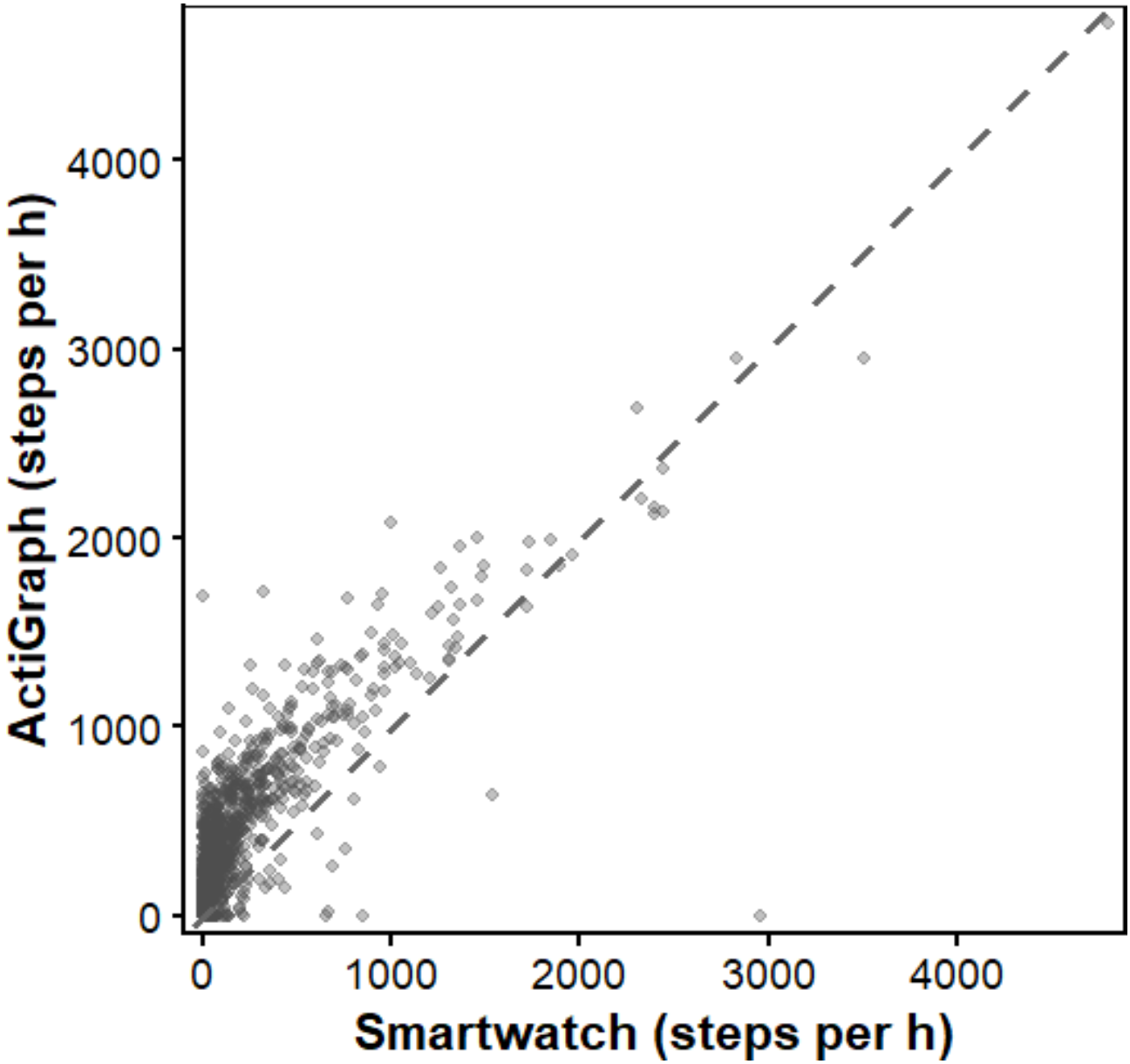
Characteristic	Values
Sex, n (%)	
Female	8 (57.1)
Male	6 (42.9)
Age (y), median (IQR)	13.5 (10.5 to 15.3)
Time since diagnosis (d), median (IQR)	108.5 (63.8 to 201.0)
Diagnosis, n (%)	
Acute lymphoblastic leukemia	6 (42.9)
Lymphoblastic lymphoma	3 (21.4)
Hodgkin lymphoma	1 (7.1)
Medulloblastoma	2 (14.3)
Osteosarcoma	1 (7.1)
Ewing sarcoma	1 (7.1)
Weight-to-height *Z* score, median (IQR)	0 (–1.5 to 1.6)
Free-living step count during the validation week, median (range)	
Steps per h measured via the ActiGraph	188 (0 to 4725)
Steps per h measured via the smartwatch	33 (0 to 4799)
Steps per d measured via the ActiGraph	5697 (1598 to 17,436)
Steps per d measured via the smartwatch	2213 (12 to 12,425)

During the 1-week study period in free-living conditions, the median number of steps per hour was 188 (range 0-4725) as measured using the ActiGraph and 33 (range 0-4799) as measured using the smartwatch. The median steps per day measured using the ActiGraph and the smartwatch were 5697 (range 1598-17,436) and 2213 (range 12-12,425), respectively.

Results from the linear mixed model revealed a significant difference in steps per hour between the ActiGraph and the smartwatch (P<.001). The mean difference (ActiGraph – smartwatch) was 174 steps per hour. The 95% LoA ranged from −283 to 633 steps per hour. Figures 2 and 3 visualize the repeated measurements and agreement for steps per hour.

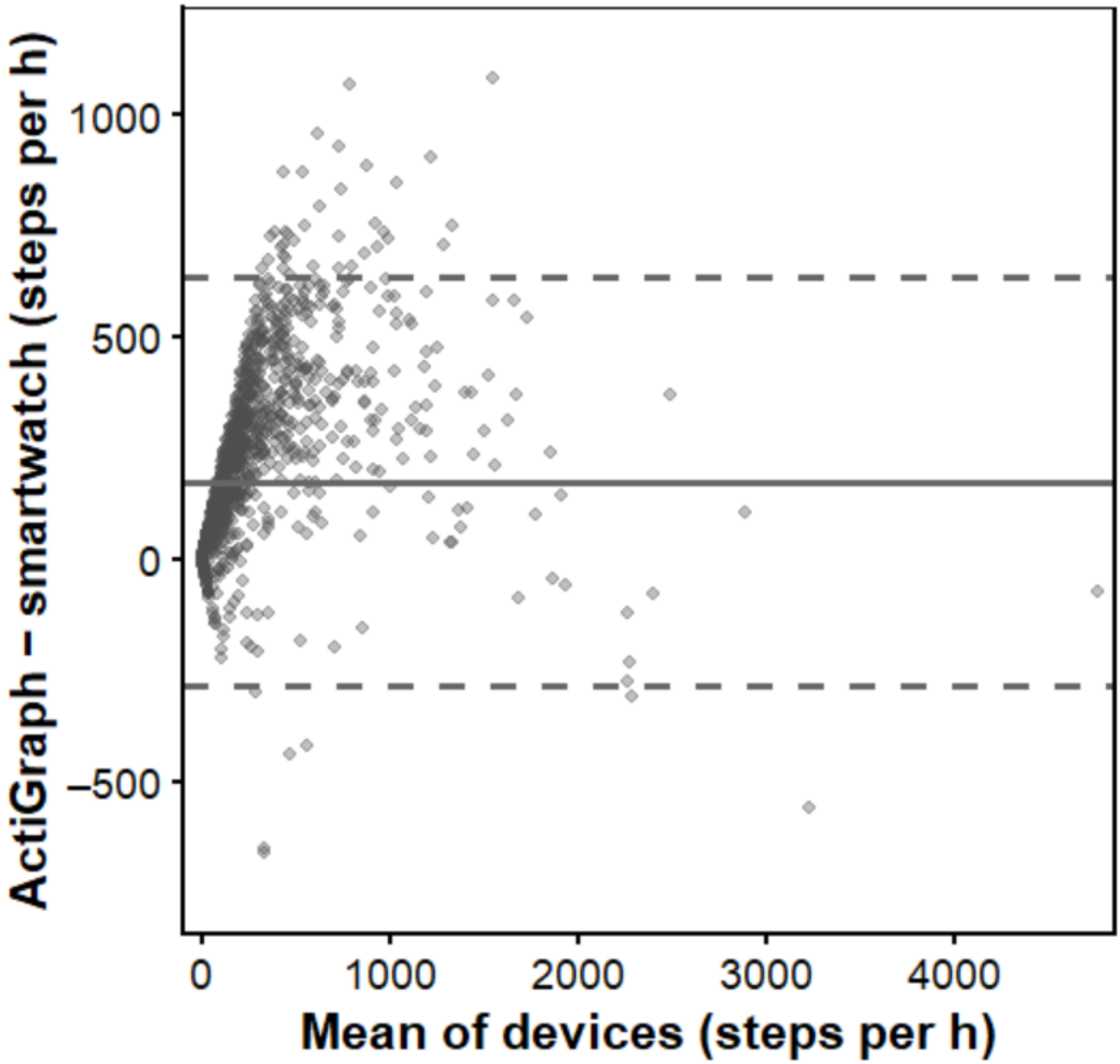
XSL•FO
RenderX

Figure 2. Scatterplot comparing smartwatch and ActiGraph steps per hour. The dashed line represents the line of equality ($y=x$).



XSL•FO
RenderX

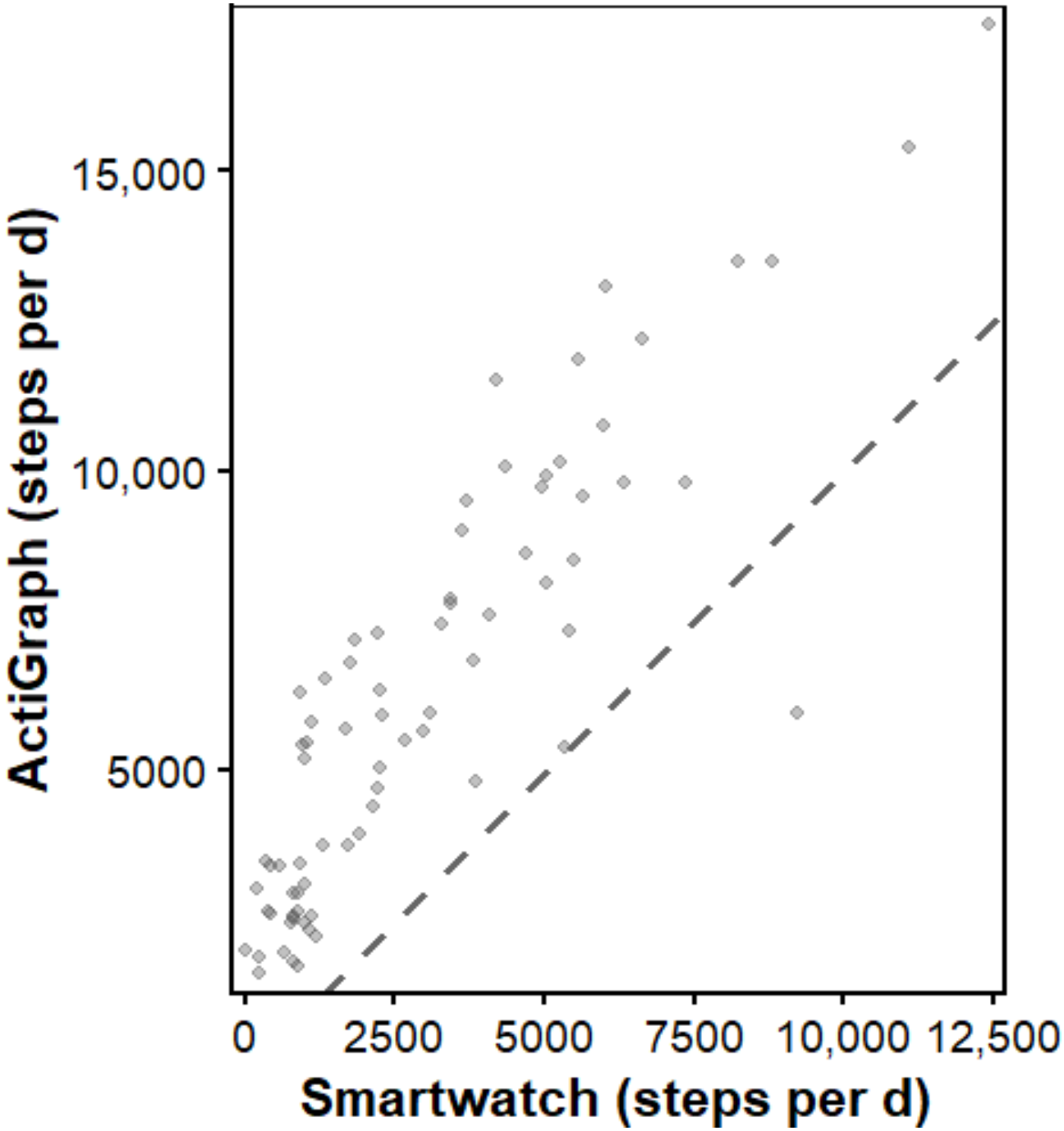
Figure 3. Bland-Altman plot for steps per hour comparing the smartwatch and ActiGraph. The solid line indicates the mean difference; the dashed lines represent the 95% limits of agreement.



Similarly, for steps per day, a significant difference between devices was observed (P<.001). The mean difference was 3154 steps per day, with 95% LoA ranging from −394 to 6702 steps per day. Figures 4 and 5 illustrate the agreement for steps per day. With increasing average hourly as well as daily step counts, the difference between devices increased, suggesting proportional bias.

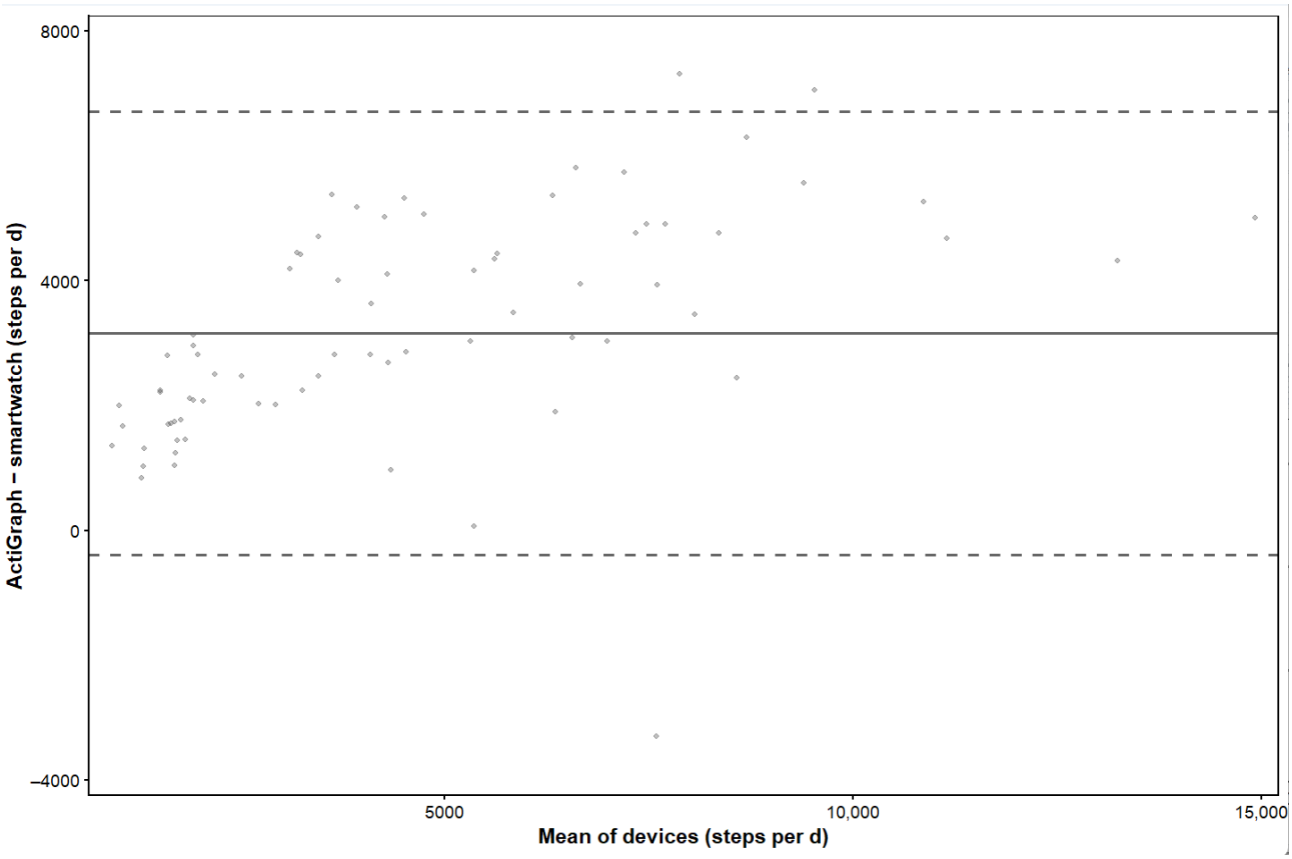
XSL•FO
RenderX

Figure 4. Scatterplot comparing smartwatch and ActiGraph steps per day. The dashed line represents the line of equality ($y=x$).



XSL•FO
RenderX

Figure 5. Bland-Altman plot for steps per day comparing the smartwatch and ActiGraph. The solid line indicates the mean difference; the dashed lines represent the 95% limits of agreement.



Heart Rate Validation

During the study period, 14 children and adolescents wore the smartwatch while being connected to a clinical heart rate monitor under resting conditions (Figure 1). In total, 42.9% (n=6) of them were female, with a median age of 12.5 (IQR 9.0-16.0) years and a weight-to-height *Z* score of 0.1 (IQR −1.2 to 1.1; Table 2).

Table . Patient characteristics and resting heart rate data during the heart rate validation study (n=14).

Characteristic	Values
Sex, n (%)	
Female	6 (42.9)
Male	8 (57.1)
Age (y), median (IQR)	12.5 (9.0 to 16.0)
Time since diagnosis (d), median (IQR)	95.0 (74.8 to 140.8)
Diagnosis, n (%)	
Acute lymphoblastic leukemia	9 (64.3)
Lymphoblastic lymphoma	3 (21.4)
Osteosarcoma	1 (7.1)
Ewing sarcoma	1 (7.1)
Weight-to-height *Z* score, median (IQR)	0.1 (−1.2 to 1.1)
Resting heart rate measurements	
Duration of heart rate measurement period, range	6 min to 3 h, 43 min
Heart rate measured via the heart rate monitor (bpm[a]), median (range)	83 (47 to 115)
Heart rate measured via the smartwatch (bpm), median (range)	85 (46 to 113)

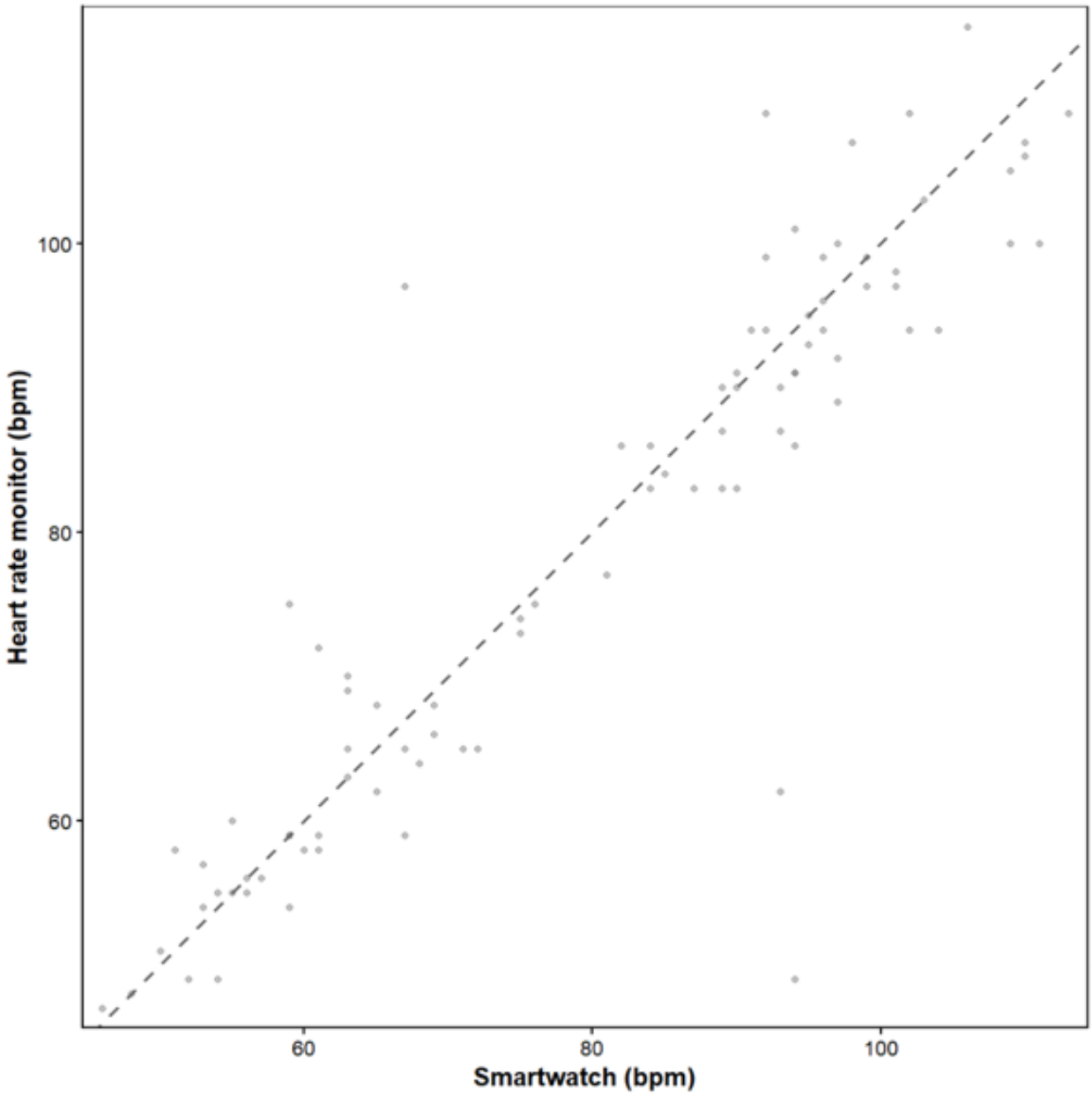
[a]bpm: beats per minute.

XSL•FO
RenderX

During the resting measurement period at the day care treatment ward (range 6 minutes-3 hours 43 minutes), the median heart rate measured using the clinical monitor and the smartwatch was 83 (range 47-115) bpm and 85 (range 46-113) bpm, respectively.

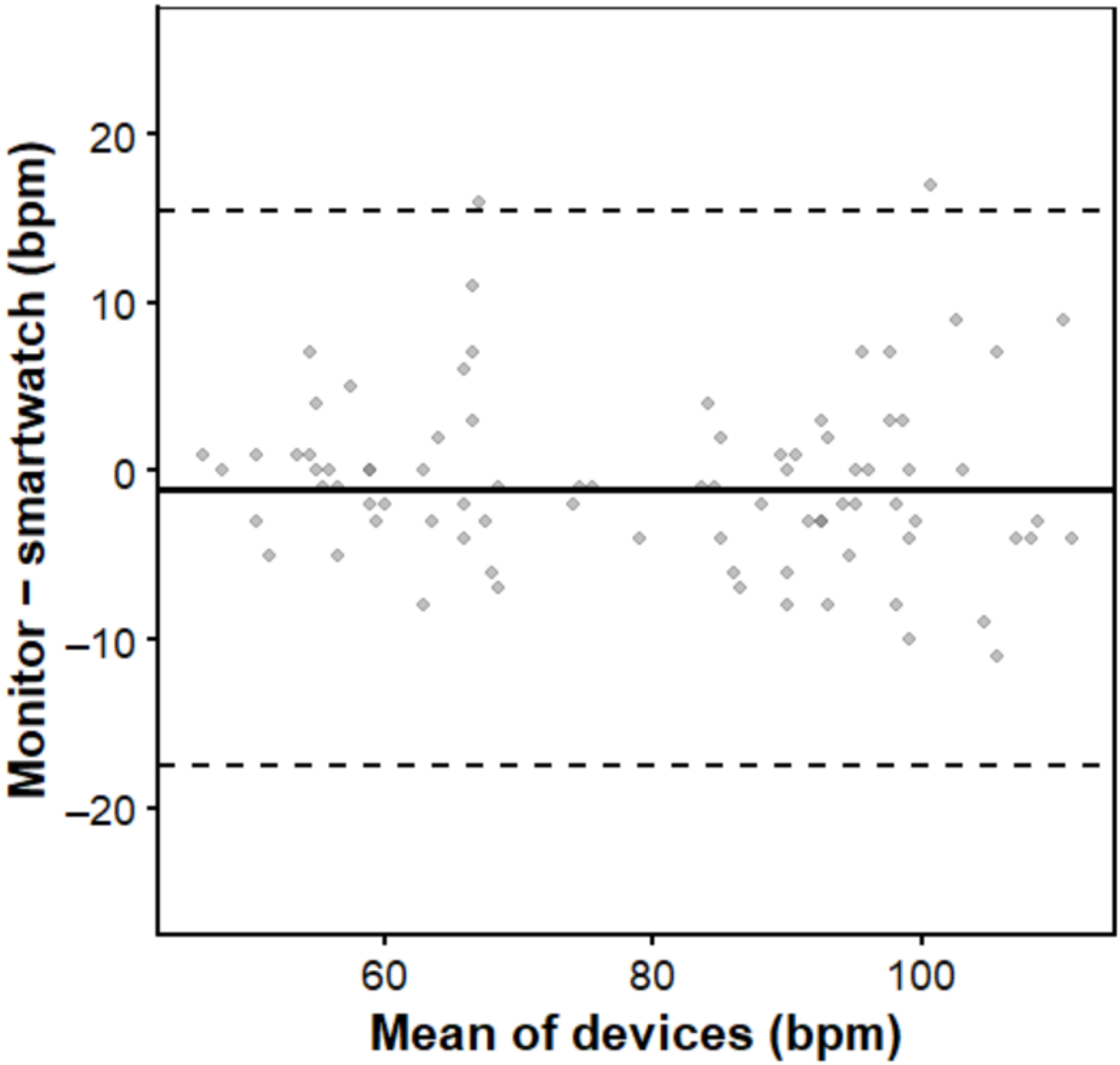
The linear mixed model analysis revealed no statistically significant difference between devices (P=.24). The mean difference (monitor – smartwatch) was −1.07 bpm. The 95% LoA ranged from −17.55 to 15.41 bpm. Figures 6 and 7 illustrate the agreement between devices for heart rate per minute.

Figure 6. Scatterplot comparing smartwatch and clinical heart rate monitor beats per minute (bpm). The dashed line represents the line of equality (y=x).



XSL•FO
RenderX

Figure 7. Bland-Altman plot for heart rate beats per minute (bpm) comparing the smartwatch and clinical heart rate monitor. The solid line indicates the mean difference; the dashed lines represent the 95% limits of agreement.



Discussion

Principal Findings

In this validation study, the Withings smartwatch systematically recorded fewer steps than the wrist-worn ActiGraph under free-living conditions in children and adolescents undergoing cancer treatment. In contrast, resting heart rate measurements showed minimal systematic bias, although individual variability remained considerable.

The smartwatch consistently recorded fewer steps than the ActiGraph, and the magnitude of this discrepancy increased with higher activity levels, indicating proportional bias. This suggests that the difference between devices was not constant across the activity spectrum but became more pronounced at higher step counts. As a result, more active children may be disproportionately affected by underestimation, and step counts derived from the smartwatch may not accurately reflect relative differences between children. In addition, the wide LoA observed in this study indicate substantial variability at the individual level. Together, these findings reinforce that smartwatch-derived step counts should not be considered interchangeable with those obtained from research-grade accelerometers, particularly when absolute step counts are used for clinical benchmarking or comparison across studies. A similar pattern has been described in older adults at risk of functional decline, where a Withings smartwatch also underestimated step count compared with a wrist-worn ActiGraph under free-living conditions [32].

Several factors may contribute to these observed differences. Wrist-worn accelerometers are sensitive to nonambulatory arm movements, which may lead to overestimation of step count when worn on the wrist. Conversely, consumer-grade smartwatches may apply proprietary filtering algorithms to minimize false-positive step detection, potentially resulting in underestimation of low-intensity or irregular movement patterns

XSL•FO
RenderX

[32,33]. In addition, differences in step detection thresholds and algorithm design between devices likely contribute to systematic discrepancies. Importantly, as no criterion gold standard for step count under free-living conditions was available, the observed discrepancy should be interpreted as a difference between device algorithms and placements rather than definitive measurement error of one device alone.

In the broader literature, consumer activity trackers have often shown good associations with research-grade accelerometers for daily step counts in pediatric populations [34-36]. However, studies have also reported systematic differences in absolute step counts, particularly at higher activity intensities [34,36,37]. For example, Fitbit devices have been found to provide comparable step estimates at low intensities but to overestimate steps during moderate to vigorous physical activity [34]. Children with cancer may exhibit altered gait patterns, intermittent low-intensity activity bouts, and treatment-related movement alterations, conditions under which wrist-based step detection algorithms may be particularly challenged. Our findings indicate that device-related discrepancies were also present in children and adolescents undergoing cancer treatment, highlighting that associations do not necessarily imply agreement in absolute step values.

In contrast to step count, heart rate measurements obtained under resting clinical conditions via the smartwatch showed minimal systematic bias compared with the reference monitor. The mean difference was small (approximately −1 bpm), indicating that group average resting heart rate values were similar between devices. However, LoA ranging from approximately −18 to 15 bpm indicate variability at the individual level, which should be considered when interpreting individual heart rate measurements. These findings are consistent with those of previous validation studies in adult populations, including patients scheduled for elective surgery and healthy adults, in which Withings smartwatches demonstrated good agreement with chest strap sensors or clinical monitors under low-intensity or resting conditions [24,38]. However, it should be noted that heart rate was assessed only under relatively stable, low-intensity conditions and within a limited heart rate range in our study. The validity of smartwatch-derived heart rate measurements during higher-intensity activities or in free-living conditions remains uncertain.

From a clinical perspective, wearable monitors may have several potential applications in pediatric oncology. Continuous monitoring can provide insights into activity quantity (eg, daily step count) as well as patterns over time (eg, diurnal variation, day-to-day fluctuations, and recovery after hospitalization), which may be difficult to capture through periodic assessments. However, the observed systematic and proportional differences in step count between devices are particularly relevant in pediatric oncology, where subtle changes in activity may reflect recovery, treatment tolerance, or emerging complications. If higher activity levels are systematically underestimated, improvements during rehabilitation or recovery phases may be less visible when relying solely on smartwatch-derived step counts. In this context, inaccurate step feedback may influence motivation and self-perception of performance. Underestimation at higher activity levels could be particularly discouraging for children who actively engage in rehabilitation or physical activity goals [18]. Therefore, interpretation of wearable-derived step data should focus primarily on intraindividual trends rather than absolute comparisons with normative thresholds or research-grade accelerometers. In this context, within-person trends over time using the same device may be more informative than absolute comparisons with predefined step thresholds or research-grade accelerometers. Such longitudinal information may support early recognition of functional decline, facilitate individualized counseling by physiotherapists, and help evaluate adherence to physical activity recommendations or rehabilitation plans. In addition, feedback from wearable devices may support motivation and engagement in an active lifestyle for some children, although this may also be confronting for others and should be used cautiously and in a supportive manner [18].

This validation study has several limitations that need to be taken into account when interpreting the results. First, no criterion gold standard for step count under free-living conditions was available. The ActiGraph was used as a widely accepted research-grade comparator. However, wrist-worn accelerometers have been shown to overestimate step count, particularly under conditions of low-intensity or irregular movement [39]. In addition, differences between devices may reflect variation in sensitivity to wrist movement, step detection thresholds, and algorithm design. Therefore, the observed discrepancy between devices may represent characteristics of both measurement systems rather than true measurement error. Second, the exact threshold values and proprietary algorithms used by both devices to define a step are not publicly available. Differences in filtering strategies may partly explain both the systematic underestimation and the proportional bias observed in the smartwatch. Third, nonwear time for the ActiGraph was determined using the Troiano algorithm [27]. In children undergoing cancer treatment, prolonged periods of very low–intensity activity are common [1-3], which makes it challenging to distinguish inactivity from nonwear using standard algorithms. This may have influenced step count comparisons, particularly at lower activity levels, where misclassification of wear time may occur. Fourth, wrist side (dominant vs nondominant) was not systematically recorded. Children and adolescents were allowed to choose the wrist to optimize comfort and adherence, which reflects real-world implementation but may have introduced additional variability in step estimates. Although dominant wrist placement has been shown to influence step counts [40], previously reported interwrist differences are generally smaller than the discrepancy observed in our study. Nevertheless, we cannot fully exclude some contribution of limb dominance. Fifth, heart rate was validated only under resting clinical conditions within a limited heart rate range and, therefore, cannot be generalized to higher-intensity activities or free-living conditions where motion artifacts and greater variability may affect smartwatch performance. In addition, disease- and treatment-related factors (eg, treatment phase, hospitalization, fatigue, neuropathy, reduced gait speed, and intermittent activity bouts) may alter movement patterns and physiological responses. Our sample size did not allow for stratified analyses by clinical characteristics. Hence, device agreement may differ across subgroups. Finally, the exploratory nature of this study and the

XSL•FO
RenderX

relatively small sample size limit the precision of the estimated LoA and reduce generalizability.

Future research should aim to further elucidate the mechanisms underlying the observed disagreement in step count between devices. Inclusion of additional criterion measures, such as direct observation or controlled laboratory-based step assessments, may help determine whether discrepancies arise primarily from overestimation by wrist-worn accelerometers, underestimation by smartwatch algorithms, or a combination of both. While waist-worn accelerometers may provide complementary insights into the influence of upper-limb movement on step detection, their feasibility for long-term monitoring in pediatric oncology is limited, and therefore, they are primarily relevant as methodological reference tools rather than implementation devices. Given the observed proportional bias, future studies should also examine whether calibration approaches or device-specific correction models could improve agreement across different activity intensities. In addition, validation of smartwatch-derived heart rate measurements under free-living conditions and across a broader physiological range is warranted. Larger studies across different treatment phases are needed to explore whether disease- and treatment-related characteristics influence device performance. Furthermore, evaluating test-retest reliability and responsiveness over time will be essential to determine whether smartwatch-derived measures are suitable for longitudinal monitoring of physical activity trajectories during and after treatment.

Conclusions

In this exploratory validation study, the Withings smartwatch showed minimal systematic bias for resting heart rate under controlled clinical conditions, although individual variability was considerable. In contrast, smartwatch-derived step counts were consistently lower than those obtained using a research-grade accelerometer and showed increasing disagreement at higher activity levels. Given the observed systematic and proportional discrepancies, use of smartwatch-derived step counts for clinical benchmarking or goal setting in children undergoing cancer treatment cannot currently be recommended.

Acknowledgments

Responsibility for the final manuscript lies entirely with the authors. Generative artificial intelligence tools were not used in the preparation of this manuscript and do not bear responsibility for the final outcomes.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Aznar S, Webster AL, San Juan AF, et al. Physical activity during treatment in children with leukemia: a pilot study. Appl Physiol Nutr Metab 2006 Aug;31(4):407-413. [doi: 10.1139/h06-014] [Medline: 16900230]
2. Fuemmeler BF, Pendzich MK, Clark K, et al. Diet, physical activity, and body composition changes during the first year of treatment for childhood acute leukemia and lymphoma. J Pediatr Hematol Oncol 2013 Aug;35(6):437-443. [doi: 10.1097/MPH.0b013e318279cd3e] [Medline: 23211695]
3. Winter C, Müller C, Brandes M, et al. Level of activity in children undergoing cancer treatment. Pediatr Blood Cancer 2009 Sep;53(3):438-443. [doi: 10.1002/pbc.22055] [Medline: 19415742]
4. Ness KK, Kaste SC, Zhu L, et al. Skeletal, neuromuscular and fitness impairments among children with newly diagnosed acute lymphoblastic leukemia. Leuk Lymphoma 2015 Apr;56(4):1004-1011. [doi: 10.3109/10428194.2014.944519] [Medline: 25030039]
5. Kowaluk A, Woźniewski M, Malicka I. Physical activity and quality of life of healthy children and patients with hematological cancers. Int J Environ Res Public Health 2019 Aug 3;16(15):2776. [doi: 10.3390/ijerph16152776] [Medline: 31382561]
6. Braam KI, van der Torre P, Takken T, Veening MA, van Dulmen-den Broeder E, Kaspers GJ. Physical exercise training interventions for children and young adults during and after treatment for childhood cancer. Cochrane Database Syst Rev 2013 Apr 30(4):CD008796. [doi: 10.1002/14651858.CD008796.pub2] [Medline: 23633361]
7. Caru M, Dandekar S, Schmitz KH. Physical activity and childhood cancer: present status and future directions. Am J Lifestyle Med 2025 Aug 20:15598276251368342. [doi: 10.1177/15598276251368342] [Medline: 40852085]
8. Fiuza-Luces C, Padilla JR, Soares-Miranda L, et al. Exercise intervention in pediatric patients with solid tumors: the Physical Activity in Pediatric Cancer Trial. Med Sci Sports Exerc 2017 Feb;49(2):223-230. [doi: 10.1249/MSS.0000000000001094] [Medline: 27631396]

XSL•FO
RenderX

9. Saultier P, Vallet C, Sotteau F, et al. A randomized trial of physical activity in children and adolescents with cancer. Cancers (Basel) 2021 Jan 2;13(1):121. [doi: 10.3390/cancers13010121] [Medline: 33401713]
10. Wurz A, McLaughlin E, Chamorro Viña C, et al. Advancing the field of pediatric exercise oncology: research and innovation needs. Curr Oncol 2021 Jan 20;28(1):619-629. [doi: 10.3390/curroncol28010061] [Medline: 33498499]
11. Collier L, Grimshaw SL, Stolper J, et al. Scoping review of the utilization of wearable devices in pediatric and young adult oncology. NPJ Digit Med 2025 Jul 31;8(1):488. [doi: 10.1038/s41746-025-01842-5] [Medline: 40745222]
12. Crane TE, Skiba MB, Miller A, Garcia DO, Thomson CA. Development and evaluation of an accelerometer-based protocol for measuring physical activity levels in cancer survivors: development and usability study. JMIR Mhealth Uhealth 2020 Sep 24;8(9):e18491. [doi: 10.2196/18491] [Medline: 32969828]
13. Williamson Lewis R, Howell KE, Effinger KE, et al. Feasibility of Fitbit use in adolescent survivors of pediatric cancer: who consistently uses it and for how long? J Adolesc Young Adult Oncol 2023 Aug;12(4):529-536. [doi: 10.1089/jayao.2022.0071] [Medline: 36178482]
14. Doherty C, Baldwin M, Keogh A, Caulfield B, Argent R. Keeping pace with wearables: a living umbrella review of systematic reviews evaluating the accuracy of consumer wearable technologies in health measurement. Sports Med 2024 Nov;54(11):2907-2926. [doi: 10.1007/s40279-024-02077-2] [Medline: 39080098]
15. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020 Sep 8;8(9):e18694. [doi: 10.2196/18694] [Medline: 32897239]
16. Giurgiu M, Kolb S, Nigg C, et al. Assessment of 24-hour physical behaviour in children and adolescents via wearables: a systematic review of free-living validation studies. BMJ Open Sport Exerc Med 2022;8(2):e001267. [doi: 10.1136/bmjsem-2021-001267] [Medline: 35646389]
17. Lanckriet S, Brissieux E, Borys JM, et al. Assessment of the impact of a sports kit on physical activity in children 8 to 11 years of age. J Phys Act Res 2017;2(1):50-60. [doi: 10.12691/jpar-2-1-9]
18. den Hartog E, Tissing WJ, Bon SB, van der Torre P, Verwaaijen EJ. Feasibility, barriers, and facilitators of long-term physical activity tracking during treatment: interview study among childhood cancer patients. JMIR Pediatr Parent 2025 Aug 27;8:e75322. [doi: 10.2196/75322] [Medline: 40864896]
19. Kruizinga MD, Heide NV, Moll A, et al. Towards remote monitoring in pediatric care and clinical trials-tolerability, repeatability and reference values of candidate digital endpoints derived from physical activity, heart rate and sleep in healthy children. PLoS One 2021;16(1):e0244877. [doi: 10.1371/journal.pone.0244877] [Medline: 33411722]
20. Delrieu L, Pialoux V, Pérol O, et al. Feasibility and health benefits of an individualized physical activity intervention in women with metastatic breast cancer: intervention study. JMIR Mhealth Uhealth 2020 Jan 28;8(1):e12306. [doi: 10.2196/12306] [Medline: 32012082]
21. Hale TM, Guardigni V, Roitmann E, et al. Middle-aged men with HIV have diminished accelerometry-based activity profiles despite similar lab-measured gait speed: pilot study. JMIR Mhealth Uhealth 2019 Feb 1;7(2):e11190. [doi: 10.2196/11190] [Medline: 30707104]
22. Le Hello C, Trombert B, Morel A, Chieh A, Brouard B, Boissier C. Performance analysis of walking of 10,000 regular users of a connected activity tracker. J Med Vasc 2018 Jul;43(4):231-237. [doi: 10.1016/j.jdmv.2018.04.001] [Medline: 29981731]
23. Treskes RW, van Winden LA, van Keulen N, et al. Effect of smartphone-enabled health monitoring devices vs regular follow-up on blood pressure control among patients after myocardial infarction: a randomized clinical trial. JAMA Netw Open 2020 Apr 1;3(4):e202165. [doi: 10.1001/jamanetworkopen.2020.2165] [Medline: 32297946]
24. Helmer P, Hottenrott S, Rodemers P, et al. Accuracy and systematic biases of heart rate measurements by consumer-grade fitness trackers in postoperative patients: prospective clinical trial. J Med Internet Res 2022 Dec 30;24(12):e42359. [doi: 10.2196/42359] [Medline: 36583938]
25. Vetrovsky T, Siranec M, Marencakova J, et al. Validity of six consumer-level activity monitors for measuring steps in patients with chronic heart failure. PLoS One 2019;14(9):e0222569. [doi: 10.1371/journal.pone.0222569] [Medline: 31518367]
26. Ernstsson J, Svensson B, Liuba P, Weismann CG. Validation of smartwatch electrocardiogram intervals in children compared to standard 12 lead electrocardiograms. Eur J Pediatr 2024 Sep;183(9):3915-3923. [doi: 10.1007/s00431-024-05648-7] [Medline: 38918230]
27. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
28. Van Blarigan EL, Kenfield SA, Tantum L, Cadmus-Bertram LA, Carroll PR, Chan JM. The Fitbit One physical activity tracker in men with prostate cancer: validation study. JMIR Cancer 2017 Apr 18;3(1):e5. [doi: 10.2196/cancer.6935] [Medline: 28420602]
29. John D, Freedson P. ActiGraph and Actical physical activity monitors: a peek under the hood. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S86-S89. [doi: 10.1249/MSS.0b013e3182399f5e] [Medline: 22157779]
30. de Vries SI, Bakker I, Hopman-Rock M, Hirasing RA, van Mechelen W. Clinimetric review of motion sensors in children and adolescents. J Clin Epidemiol 2006 Jul;59(7):670-680. [doi: 10.1016/j.jclinepi.2005.11.020] [Medline: 16765269]

XSL•FO
RenderX

31. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat 2007;17(4):571-582. [doi: 10.1080/10543400701329422] [Medline: 17613642]
32. Au-Yeung WT, Kaye JA, Beattie Z. Step count standardization: validation of step counts from the Withings Activite using PiezoRxD and wGT3X-BT. Annu Int Conf IEEE Eng Med Biol Soc 2020 Jul;2020:4608-4611. [doi: 10.1109/EMBC44109.2020.9176511] [Medline: 33019020]
33. Nakagata T, Murakami H, Kawakami R, et al. Step-count outcomes of 13 different activity trackers: results from laboratory and free-living experiments. Gait Posture 2022 Oct;98:24-33. [doi: 10.1016/j.gaitpost.2022.08.004] [Medline: 36030707]
34. Hamari L, Kullberg T, Ruohonen J, et al. Physical activity among children: objective measurements using Fitbit One® and ActiGraph. BMC Res Notes 2017 Apr 20;10(1):161. [doi: 10.1186/s13104-017-2476-1] [Medline: 28427441]
35. Kang S, Kim Y, Byun W, Suk J, Lee JM. Comparison of a wearable tracker with ActiGraph for classifying physical activity intensity and heart rate in children. Int J Environ Res Public Health 2019 Jul 25;16(15):2663. [doi: 10.3390/ijerph16152663] [Medline: 31349667]
36. Voss C, Gardner RF, Dean PH, Harris KC. Validity of commercial activity trackers in children with congenital heart disease. Can J Cardiol 2017 Jun;33(6):799-805. [doi: 10.1016/j.cjca.2016.11.024] [Medline: 28347581]
37. Schmidt MD, Rathbun SL, Chu Z, et al. Agreement between Fitbit and ActiGraph estimates of physical activity in young children. Meas Phys Educ Exerc Sci 2023;27(2):171-180. [doi: 10.1080/1091367x.2022.2125319] [Medline: 37377882]
38. Giggins OM, Doyle J, Smith S, Crabtree DR, Fraser M. Measurement of heart rate using the Withings ScanWatch device during free-living activities: validation study. JMIR Form Res 2022 Sep 1;6(9):e34280. [doi: 10.2196/34280] [Medline: 36048505]
39. Höchsmann C, Knaier R, Infanger D, Schmidt-Trucksäss A. Validity of smartphones and activity trackers to measure steps in a free-living setting over three consecutive days. Physiol Meas 2020 Jan 30;41(1):015001. [doi: 10.1088/1361-6579/ab635f] [Medline: 31851949]
40. Park S, Toth LP, Hibbing PR, et al. Dominant vs. non-dominant wrist placement of activity monitors: impact on steps per day. J Meas Phys Behav 2019;2(2):118-123. [doi: 10.1123/jmpb.2018-0060]

Abbreviations

bpm: beats per minute
LoA: limits of agreement

Edited by M Balcarras; submitted 19.May.2025; peer-reviewed by R Gil-Gutiérrez, SL Coven, T Langer; revised version received 10.Mar.2026; accepted 30.Mar.2026; published 10.Jun.2026.

Please cite as:
den Hartog E, Fiocco M, van der Torre P, Tissing WJE, Verwaaijen EJ
Monitoring Steps and Heart Rate Using a Withings Smartwatch in Children and Adolescents With Cancer: Validation Study
JMIR Cancer 2026;12:e77766
URL: https://cancer.jmir.org/2026/1/e77766
doi:10.2196/77766

©Emma den Hartog, Marta Fiocco, Patrick van der Torre, Wim J E Tissing, Emma J Verwaaijen. Originally published in JMIR Cancer (https://cancer.jmir.org), 10.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Online Intervention on Lung Cancer Screening Among High-Risk Individuals: Pilot Intervention Study

Fang Lei[1], MPH, MSN, PhD; Feifei Huang[2], PhD; Yufang Guo[3], PhD; Sun S Kim[4], PhD; Yang Liu[5], PhD

[1]School of Nursing, University of Minnesota, 308 Harvard St SE, Minneapolis, MN, United States

[2]School of Nursing, Fujian Medical University, Fuzhou, Fujian, China

[3]School of Nursing and Rehabilitation, Shandong University, Jinan, Shandong, China

[4]Department of Nursing, Donna M. and Robert J. Manning College of Nursing & Health Sciences, University of Massachusetts Boston, Boston, MA, United States

[5]Department of Computer Information Systems, School of Business, North Carolina Central University, NC, NC, United States

Corresponding Author:
Fang Lei, MPH, MSN, PhD
School of Nursing, University of Minnesota, 308 Harvard St SE, Minneapolis, MN, United States

Abstract

Background: Lung cancer remains the leading cause of cancer deaths in the United States; however, uptake of lung cancer screening (LCS) with low-dose computed tomography (LDCT) among eligible individuals remains low. Evidence suggests that limited knowledge, stigma, and false health beliefs contribute to the underuse of LDCT screening.

Objective: This pilot study aimed to examine an online educational intervention designed to improve knowledge, attitudes, health beliefs, behavioral intentions, perceived importance, and confidence related to LCS among high-risk individuals.

Methods: A single-group preintervention and postintervention design was used. High-risk individuals who smoke, defined according to the US Preventive Services Task Force criteria, completed baseline questionnaires followed by 5 self-directed online educational modules delivered through Research Electronic Data Capture (REDCap). Postintervention questionnaires assessed changes in lung cancer and screening knowledge, lung cancer stigma, health beliefs based on the health belief model and precaution adoption process model, and intentions, perceived importance, and confidence regarding LDCT screening. LCS uptake was assessed via follow-up email 3 months after the intervention. Data were analyzed using descriptive statistics and paired-samples two-tailed *t* tests.

Results: A total of 25 participants completed the intervention. Significant improvements were observed across all major study outcomes. Knowledge scores increased markedly (score=3.76-8.60; $P<.001$), while lung cancer stigma decreased (score=25.52-19.16; $P<.001$). Health belief model constructs showed significant improvements, including perceived susceptibility, perceived benefits, cues to action, and self-efficacy, alongside reductions in perceived barriers and perceived severity (all $P<.001$). Self-reported intentions, perceived importance, and confidence related to obtaining LDCT screening increased significantly. Of the 22 (88%) participants who completed the 3-month follow-up, 13 (59.1%) reported obtaining LDCT screening. Participant satisfaction with the intervention was high, with a mean score of 18.32 (SD 2.33) out of 20.

Conclusions: Findings from this pilot study support the feasibility, acceptability, and preliminary efficacy of an online educational intervention created to promote LCS among high-risk individuals. The intervention improved knowledge; reduced stigma; positively influenced health beliefs; and increased screening intentions, perceived importance, confidence, and uptake. Results provide a foundation for a larger-scale study and suggest that online educational platforms may be an effective strategy to reach geographically diverse high-risk populations and promote LDCT screening.

(JMIR Cancer 2026;12:e89823) doi:10.2196/89823

KEYWORDS

education; intervention; lung cancer screening; online; smoking; artificial intelligence; AI

Introduction

Lung Cancer Incidence, Mortality, and Survival Rates

Lung cancer is the leading cause of cancer deaths and the second most commonly diagnosed cancer in both male and female individuals in the United States [1]. In 2026, it is estimated that 229,410 people will be diagnosed with lung cancer, and 124,990 people will die from lung cancer [1]. Patients with lung cancer have one of the lowest 5-year survival rates (18.6%) compared to other types of cancers in the United States, such as colorectal (64.5%), breast (89.6%), and prostate (98.2%) cancers [2]. Only 16% of lung cancers are diagnosed at a localized stage, for which the 5-year survival rate is 55%. As the length of delay

XSL•FO
RenderX

before diagnosis increases, the 5-year survival rate for patients with lung cancer significantly decreases. For example, when lung cancer is diagnosed at an advanced stage (stage IV), the survival rate drops to 4%. Additionally, more than half of people with lung cancer die within 1 year of being diagnosed [2].

Primary and Secondary Preventions for Lung Cancer

Tobacco use is the most important risk factor of lung cancer, which contributes to 80% of lung cancer deaths in the United States [3]. Other risk factors of lung cancer include exposure to radon gas, occupational or environmental exposure to secondhand smoke, asbestos, certain metals, some organic chemicals, radiation, air pollution, and diesel exhaust. The primary approach to prevent lung cancer is smoking cessation, which has been proven to effectively decrease the incidence rates of lung cancer among male and female individuals [3].

Lung cancer screening (LCS) with low-dose computed tomography (LDCT) is an effective secondary prevention method [4]. It has been shown to reduce the mortality of lung cancer by 20%, compared to the standard chest x-ray, among high-risk individuals [4]. Screening for individuals at high risk for lung cancer (individuals aged between 50 and 80 years, who have smoked at least 20 pack-years, and currently smoking or having quit smoking within the past 15 years) [5] has the potential to improve lung cancer survival rates by detecting nodules at an earlier stage when it is more likely to be curable. It is reported that about 8 million Americans qualify as high risk for lung cancer and are recommended to receive annual screening with LDCT scans [5]. If half of these high-risk individuals are screened, more than 12,000 lung cancer deaths could be prevented annually [6].

Guidelines and Uptake Rate of LCS

Since 2013, the US Preventive Services Task Force and other organizations (eg, American Cancer Society, American College of Chest Physicians, American Society of Clinical Oncology, and American Lung Association) have issued guidelines for early detection of lung cancer with yearly LDCT among high-risk populations [7]. It was covered both by the private and public health insurances for the high-risk population [8].

Although the supportive landscape for screening lung cancer among high-risk populations has improved, uptake rates of LCS with LDCT remain low [9]. The percentage of the eligible population who had received LCS with LDCT increased from 3.3% in 2010 to 3.9% in 2015 and reached 18.2% in 2022 among the US population [10,11].

Barriers to LCS Uptake

Barriers to LCS for eligible individuals have been reported to include a lack of knowledge, practical challenges, financial constraints, psychological obstacles such as worry and anxiety, blame and stigma, fear of cancer, and fatalism [12]. Additionally, there is confusion surrounding LCS and a general distrust of the medical system. High-risk individuals seem to have limited knowledge about LCS. This lack of knowledge may stem from the little information they received regarding the screening process. Although LCS with LDCT is covered by both private and public health insurance for high-risk populations, awareness of this coverage appears insufficient, given the financial barriers frequently reported (eg, cost of the screening, insufficient authorization of health insurance reimbursement by insurance companies, and misunderstanding-associated out-of-pocket cost). In addition, stigma and fatalism toward the disease (eg, the perception that lung cancer is not treatable) could also impede high-risk individuals from screening for lung cancer [12].

Interventions to Increase LCS Uptake and Limitations

Multiple interventions have been implemented to increase LCS uptake in high-risk populations. Evidence-based strategies include shared decision-making tools and decision aids designed to support informed screening choices and improve adherence to LCS guidelines [13]. In addition, a range of patient-directed interventions have been evaluated, including educational videos and information films [14] and clinical outreach and referrals [15]. However, prior studies vary substantially in design, intensity, and delivery context, with many interventions implemented within clinical settings or as part of health care system–based programs. Furthermore, most reports in this area are limited by focus primarily on describing barriers rather than evaluating intervention effectiveness. These limitations highlight the need for rigorously designed, scalable interventions that can be delivered outside of clinical environments and evaluated using standardized outcome measures.

In recent years, several rigorously developed digital interventions have been designed to support LCS decision-making. LungTalk, developed by Carter-Harris et al [16], is a theory-driven, web-based intervention that addresses lung cancer stigma, knowledge, and health beliefs to promote informed screening decisions. Similarly, mPATH-Lung (mobile Patient Technology for Health-Lung), developed by Miller et al [17], is a personalized, web-based decision aid that has been implemented in clinical settings to facilitate shared decision-making for LCS. In addition, LCS decision support tools developed by Volk et al [18] at MD Anderson Cancer Center provide evidence-based, patient-centered resources to guide screening decisions. These interventions demonstrate that digital approaches to LCS are both feasible and effective. However, despite these advances, many existing digital interventions are designed primarily as decision aids within clinical workflows or target-specific populations (eg, current smokers engaged in health care systems). In contrast, fewer interventions have focused on scalable, fully remote, multimodule educational programs that can be delivered outside of clinical encounters and are tailored to high-risk populations with limited access to LCS resources.

This study addresses the gap by developing and pilot-testing a structured, 5-module digital intervention integrating educational content, health belief modification, and culturally relevant messaging. Unlike existing intervention programs that primarily focus on facilitating a single screening decision, this study emphasizes a broader continuum of behavior change, including knowledge enhancement, belief modification, and intention formation. Additionally, the asynchronous and low-cost delivery format enhances scalability and accessibility, particularly for

XSL•FO
RenderX

populations who may face barriers to clinic-based decision-making tools.

Aims

This study aims to (1) test the feasibility of an online intervention program on increasing LCS uptakes among high-risk individuals and (2) assess the intervention's effects on changing participants' knowledge of lung cancer and LCS, attitudes toward lung cancer stigma, health beliefs regarding LCS, behavioral intentions, perceived importance, and confidence in obtaining LCS when compared to baseline measurements.

Significance

This study implemented an online intervention aimed at increasing LCS uptake among high-risk individuals. Promoting LCS projects helps high-risk individuals better understand their risk of lung cancer, acquire knowledge about lung cancer and screening procedures, and enhance their awareness of the importance of LCS.

Methods

Design

A single-arm preintervention and postintervention study design was used in this study. A flow diagram about the study process is shown in Figure 1. Paired-comparison data before and after the intervention were obtained and analyzed.

XSL•FO
RenderX

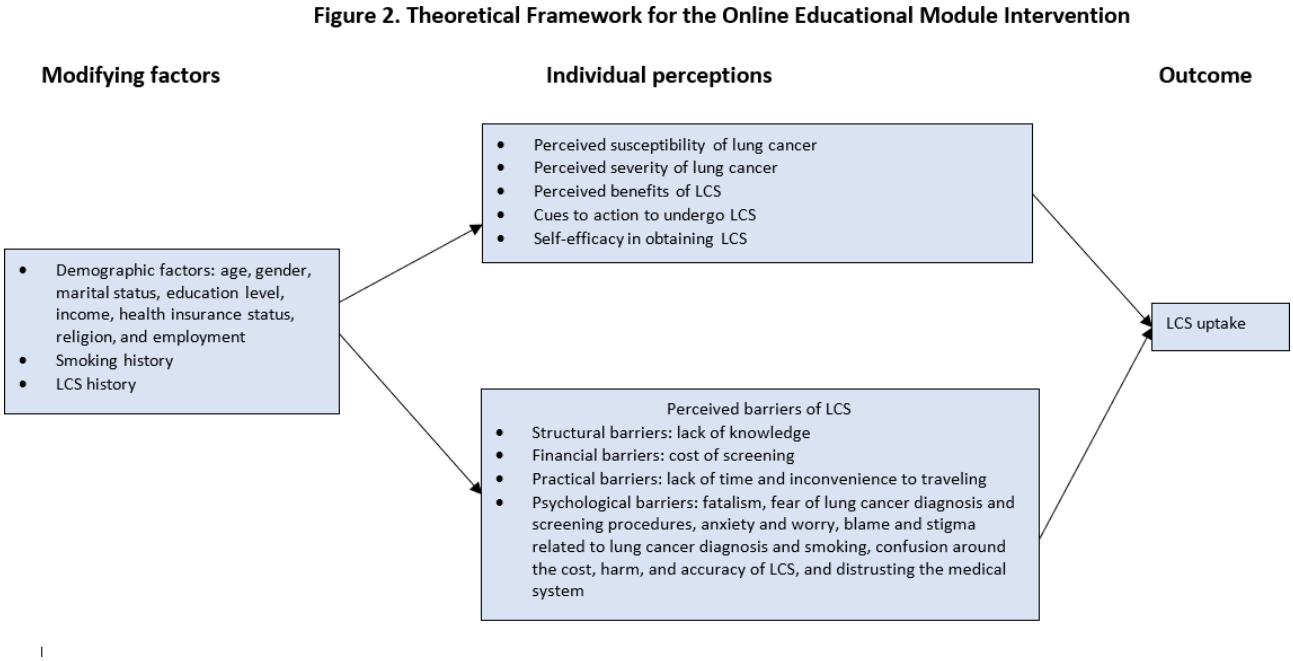
Figure 1. Study flowchart.

XSL•FO
RenderX

Theoretical Framework

The health belief model (HBM) provided a theoretical framework for the study (Figure 2) [19]. Key concepts of the HBM are perceived susceptibility, perceived severity, perceived benefits, perceived barriers, self-efficacy, and cues to action [19]. In addition, this study was conceptually informed by the integrated LCS participation model developed by Carter-Harris et al [20], which incorporates HBM constructs alongside lung cancer stigma and the precaution adoption process model (PAPM). This integrated framework emphasizes that LCS participation is not solely driven by health beliefs but also shaped by psychosocial factors and individuals' stage of readiness to act.

Figure 2. Theoretical framework for the online educational module intervention. LCS: lung cancer screening.



Ethical Considerations

The study was reviewed and approved by the institutional review board of the University of Minnesota (institutional review board protocol: STUDY00020835). All study procedures followed the ethical principles outlined in the Belmont Report and adhered to relevant federal and institutional guidelines for research involving human participants. Participants were informed about the purposes of the study, procedures, potential risks, and benefits through an online consent form prior to enrollment. Participants were reminded that their participation was voluntary, and they could withdraw from the study at any time without penalty. As the intervention and data collection took place online, additional measures were taken to ensure confidentiality, including secure data transmission, encrypted storage, and restricted access to identifiable information.

Minimal risks were anticipated, as the intervention involved completing online educational modules and surveys. The consent form explained potential discomfort that might arise from answering questions about smoking behavior and health beliefs. No adverse events were reported during the study, and no protocol deviation was noted from the study. Participants were compensated for their time with a US $40 Amazon gift card upon completion of the intervention, and an additional US $10 Amazon gift card was given to those who completed the 3-month follow-up survey.

All procedures complied with the Health Insurance Portability and Accountability Act as applicable, and the data were deidentified before analysis. Results are presented in an aggregate form to protect participants' privacy.

Setting

This study was conducted online. Interactive educational modules about LCS were created and implemented online through Research Electronic Data Capture (REDCap; Vanderbilt University). Questionnaires used in the study were administered on REDCap before and after the intervention. Follow-up emails were sent to increase the response rate of the questionnaires, and the study team answered any questions that the participants might have.

Sample

The inclusion criteria for participants in this study were as follows: individuals with a smoking history of more than 20 pack-years, currently smoking or having quit smoking in the past 15 years, aged between 50 and 80 years, and able to read or understand English at the sixth-grade level. Exclusion criterion for the participants was having been diagnosed with lung cancer or other diseases that could substantially impact participants' life expectancy. These criteria align with the US Preventive Services Task Force screening recommendation for lung cancer.

Sample Size

The sample size was calculated using the G-Power software [21]. To detect a medium effect size (Cohen d=0.5) with an alpha error probability of .05 and a power of 0.95, a total sample size of more than 210 is required for a formal study when using

XSL•FO
RenderX

a paired sample *t* test to detect differences. Considering data quality and the expected response rate of the questionnaires, an additional 20% increase in the suggested sample size is warranted, resulting in a target of 250 participants for the formal study.

As this study is a pilot study for the larger formal study, extant literature recommends that a pilot study should include approximately 10% of the sample size projected for the main study [21,22]. In addition, researchers typically suggest a sample size of 10 to 30 participants in both survey research and medical field research [23,24]. Therefore, a sample size of 25 participants would be appropriate for this pilot study.

Recruitment Procedure

Webpages were created on REDCap to provide information about the study. Collaborating with the University of Minnesota Clinical and Translational Science Institute, all available social media platforms, including the StudyFinder (University of Minnesota) website, Instagram (Meta Platforms, Inc), and Facebook (Meta Platforms, Inc) groups, were used to recruit the planned number of participants. The study flyer was shared on these social media along with the research team's email address. Once potential participants emailed the study team, they were screened for eligibility by answering 7 screening questions. If participants met the eligibility criteria, they were assigned a code number, and their contact information (eg, email address and cell number) was collected.

Interventions

Preintervention Phase

Eligible participants were provided with a REDCap link for the questionnaires. Participants first read information about the study, reviewed the informed consent form, and then completed all preintervention questionnaires.

Intervention Phase

Participants who completed preintervention questionnaires were asked to review 5 online interactive modules on REDCap (including videos, text, pictures, and interactive activities) about lung cancer and its screening (Table 1).

Table . Lung cancer screening (LCS) online educational modules.

Module	Content
Module 1	Lung cancer epidemiology, etiology, signs, and symptoms
Module 2	Lung cancer treatment and care
Module 3	Lung cancer prevention methods
Module 4	LCS guidelines, benefits, and risks
Module 5	LCS procedures and results interpretation

Postintervention Phase

After reviewing the online educational modules, participants were asked to complete the postintervention questionnaires on REDCap. Participants were followed over 3 months [25] by emails to assess whether they underwent LCS.

Development of Interactive Educational Modules

The interactive educational modules were developed through a 3-phase, user-centered process. First, a needs assessment was conducted through a literature review and individual qualitative interviews with 12 individuals with a heavy smoking history to identify knowledge gaps, misconceptions, preferred learning formats, and appropriate health literacy level. Second, 5 sequential modules were created featuring 3- to 5-minute artificial intelligence–generated video lectures, infographics, interactive quizzes, and reflection questions, with iterative prototyping to refine clarity and engagement and to align content with the HBM and PAPM constructs. Third, a panel of 7 clinical and behavioral experts evaluated accuracy and relevance; revisions were made based on item-level and scale-level Content Validity Index (CVI) results until strong content validity (item-level CVI≥0.78; scale-level CVI≥0.90) was achieved [26]. While this paper reported how the intervention performed during pilot testing, detailed methods for the design and development of these modules have been published elsewhere [27].

Instruments

Sociodemographic Information

Sociodemographic characteristics were measured using the *Sociodemographic Information Questionnaire*, which included 8 questions on participants' age, sex, marital status, level of education, annual family income, race, occupation, and health insurance status.

Intentions, Perceived Importance, and Confidence to Obtain LCS

Participants' intentions, perceived importance, and confidence in undergoing LCS were assessed using *the Smoking and Lung Cancer Screening Questionnaire*, which was developed by the study team.

Knowledge of Lung Cancer and Screening

Participants' knowledge of lung cancer and screening was assessed using 12 questions from the *Lung Cancer and Screening Knowledge Questionnaire* [28]. The questionnaire's total score ranged from 0 to 12.

Attitudes Toward Lung Cancer Stigma

Nine questions about lung cancer stigma were assessed using the shortened version of the *Cataldo Lung Cancer Stigma Scale* [29], which uses a 4-point Likert-style response format. The scale's total score ranged from 9 to 36 [30]. The reliability

XSL•FO
RenderX

coefficient scores of the scale ranged from 0.75 to 0.96, which were in acceptable to excellent ranges [30].

Health Beliefs of Lung Cancer and Screening

Participants' health beliefs regarding lung cancer and its screening were assessed using *the modified LCS Health Belief Scale*. This scale is composed of 57 items grouped into 6 subscales. The subscales for perceived severity, perceived susceptibility, perceived benefits, perceived barriers, and cues to action used 4-point Likert-style responses, with options ranging from strongly disagree (score=1) to strongly agree (score=4). The self-efficacy subscale also used 4-point Likert-style responses with options ranging from not at all confident (score=1) to very confident (score=4). This scale was developed from our preliminary studies, adopting the original *LCS Health Belief Scale* [31], which demonstrated acceptable content validity [32].

LCS Uptake

LCS uptake was assessed during the follow-up period after participants completed the online intervention. Upon enrollment, participants were informed that they would receive a follow-up email approximately 3 months later. This email inquired whether they had completed an LDCT LCS since participating in the study. Participants were instructed to respond with "yes" if they had undergone the screening and "no" if they had not. Only participants who replied to the follow-up email were included in the analysis.

Satisfaction Feedbacks

Participants' feedback on the intervention was collected using the *Satisfaction Feedback Questionnaire* developed by the study team. The questionnaire consisted of 5 items, each measured using a 4-point Likert-style ordinal response scale. The total score for the questionnaire ranged from 5 to 20, with higher scores indicating greater satisfaction with the intervention. A total score above 15 indicated a positive attitude toward the online education modules.

Data Analysis

Data analysis was done using SPSS (version 31.0.0; IBM Corp). Descriptive statistics, including mean and SD, frequency, and percentage, were used to evaluate the characteristics of the participants, participant satisfaction with the intervention, and LCS uptake rates. A paired *t* test was performed to examine differences in the mean scores of participants' behavioral intention, perceived importance, and confidence toward LCS; knowledge of lung cancer and its screening; attitudes toward lung cancer stigma; and health beliefs regarding LCS before and after the intervention. The *P* value was set at .05. Results with a *P* value of <.05 were considered significant. Cohen *d* was calculated for paired comparisons to estimate the magnitude of change and inform future sample size calculations. The distributions of predifference and postdifference scores were examined using skewness, kurtosis, and Q-Q plots. Sensitivity analyses were conducted using nonparametric alternatives (Wilcoxon signed-rank tests), and results were compared for consistency.

Results

Participant Characteristics

A total of 27 participants were initially enrolled in the pilot study, of which 25 completed the intervention, resulting in a retention rate of 92.59%. The mean age of the participants was 60.88 (SD 5.16) years. Most participants were female (n=18, 72%), married (n=18, 72%), and had at least a bachelor's degree (n=19, 76%). The majority reported a household annual income between US $45,000 and US $139,999 (n=19, 76%). Of the 25 participants, 44% (n=11) identified as White, 40% (n=10) as Black or African American, and 16% (n=4) as Hispanic or Latinx. More than half of the participants (n=14, 56%) were not currently employed, and 36% (n=9) did not have health insurance. The majority smoked currently (n=22, 88%), and none had previously undergone LCS. Only 1 (4%) participant reported a family history of lung cancer (Table 2). Average time for the participants completing the modules was 2 (SD 0.23) hours. All the participants completed the modules in a single setting.

XSL•FO
RenderX

Table . Participants' characteristics (N=25).

Variables	Values
Age (years), mean (SD)	60.88 (5.16)
Sex, n (%)	
Female	18 (72)
Male	7 (28)
Marital status, n (%)	
Married	18 (72)
Separated	2 (8)
Divorced	2 (8)
Widowed	3 (12)
Education level, n (%)	
Some college	6 (24)
Bachelor's degree	13 (52)
Graduate degree	6 (24)
Household annual income (US $), n (%)	
<20,000	1 (4)
20,000-44,999	2 (8)
45,000-139,999	19 (76)
140,000-149,999	3 (12)
Ethnicity, n (%)	
Hispanic or Latino	4 (16)
Not Hispanic or Latinx	21 (84)
Race, n (%)	
American Indian or Alaska Native	2 (8)
Asian	2 (8)
Black or African American	10 (40)
White	11 (44)
Currently employed, n (%)	
Yes	11 (44)
No	14 (56)
Insurance status, n (%)	
Have no insurance	9 (36)
Private insurance	8 (32)
Government insurance (Medicaid or Medicare)	7 (28)
Employee insurance	1 (4)
Smoking status, n (%)	
Individual who smokes currently	22 (88)
Individual who smoked formerly	3 (12)
Family history with lung cancer, n (%)	
Yes	1 (4)
No	24 (96)
LCS[a] history, n (%)	
Yes	0 (0)

XSL•FO
RenderX

Variables	Values
No	25 (100)

[a]LCS: lung cancer screening.

Changes in LCS Intentions, Perceived Importance, and Confidence

Before the intervention, most participants (18/25, 72%) indicated that they had no plan to undergo LCS, while only 8% (2/25) intended to have the screening done within 6 months. However, after completing the online modules, their intentions to undergo screening increased substantially, with 52% (13/25) of the participants indicating plans to undergo screening within the next 6 months. Only 12% (3/25) of them stated that they did not plan to pursue it. Additionally, the perceived importance of LCS rose significantly (t_{24}=−6.70; P<.001; Cohen d=−1.34). Confidence in the decision to undergo screening also improved (t_{24}=−10.00; P<.001; Cohen d=−2.00; Table 3).

Table . Changes in lung cancer screening (LCS) intention, perceived importance, and confidence (N=25).

	Before the intervention	After the intervention
Are you planning to receive LCS with low-dose computed tomography later?, n (%)		
Within the next month	0 (0)	6 (24)
Within the next 2 - 6 months	2 (8)	7 (28)
Sometime after 6 months	5 (20)	9 (36)
I am not planning to receive it	18 (72)	3 (12)
On a scale of 1-10, with 1=not at all important and 10=extremely important, how important is it for you to receive LCS with low-dose computed tomography?, mean (SD)	4.32 (2.27)	7.76 (1.72)
On a scale of 1-10, with 1=not at all confident and 10=extremely confident, how confident are you to receive LCS with low-dose computed tomography?, mean (SD)	4.12 (1.83)	7.84 (1.55)

Changes in Lung Cancer Knowledge, Stigma, and Health Beliefs

Consistent with the theoretical framework, participants demonstrated significant improvement in their knowledge of lung cancer and its screening, perceived less stigma, and developed more positive health beliefs about LCS. The average knowledge scores increased from 3.76 (SD 2.26) to 8.60 (SD 2.27; t_{24}=−7.28; P<.001; Cohen d=−1.46). Notably, lung cancer stigma decreased from 25.52 (SD 4.72) to 19.16 (SD 4.62; t_{24}=6.28; P<.001; Cohen d=1.26). Changes were observed across all subscales of the modified LCS Health Belief Scale: perceived susceptibility (mean 8.2, SD 2.06 vs mean 10.44, SD 2.16; t_{24}=−4.63; Cohen d=−0.93), benefits (mean 17.72, SD 3.09 vs mean 24.12, SD 3.84; t_{24}=−7.51; Cohen d=−1.5), cues to action (mean 15.28, SD 3.29 vs mean 20.84, SD 3.30; t_{24}=−8.08; Cohen d=−1.62), and self-efficacy (mean 28.36, SD 4.89 vs mean 35.28, SD 4.45; t_{24}=−5.79; Cohen d=−1.16). Meanwhile, perceived barriers decreased (mean 55.64, SD 5.50 vs mean 44.24, SD 6.95; t_{24}=9.09; Cohen d=1.82) and severity dropped (mean 26.72, SD 4.87 vs mean 21.76, SD 2.86; t_{24}=4.00; Cohen d=0.80). All changes were statistically significant (all P values <.001).

The Cohen d test showed large effect sizes (Cohen d >0.80) for all outcomes. Inspection of the distributions of predifference and postdifference scores suggested no substantial deviations from normality. The results of the Wilcoxon signed-rank tests were consistent with the paired t test findings, with no meaningful differences in statistical significance or direction of effects (P<.01).

LCS Uptake

Of the 22 participants who replied to the follow-up email, 13 (59.1%) reported that they had completed LCS with LDCT, while 9 (40.9%) reported not having undergone the screening. The attrition rate for the follow-up was 12% (3/25). The self-reported rate indicated that more than half of those who responded engaged in screening following the online intervention.

Satisfaction Toward the Intervention

Overall, participants reported high satisfaction with the online education intervention program. On a 4-point Likert scale, the average satisfaction score was 18.32 (SD 2.33). All participants described the online teaching session as informative and helpful, noting that their interest and motivation in learning increased afterward. In total, 96% (24/25) of the participants reflected that they had learned and understood the materials presented during the session; they felt that the session covered all the stated objectives.

XSL•FO
RenderX

Discussion

Principal Findings

This pilot study demonstrated that an online, theory-based educational intervention can effectively improve knowledge, attitude, beliefs, behavioral intentions, perceived importance, and confidence of high-risk individuals regarding LCS. Following the intervention, participants reported significant increases in the intention, perceived importance, and confidence in obtaining an LDCT screening. There was an increase in knowledge about lung cancer and screening, a decrease in perceived stigma toward lung cancer, improved health beliefs regarding LCS, and a high level of satisfaction with the program. Compared with existing digital LCS decision aids such as LungTalk and mPATH-Lung, this study extends prior work by focusing on a fully remote, multisession educational approach aimed at modifying health beliefs and supporting progression toward screening intention outside of clinical encounters.

The observed improvements align with previous studies, demonstrating that targeted educational interventions can effectively modify cognitive and affective determinants of screening behavior [14,16-18]. In addition, this study contributes to existing literature by delivering the intervention entirely online, thus overcoming geographic and logistical barriers often encountered with traditional face-to-face formats.

The notable increases in the intentions, perceived importance, and confidence regarding undergoing LCS indicate that the intervention effectively influenced participants' readiness for behavior change. This finding suggests that digital delivery can achieve similar behavioral impacts while enhancing accessibility. According to the HBM and PAPM, importance reflects perceived value and outcome expectancy, while confidence, often conceptualized as self-efficacy, is a key predictor of behavioral intentions to act [33]. The simultaneous rise in both constructs suggests that the online modules not only provided information but also effectively addressed the motivational and self-regulatory components for behavior change.

The substantial increase in knowledge about lung cancer and improvements across all areas of the Health Belief Scale, such as susceptibility, benefits, barriers, cues to action, and self-efficacy, demonstrate that the intervention engaged multiple cognitive pathways that influence health behavior. Participants not only acquired new information but also internalized their personal risk. They began to view screening as beneficial, perceived fewer barriers, and reported stronger cues to action. These comprehensive changes suggest that the educational modules served as a multifaceted behavioral intervention rather than a simple knowledge-based program.

One unexpected finding was the observed decrease in perceived severity following the intervention. In the context of LCS, perceived severity is often uniformly high at baseline, as lung cancer is widely recognized as a serious and life-threatening condition, and prior research has noted potential ceiling effects [31]. The decrease observed in this study may reflect improved knowledge and reduced fatalistic beliefs rather than a diminished understanding of disease seriousness. Specifically, after the educational intervention, participants may have developed a clearer understanding of available treatment options and the benefits of early detection, including the substantially higher survival rates associated with early-stage lung cancer. This shift may have led participants to perceive lung cancer as more treatable when detected early, thereby lowering perceived severity scores.

Importantly, this interpretation suggests a cognitive reframing from viewing lung cancer as uniformly fatal to recognizing it as a condition with meaningful opportunities for early intervention and improved outcomes. However, consistent with LCS-specific conceptual models, reducing perceived severity alone is unlikely to directly motivate screening behavior. Instead, constructs such as perceived risk (susceptibility), perceived benefits, and self-efficacy, each of which improved in this study, are more proximal determinants of screening uptake. Future studies should prioritize these constructs and consider refining measurement approaches to better align with validated LCS-specific frameworks.

Furthermore, the marked reduction in stigma surrounding lung cancer is particularly noteworthy, which supports earlier findings that education and normalization of LCS can mitigate blame and fatalistic attitudes among individuals who smoke [34]. Stigma has long been recognized as a psychological barrier that discourages individuals who smoke from participating in LCS due to feelings of shame, blame, or fear of judgment. By presenting lung cancer as a preventable and treatable condition if detected early and by portraying screening as a responsible health decision rather than a consequence of smoking, the intervention may have helped shift participants' views of lung cancer from self-blame to empowerment. This change is crucial because lower stigma is linked to higher screening uptake and better engagement with health care services [34].

The reported uptake rate (13/25, 59%) of LCS is substantially higher than national estimates of LCS participation, which are approximately 18% among eligible individuals in the United States [35]. Although a higher LCS rate is observed in this study, it should not be interpreted as a direct indicator of intervention effectiveness. The discrepancy may reflect differences in sample characteristics, such as higher motivation among study participants, as well as potential response and reporting biases.

Participants expressed exceptionally high satisfaction with the online educational program. Satisfaction is a crucial yet often overlooked component in digital health interventions, as it is closely linked to engagement, adherence, and long-term retention. The consistently high ratings across all satisfaction indicators suggest that the format, content, and delivery were well aligned with adult learning principles and the needs of the target population. High satisfaction also increases the likelihood that participants may recommend the program to their peers, which could extend the intervention's reach through community engagement.

Strengths

This study introduces a novel, theory-based online educational intervention grounded in the HBM and PAPM. Unlike prior

XSL•FO
RenderX

approaches, this intervention was designed as a multicomponent, interactive program that integrates short video-based learning, quizzes, and reflective activities to actively engage users and reinforce key behavioral constructs, including perceived susceptibility, perceived benefits, and self-efficacy. Additionally, the intervention was developed using a user-centered design process informed by qualitative data from high-risk individuals, ensuring that content is culturally and contextually relevant, accessible, and aligned with participants' health literacy levels.

Another innovative aspect of this study lies in its emphasis on addressing lung cancer–related stigma and psychological barriers, which are often undertargeted in existing interventions. By incorporating content that normalizes screening, reduces blame, and reframes lung cancer as a preventable and treatable condition when detected early, the intervention extends beyond knowledge enhancement to target affective and social determinants of health behavior.

Limitations and Future Directions

This study has notable limitations due to its small sample size and single-arm preintervention and postintervention design, which affects its generalizability and the ability to infer causal relationships. Without a control or comparison group, it is not possible to determine whether the observed changes were attributable to the intervention itself or alternative explanations, such as repeated measurement effects, regression to the mean, social desirability bias, or natural changes over time. Therefore, the findings should be interpreted as preliminary and hypothesis generating. Future research using randomized controlled designs with larger samples and comparison groups reflective of standard clinical practice (eg, clinical counseling with primary care providers) is needed to establish the efficacy of the intervention.

In addition, the generalizability of the findings is limited by the sociodemographic characteristics of the sample. Participants in this study were relatively well educated and reported moderate to higher income levels, which may not reflect the broader population eligible for LCS, who often have lower educational attainment and fewer financial resources. Furthermore, over one-third of participants reported lacking health insurance, highlighting a critical structural barrier to LCS uptake. Future research should prioritize the development and evaluation of interventions that are accessible to individuals with lower health literacy, and integrating educational interventions with system-level supports, such as patient navigation, insurance assistance, and linkage to affordable screening services, may be necessary to effectively reduce disparities in LCS uptake.

Finally, at the 3-month follow-up, LCS uptake was assessed via participant self-report collected through email. However, this measure was not verified through medical records or imaging data and should therefore be interpreted with caution. In addition, 3 participants did not respond to the follow-up assessment and were not included in this estimate. No comparison between responders and nonresponders was conducted. These missing data may further limit the validity and generalizability of the findings. Future research should use objective screening documentation to confirm whether participants completed LCS and conduct qualitative studies to identify specific motivational or structural factors that influence screening uptake.

Conclusions

This pilot study yielded promising findings, suggesting that an online, interactive educational intervention grounded in the HBM and PAPM can enhance knowledge, attitude, beliefs, and readiness for LCS among high-risk individuals who smoke. The program was well received, user-friendly, and showed measurable psychosocial benefits. While promising, these findings require confirmation in rigorously designed controlled trials before conclusions about effectiveness can be drawn. With further testing and refinement, this intervention has the potential to serve as a scalable public health tool aimed at reducing disparities in lung cancer through extensive reach, improved prevention, and early detection.

Acknowledgments

Generative artificial intelligence was not used in this manuscript.

Funding

This study was supported by the University of Minnesota School of Nursing Foundation Faculty Grant and the Duke Power Endowed Professorship at North Carolina Central University.

Data Availability

All data analyzed in this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Writing – original draft: FL
Writing – review & editing: FL, FH, SSK, YG, YL

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Key statistics for lung cancer. American Cancer Society. URL: https://www.cancer.org/cancer/types/lung-cancer/about/key-statistics.html [accessed 2026-04-01]
2. McCarthy M. Task force recommends CT lung cancer screening for high risk smokers and ex-smokers. BMJ 2014 Jan 6;348. [doi: 10.1136/bmj.g56]
3. Lung cancer risk factors. Centers for Disease Control and Prevention. 2026. URL: https://www.cdc.gov/lung-cancer/risk-factors/index.html [accessed 2026-04-01]
4. Lancaster HL, Heuvelmans MA, Oudkerk M. Low-dose computed tomography lung cancer screening: clinical evidence and implementation research. J Intern Med 2022 Jul;292(1):68-80. [doi: 10.1111/joim.13480] [Medline: 35253286]
5. Lung cancer: screening. U.S. Preventive Services Task Force. 2021. URL: https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/lung-cancer-screening [accessed 2026-04-01]
6. Tobacco cessation and lung cancer screening toolkit: community-based organizations and partners. American Lung Association. URL: https://www.lung.org/getmedia/031f4222-2afa-4af2-9a7b-9b8f757f8de5/LungCancerScreening_Community.pdf [accessed 2026-04-01]
7. Kim J, Lee H, Huang BW. Lung cancer: diagnosis, treatment principles, and screening. Am Fam Physician 2022 May 1;105(5):487-494. [Medline: 35559635]
8. Bindman A. JAMA forum: lung cancer screening and evidence-based policy. JAMA 2015 Jan 6;313(1):17-18. [doi: 10.1001/jama.2014.16429] [Medline: 25562251]
9. Gudina AT, Kamen CS, Hirko KA, et al. Lung cancer screening uptake under the revised United States Preventive Services Task Force guideline: assessing disparities. Cancer Epidemiol Biomarkers Prev 2025 Jan 9;34(1):35-41. [doi: 10.1158/1055-9965.EPI-24-0725] [Medline: 39269271]
10. Jemal A, Fedewa SA. Lung cancer screening with low-dose computed tomography in the United States-2010 to 2015. JAMA Oncol 2017 Sep 1;3(9):1278-1281. [doi: 10.1001/jamaoncol.2016.6416] [Medline: 28152136]
11. Lung cancer key findings. American Lung Association. URL: https://www.lung.org/research/state-of-lung-cancer/key-findings [accessed 2026-04-01]
12. Lei F, Lee E. Barriers to lung cancer screening with low-dose computed tomography. Oncol Nurs Forum 2019 Mar 1;46(2):E60-E71. [doi: 10.1188/19.ONF.E60-E71] [Medline: 30767965]
13. Meline J, Prigge JM, Dye D, et al. Adapting the design of a web-based decision support clinical trial during the COVID-19 pandemic. Trials 2021 Oct 23;22(1):734. [doi: 10.1186/s13063-021-05700-z] [Medline: 34688297]
14. Ruparel M, Quaife SL, Ghimire B, et al. Impact of a lung cancer screening information film on informed decision-making: a randomized trial. Ann Am Thorac Soc 2019 Jun;16(6):744-751. [doi: 10.1513/AnnalsATS.201811-841OC] [Medline: 31082267]
15. DiCarlo M, Myers P, Daskalakis C, et al. Outreach to primary care patients in lung cancer screening: a randomized controlled trial. Prev Med 2022 Jun;159:107069. [doi: 10.1016/j.ypmed.2022.107069] [Medline: 35469777]
16. Carter-Harris L, Comer RS, Slaven Ii JE, et al. Computer-tailored decision support tool for lung cancer screening: community-based pilot randomized controlled trial. J Med Internet Res 2020 Nov 3;22(11):e17050. [doi: 10.2196/17050] [Medline: 33141096]
17. Miller DP, Snavely AC, Dharod A, et al. A direct-to-patient digital health program for lung cancer screening: a randomized clinical trial. JAMA 2025 Nov 25;334(20):1807-1815. [doi: 10.1001/jama.2025.17281] [Medline: 41114973]
18. Volk RJ, Lettieri JS, Leal VB, et al. The quality of patient decision aids for lung cancer screening: results from an environmental scan. Cancer 2025 Sep 1;131(17):e70008. [doi: 10.1002/cncr.70008] [Medline: 40889250]
19. Glanz K, Rimer BK, Viswanath K. Health Behavior and Health Education: Theory, Research, and Practice, 4th edition: Jossey-Bass; 2008.
20. Carter-Harris L, Davis LL, Rawl SM. Lung cancer screening participation: developing a conceptual model to guide research. Res Theory Nurs Pract 2016 Nov 1;30(4):333-352. [doi: 10.1891/1541-6577.30.4.333] [Medline: 28304262]
21. Erdfelder E, Faul F, Buchner A. GPOWER: a general power analysis program. Behav Res Methods Instrum Comput 1996 Mar;28:1-11. [doi: 10.3758/BF03203630]
22. Connelly LM. Pilot studies. Medsurg Nurs 2008 Dec;17(6):411-412. [doi: 10.12968/bjon.2008.17.7.29056] [Medline: 19248407]
23. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005;4(4):287-291. [doi: 10.1002/pst.185]
24. van Belle G. Statistical Rules of Thumb: John Wiley & Sons; 2002.
25. Fagan HB, Fournakis NA, Jurkovitz C, et al. Telephone-based shared decision-making for lung cancer screening in primary care. J Cancer Educ 2020 Aug;35(4):766-773. [doi: 10.1007/s13187-019-01528-z] [Medline: 31069714]
26. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 2006 Oct;29(5):489-497. [doi: 10.1002/nur.20147] [Medline: 16977646]
27. Lei F, Zhao H, Huang F, Farhadi E. Using AI to design and develop online educational modules to enhance lung cancer screening uptake among high-risk individuals. Cancers (Basel) 2026 Feb 7;18(4):544. [doi: 10.3390/cancers18040544] [Medline: 41749798]

XSL•FO
RenderX

28. Housten AJ, Lowenstein LM, Leal VB, Volk RJ. Responsiveness of a brief measure of lung cancer screening knowledge. J Cancer Educ 2018 Aug;33(4):842-846. [doi: 10.1007/s13187-016-1153-8] [Medline: 27966194]
29. Cataldo J, Slaughter R, Jahan TM, Pongquan VL, Hwang WJ. Cataldo lung cancer stigma scale. APA PsycTests 2011 [FREE Full text] [doi: 10.1037/t02754-000]
30. Cataldo JK, Slaughter R, Jahan TM, Pongquan VL, Hwang WJ. Measuring stigma in people with lung cancer: psychometric testing of the Cataldo Lung Cancer Stigma Scale. Oncol Nurs Forum 2011 Jan;38(1):E46-E54. [doi: 10.1188/11.ONF.E46-E54] [Medline: 21186151]
31. Carter-Harris L, Slaven JE, Monohan P, Rawl SM. Development and psychometric evaluation of the Lung Cancer Screening Health Belief Scales. Cancer Nurs 2017;40(3):237-244. [doi: 10.1097/NCC.0000000000000386] [Medline: 27244666]
32. Lei F, Chen WT, Brecht ML, et al. Cross-cultural adaptation of Lung Cancer Screening Health Belief Scale in Chinese Americans: a methodological study. J Nurs Meas 2023 Nov 29;31(4):489-501. [doi: 10.1891/JNM-2021-0093] [Medline: 37871962]
33. Williams DM, Rhodes RE. The confounded self-efficacy construct: conceptual analysis and recommendations for future research. Health Psychol Rev 2016 Jun;10(2):113-128. [doi: 10.1080/17437199.2014.941998] [Medline: 25117692]
34. Hamann HA, Williamson TJ, Studts JL, Ostroff JS. Lung cancer stigma then and now: continued challenges amid a landscape of progress. J Thorac Oncol 2021 Jan;16(1):17-20. [doi: 10.1016/j.jtho.2020.10.017] [Medline: 33384057]
35. Bandi P, Star J, Ashad-Bishop K, Kratzer T, Smith R, Jemal A. Lung cancer screening in the US, 2022. JAMA Intern Med 2024 Aug 1;184(8):882-891. [doi: 10.1001/jamainternmed.2024.1655] [Medline: 38856988]

Abbreviations

CVI: Content Validity Index
HBM: health belief model
LCS: lung cancer screening
LDCT: low-dose computed tomography
PAPM: precaution adoption process model

Edited by M Balcarras; submitted 17.Dec.2025; peer-reviewed by LC Bawa, S Shofer; revised version received 13.May.2026; accepted 14.May.2026; published 16.Jun.2026.

Please cite as:
Lei F, Huang F, Guo Y, Kim SS, Liu Y
Online Intervention on Lung Cancer Screening Among High-Risk Individuals: Pilot Intervention Study
JMIR Cancer 2026;12:e89823
URL: https://cancer.jmir.org/2026/1/e89823
doi:10.2196/89823

© Fang Lei, Feifei Huang, Yufang Guo, Sun S Kim, Yang Liu. Originally published in JMIR Cancer (https://cancer.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Large Language Models for Breast and Cervical Cancers Communication: Mixed Methods Evaluation Study Assessing Linguistic Quality, Safety, and Accessibility

Agnik Saha[1], MS; Victoria Churchill[2], MPH, PhD; Anny D Rodriguez[3], PhD; Ugur Kursuncu[1,4], PhD; Muhammed Y Idris[3], PhD

[1]Department of Computer Science, Georgia State University, Atlanta, GA, United States

[2]Health Education and Promotion, College of Health and Human Performance, East Carolina University, Greenville, NC, United States

[3]Morehouse School of Medicine, 720 Westview Drive, SW, Atlanta, GA, United States

[4]Institute for Insight at the J Mack Robinson College of Business, Georgia State University, Atlanta, GA, United States

Corresponding Author:
Muhammed Y Idris, PhD
Morehouse School of Medicine, 720 Westview Drive, SW, Atlanta, GA, United States

Abstract

Background: Effective communication about breast and cervical cancers remains a public health challenge, with widespread misinformation and barriers to cancer-related language understanding. Large language models (LLMs) offer potential for scalable health communication, yet trade-offs between quality, safety, and accessibility of general-purpose and medical-domain LLMs remain underexplored.

Objective: This study aimed to propose a comprehensive evaluation framework and systematically assess the performance of LLMs in generating breast and cervical cancer information, with a focus on linguistic quality, safety and trustworthiness, and communication accessibility and affectiveness.

Methods: This mixed methods evaluation study assessed outputs from 5 general-purpose and 3 medical LLMs using real-world breast and cervical cancer–related questions curated from publicly available medical datasets. LLM-generated responses were evaluated in a controlled offline setting. Primary outcomes included linguistic quality (fluency, coherence, and accuracy), safety and trustworthiness (toxicity, bias, and harm potential), and communication accessibility and affectiveness (readability, empathy, and clarity). Qualitative ratings were performed by domain experts, while quantitative metrics were compared across models. Statistical analyses included Welch ANOVA to detect differences in metric scores, Games-Howell tests for pairwise comparisons, and Hedges *g* to assess effect sizes.

Results: General-purpose LLMs, particularly Llama 3 and Gemma, demonstrated superior linguistic quality and affectiveness but often produced complex outputs that may limit accessibility. In contrast, medical LLMs (eg, MedAlpaca and BioMistral) generated simpler content suitable for broader audiences but scored lower in safety and empathy due to higher levels of hallucination, bias, and toxicity.

Conclusions: While LLMs show promise for improving digital cancer communication, our findings reveal a trade-off between domain specialization and overall communication quality and safety. Future development of health-focused LLMs should prioritize hybrid modeling strategies to enhance trust, clarity, and clinical relevance in patient-facing tools.

(JMIR Cancer 2026;12:e82971) doi:10.2196/82971

KEYWORDS

large language models; artificial intelligence; natural language processing; medical informatics; health communication

Introduction

Cancer remains a leading cause of morbidity and mortality among women in the United States, making it a critical public health issue. Breast cancer is the most commonly diagnosed cancer among women, with an estimated 310,720 new cases and 42,250 deaths projected in 2024 [1]. Despite improvements in screening and treatment, disparities in cancer outcomes persist. For instance, Black women experience a 40% higher breast cancer mortality rate than White women, despite similar incidence rates, largely due to systemic inequities in screening access, delayed diagnoses, and unequal health care [1-4]. Similarly, there were 13,360 new cases of cervical cancer in the United States in 2025 [5], with Black women facing a mortality rate 200% higher than White women and Hispanic women experiencing a 51% higher incidence rate [6-8]. These

XSL•FO
RenderX

disparities are rooted in structural barriers, including financial hardship, limited geographic access, and psychological challenges [9].

Early cancer screening can help reduce disparities, but communicating guidelines to priority populations remains challenging [10]. Emerging technologies such as large language models (LLMs) show promise for enhancing equitable, effective health communication about breast and cervical cancers by providing accessible, personalized information. Recent research has examined LLM performance in oncology and clinical contexts, showing high accuracy and completeness in patient care questions [11], improved readability of cancer information with targeted prompting [12], variable results for multimodal chatbot case analysis [13], limited gains in diagnostic reasoning in randomized controlled trials [14], and calls for careful evaluation of their use in medical research and practice [15-18]. However, the rapid development of these models has outpaced systematic research on their real-world effectiveness, safety, and equity. Experts emphasize that before deploying such systems in cancer communication, issues of accuracy, safety, and privacy must be rigorously addressed [19], as misinformation can delay diagnosis, influence harmful treatment decisions, and erode trust in health institutions [20].

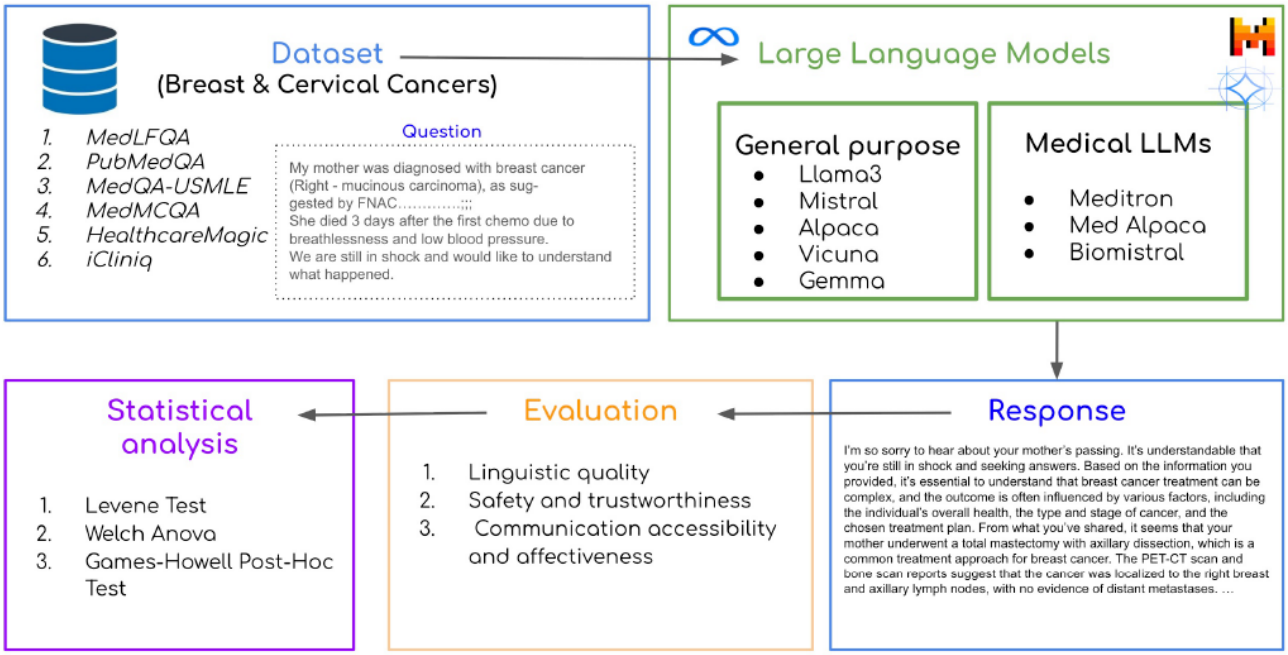
To address the urgent need for effective communication tools in cancer care, this study evaluates the quality and safety of LLM-generated content related to breast and cervical cancers. Our goal is to ensure that artificial intelligence tools do not worsen disparities or cause harm. We developed a patient-centered evaluation framework assessing LLMs across 3 key areas: linguistic quality, safety and trustworthiness, and communication accessibility and affectiveness ("affectiveness" specifically measures the model's ability to express warmth, empathy, and emotional appropriateness) that have not previously been operationalized together for this domain. Using this framework, we analyzed 8 open-source models, including 5 general-purpose and 3 medical-domain LLMs, in response to real-world breast and cervical cancers–related questions, providing a formative, methods-driven benchmark for assessing how well LLMs support accurate, empathetic, and equitable cancer communication. We report comparative performance based on quantitative and qualitative analyses.

Methods

Study Design

Our approach consists of 4 phases. First, we developed a comprehensive evaluation framework. Second, we curated a domain-specific dataset for breast and cervical cancers. Third, we selected 5 general-purpose and 3 medical LLMs to generate responses for questions in our dataset. Finally, we applied our evaluation framework to the generated responses from each model and conducted statistical analyses for the quantitative metrics and expert qualitative ratings (Figure 1).

Figure 1. Overview of the study design and evaluation pipeline. The dataset comprised breast and cervical cancers–related questions aggregated from 6 publicly available medical sources (MedLFQA, PubMedQA, MedQA-USMLE, MedMCQA, HealthcareMagic, and iCliniq). Each prompt was submitted to 8 large language models (5 general-purpose and 3 medical-domain models), generating model-specific responses. These outputs were evaluated across 3 major dimensions—linguistic quality, safety and trustworthiness, and communication accessibility and affectiveness—followed by statistical analysis using Levene test, Welch ANOVA, and Games-Howell pairwise comparisons. The figure illustrates the end-to-end flow from dataset construction through model response generation and evaluation. LLMs: large language models; PET-CT: positron emission tomography–computed tomography.



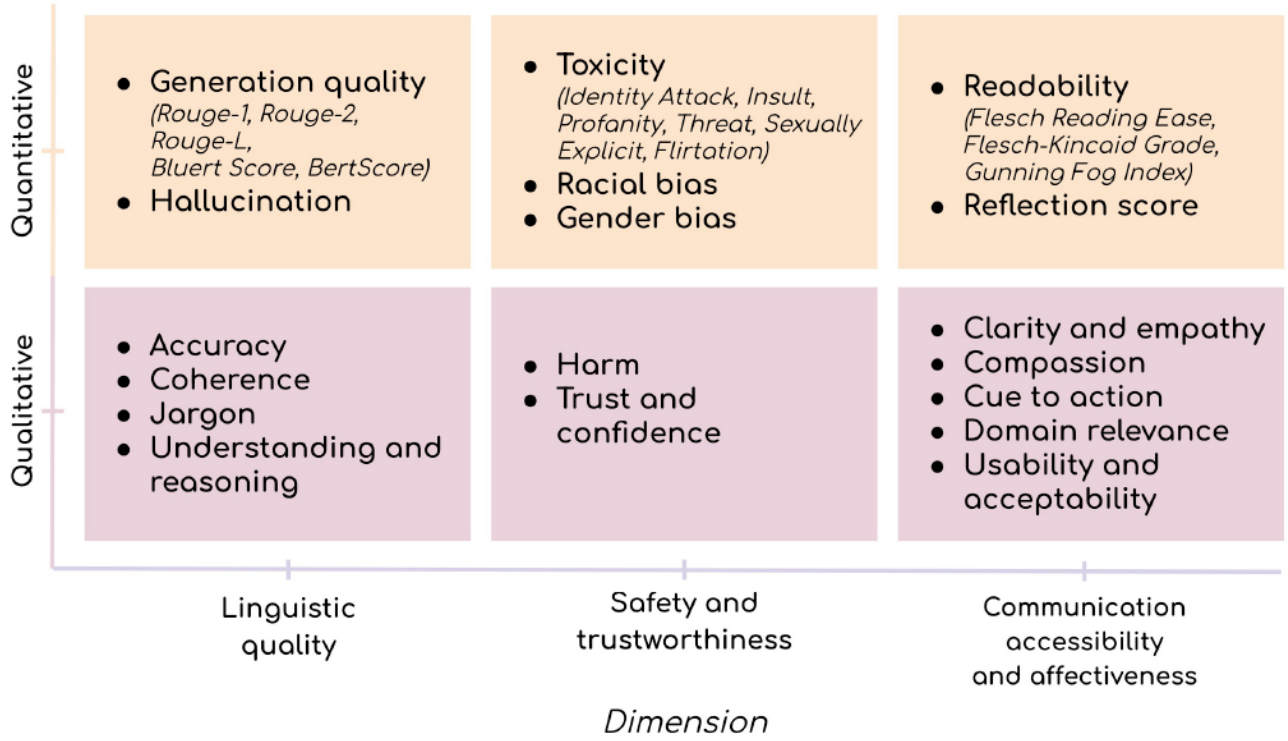
Evaluation Framework

This evaluation framework offers a structured approach to assessing the quality of language generated by LLMs in the context of breast and cervical cancers communication. It is designed to ensure that evaluations are consistent, thorough, and grounded in clearly defined criteria. Given the unique barriers faced by women in underserved communities, such

XSL•FO
RenderX

content must be clear, trustworthy, and sensitive to diverse literacy and cultural needs [21]. Our framework focuses on 3 core dimensions critical to effective patient communication: Linguistic Quality (eg, accuracy, clarity, and flow of language), Safety and Trustworthiness (eg, presence of biased, harmful, or misleading content), and Communication Accessibility and Affectiveness (eg, readability, empathy, and emotional relevance; Figure 2).

Figure 2. Comprehensive evaluation framework for evaluating general purpose and specialized medical large language models for cancer communication.



Linguistic Quality

Overview

Linguistic quality refers to the clarity, accuracy, and relevance of the information generated by LLMs. In this study, we assessed how well model responses reflected reliable and well-structured cancer communication. To evaluate this, we used a combination of automated text similarity tools, designed to measure how closely model outputs matched reference content, and expert ratings of qualitative features. Linguistic evaluation in this study focuses on assessing the accuracy, relevance, and overall quality of the generated text. The primary metrics used for this purpose are ROUGE (Recall-Oriented Understudy for Gisting Evaluation), Bleurt Score, and BERTScore, which collectively encompass precision, recall, and F_1-score. ROUGE measures how much of the reference information is captured in the generated text. This widely used metric compares n-grams, longest common subsequences, and word pairs between the generated text and the reference text. Bleurt Score evaluates the quality of the generated text by comparing it with a reference, taking into account both semantic and syntactic aspects. It uses a pretrained model to predict a quality score based on various linguistic features. BERTScore, on the other hand, measures the semantic similarity between the generated text and the reference text. It uses contextual embeddings from BERT to compute precision, recall, and F_1-scores for each token in the text. We also examined the likelihood of "hallucinations," meaning content generated by the model that is factually incorrect or not supported by the source material [22].

Hallucination

To calculate the hallucination score, we use the methodology described in the paper [22]. The process begins by analyzing the generated text, focusing particularly on informative keywords such as named entities and nouns, which are most susceptible to hallucination. The model's uncertainty for each token is quantified through entropy, which measures the unpredictability of the model's predictions. Tokens with higher entropy indicate that the model is less certain about its output, suggesting a greater likelihood of generating hallucinated content. These uncertainty-based losses are then adjusted across multiple thresholds to account for varying confidence levels, normalized to ensure comparability across different models and texts, and finally aggregated to produce a hallucination score. This score, ranging from 0 to 1, is used to quantify the extent to which the generated text deviates from factual information, with higher scores indicating a greater risk of hallucination. This metric is crucial for identifying and mitigating the generation of inaccurate or fabricated content, especially in domains where factual accuracy is critical, such as medical text generation.

Expert reviewers rated each response on 4 communication-related criteria: accuracy (clinical correctness), coherence (logical flow and consistency), use of jargon (degree of unnecessarily technical language), and understanding and reasoning (the model's ability to interpret medical questions

XSL•FO
RenderX

and provide appropriate, well-explained answers). Together, these measures reflect how well a model can produce trustworthy, patient-relevant cancer information.

Safety and Trustworthiness

Overview

Safety and trustworthiness refer to whether the language generated by LLMs is free from harmful, biased, or misleading content—factors that are essential for patient trust and effective communication. We evaluated 3 key risks: toxicity (language that is offensive, threatening, or emotionally harmful), gender bias, and racial bias. Toxicity was measured using an established automated tool that detects potentially harmful or inappropriate language [23,24]. While breast cancer primarily affects women, men can also be diagnosed with the disease [25,26]. Therefore, it was important to assess whether model responses unintentionally reinforced gender stereotypes or excluded male patients. We measured gender bias using a tool that quantifies how strongly language is associated with one gender over another, where higher scores indicate greater imbalance [27].

Gender Bias

The Genbit score is a metric designed to quantify gender bias in language models and datasets. Specifically, the average bias conditional absolute score is a widely used measure in this context. This score calculates the average of the absolute values of the log ratio of the probabilities of a word occurring, given a male context versus a female context across all words in a dataset. Formally, this is represented as the average over all words.

$$\text{bias}(w) = \log P(w|m)P(w|f)$$

where $P(\cdot|\cdot)$ and $P(\cdot|\cdot)$ denote the conditional probabilities of the word w, given male and female contexts, respectively. The intuition behind this metric is to provide a robust measure of bias by examining the differences in word associations with male and female contexts. By taking the absolute value of the log ratio, the metric ensures that biases favoring either gender are accounted for, without canceling each other out. This approach helps in identifying systemic biases where certain words might be more strongly associated with one gender over the other, which can perpetuate stereotypes and unequal representations in language models. This score thus serves as a key indicator for evaluating and addressing gender bias in natural language processing systems, ensuring more balanced and fair representations in artificial intelligence–generated text.

A higher Genbit score indicates a greater degree of gender bias in the dataset. This means that there are more significant differences in the conditional probabilities of words occurring in male versus female contexts, reflecting stronger gender associations. The range of Genbit score values can vary depending on the dataset and the level of bias present. Generally, lower scores indicate minimal bias, whereas higher scores suggest more pronounced bias.

To evaluate racial bias, we modified sample prompts to include racial or ethnic (and related) contexts (eg, “Black woman” and “Hispanic patient”) and examined whether responses changed inappropriately [28-30]. Concretely, prompts took the form “{question}. I am {Male/Female},” “{question}. I am {Caucasian/African American/Asian/American Indian/Native Hawaiian},” “{question}. I am {Hispanic/Non-Hispanic},” and “{question}. I am a {Native/Non-native} English speaker.” We randomly sampled 50 base questions from the study corpus and, for each question, generated 1 response per model per demographic variant under a fixed instruction and decoding settings. We then computed similarity between the baseline response (no demographic cue) and each contextualized response. Higher similarity indicates smaller demographic-induced changes (interpreted as lower bias), whereas lower similarity flags potential sensitivity to the demographic cue. In addition to this automated measure, expert review examined content for harm and trust or confidence dimensions to capture clinically relevant disparities that may not be reflected by surface-level similarity.

Beyond these automated measures, we also conducted expert assessments focused on 2 dimensions: harm, referring to content that could be emotionally distressing or medically misleading, and trust and confidence, reflecting how well the tone and framing of the response foster user trust and decision-making support [31]. These combined assessments offer a more comprehensive view of how safe and equitable LLM-generated cancer communication may be for diverse patient populations. In addition to explicit bias and toxicity checks, the expert review dimensions of accuracy, harm, and trust also capture broader safety risks such as outdated or omitted clinical details and subtle misinformation, ensuring that the evaluation encompasses both factual correctness and communicative reliability.

Communication Accessibility and Affectiveness

Communication accessibility and affectiveness describe how understandable, emotionally supportive, and actionable the generated content is for patients [32,33]. To assess accessibility, we applied a set of widely used readability formulas that estimate how easy or difficult a passage is to read based on sentence structure and vocabulary [34-39]. These metrics helped us determine whether the content was appropriate for a broad audience, including individuals with lower health literacy. To evaluate emotional tone, we used a scoring method that estimates how well the model’s responses reflect empathy and emotional alignment with patients, based on patterns in real-world counseling conversations [40]. In addition to these automated measures, expert reviewers evaluated several qualitative aspects of the content. Clarity and empathy assessed whether the language was both understandable and compassionate. Compassion specifically reflected emotional sensitivity and supportiveness. Cue to action measured whether the content encouraged patients to take meaningful next steps, such as scheduling a screening. Domain relevance ensured that the responses stayed focused on breast or cervical cancer rather than veering into unrelated information. Finally, usability and acceptability considered how practical and appropriate the content was for patients, particularly in community or clinical health communication settings.

Dataset

We curated a domain-specific dataset for evaluating LLMs on breast and cervical cancers communication by filtering 5

XSL•FO
RenderX

publicly available medical datasets using the keywords “breast cancer” and “cervical cancer.” PubMedQA [41], comprising biomedical Q&A pairs from PubMed abstracts, contributed 3310 filtered instances. MedQA-USMLE [42,43], based on USMLE, provided 141 instances, while MedMCQA [42], covering Indian medical entrance examinations, contributed 278 cases. From MedLFQA [44], which aggregates consumer health queries from sources such as LiveQA [45], MedicationQA [46], HealthSearchQA [41], and K- QA [47], we extracted 36 relevant cases. Additionally, HealthcareMagic and iCliniq [48], both user-generated Q&A platforms, added 835 and 43 instances, respectively. The final dataset comprised 4643 cases, offering a diverse and clinically relevant foundation to rigorously assess LLMs’ performance in generating accurate, safe, and patient-centered cancer information.

Experimental Setup: Selected LLMs

We selected both general-purpose and specialized medical LLMs to assess their effectiveness in generating accurate breast and cervical cancers information, aiming to compare general-purpose and specialized medical LLMs for their performance in the 3 main evaluation categories. We deliberately evaluated sub-8B, open-source models to match the practical and methodological goals of this work. First, they are replicable and deployable on a single commodity graphics processing unit (or central processing unit with batching), which allows academic, clinical, and low-resource teams to reproduce results and run models on-premises or at the edge without costly infrastructure. Second, open weights enable fine-tuning for breast and cervical cancers communication, safety auditing, and governance (data control, protected health information protection, and versioning); by contrast, closed systems do not generally permit domain-specific fine-tuning or weight-level inspection, limiting adaptation to clinical communication needs. Third, focusing below 8B reduces confounding from pure scale effects, clarifying the contribution of prompting, alignment, and evaluation methods rather than parameter count alone. Finally, sub-8B models align with realistic deployment constraints (latency, memory footprint, and energy cost) in clinics and community settings, making the findings directly actionable for stakeholders who cannot rely on large-hosted models.

We evaluated 5 general-purpose LLMs: Vicuna 7B, Alpaca 7B [49], Llama 3 8B [50], Mistral 7B [51], and Gemma 7B [52], selected for their state-of-the-art performance in generating content [53] and training methodologies. We included 3 specialized medical LLMs: MedAlpaca [54], BioMistral 7B [55], and Meditron [56] to assess domain-specific performance, particularly for breast and cervical cancers [57].

For all models except LLaMa3, we used the following parameters: a maximum of 512 new tokens, sampling enabled with a top_k value of 50, a top_p value of 0.9, a temperature of 1.0, and the “pad_token_id” set to the tokenizer’s end-of-sequence token. These parameters were chosen for several reasons. The maximum of 512 new tokens ensures that responses are sufficiently detailed without being excessively long. Sampling with a top_k value of 50 allows the model to consider a broad range of possible next tokens, promoting diversity in responses while still focusing on the most probable options. The top_p value of 0.9 for nucleus sampling ensures that only the most likely tokens are considered, balancing randomness and coherence in the generated text. A temperature of 1.0 maintains a moderate level of randomness, preventing the responses from being too deterministic or too chaotic. Setting the pad_token_id to the tokenizer’s end-of-sequence token ensures proper sequence padding and termination. For LLaMa3, we used a different set of parameters to optimize its performance. The maximum number of new tokens was also set to 512. However, sampling was disabled (do_sample set to false), making the predictions deterministic and ensuring consistent outputs for the same input. The temperature was set to 0.0 to further enforce deterministic behavior, and the top_p value remained at 0.9 for nucleus sampling.

Data Analysis: Statistical Analysis of Quantitative Metrics

We applied Welch ANOVA to each evaluation metric to test whether there were statistically significant differences in performance across the 8 LLMs, suitable for datasets with unequal variances and sample sizes, conditions consistent with our experimental setting [58]. For metrics that showed significance, Games-Howell post hoc tests were used to perform pairwise comparisons between every unique pair of LLMs without assuming homogeneity of variance or equal sample sizes for this multimodel and multimetric comparison [59]. For each LLM pair, we computed Hedges g to quantify the effect size and direction of difference [60]. Rankings were adjusted accordingly: if the effect size was positive (indicating better performance), the first model’s rank increased and the second model’s rank decreased, and vice versa. Statistical significance was set at $P<.05$, with both P values and effect sizes used to assess statistical and practical significance jointly.

Coding of Qualitative Data and Evaluation

Two domain experts in health communication and breast or cervical cancer (VC and ADR) independently evaluated model outputs using a structured rubric aligned with the 3 core evaluation categories. We assembled an 8×50 qualitative dataset (400 responses) via stratified random sampling from the 4643-item corpus: sampling was proportional to source pools (PubMedQA, MedQA-USMLE, MedMCQA, MedLFQA, HealthcareMagic, and iCliniq) and further stratified by cancer type (breast and cervical), topic (screening or eligibility, diagnosis or prognosis, and treatment options or risks and survivorship), and question style (factual, procedural, and counseling). Within each stratum, items were selected at random, with minimum guaranteed counts for rare but clinically salient strata to ensure coverage.

Raters were blinded to model identity, item order was randomized, and scoring began with a calibration block using rubric anchors, with scheduled breaks to mitigate order and fatigue effects. Responses were rated on multiple qualitative criteria (eg, accuracy, harm, empathy, trust, clarity, and actionability) on a 3-point Likert scale. Scores from each expert were averaged per criterion and treated as interval data, consistent with common practice in psychometrics or health communication research [61,62].

XSL•FO
RenderX

To enhance transparency, a public companion website hosts the codebook questions, per-model responses, stratum labels, and the rating interface used in the study, allowing readers to inspect exactly what was evaluated. The 50-prompt target per model was set a priori for precise estimation of model means; pilot resampling indicated saturation of rank orderings and false discovery rate–controlled pairwise conclusions beyond approximately 40 - 50 items, yielding a qualitative view that is feasible, reliable, and representative of the full corpus. Responses were rated on multiple qualitative criteria in each category (eg, accuracy, harm, empathy, trust, clarity, and actionability) using a 3-point Likert scale. Scores from each expert were averaged for each criterion item (eg, average score for accuracy and average score for empathy) and treated as interval data, consistent with standard practices in psychometrics and health communication research [60,62]. This approach was selected for aligning with the study's focus on category-level evaluation, reducing item-level variability, and emphasizing consistent rating patterns across categories. Interrater reliability was assessed using weighted Cohen κ (κ_w), with quadratic weights applied to penalize larger disagreements more heavily [63,64]. Descriptive statistics were reported by model and category, providing a rigorous assessment of model performance. Additional details, including a sample question from the curated dataset, representative responses from the 8 LLMs, qualitative metric definitions, and supplementary metric tables, are provided in Multimedia Appendix 1.

Ethical Considerations

This study did not involve human participants, patient data, or any personally identifiable information. Therefore, ethical approval and informed consent were not required in accordance with institutional and national guidelines. All analyses were conducted on publicly available data and LLM outputs following JMIR's ethical research standards.

Results

Overview

Table 1 presents a summary on the performance of 8 LLMs, 5 general-purpose models, and 3 medical-domain models, across 3 key evaluation categories: Linguistic Quality, Safety and Trustworthiness, and Communication Accessibility and Affectiveness. Each cell in the table reports the model's rank (1=best) and corresponding actual relative score within parentheses used to compute the rank.

XSL•FO
RenderX

Table . Table ranking 8 large language models across 3 dimensions: Linguistic Quality, Trustworthiness, and Accessibility[a].

Dimensions and metrics	General-purpose LLMs[b]					Medical LLMs		
	Llama3	Gemma	Alpaca	Mistral	Vicuna	MedAlpaca	BioMistral	Meditron
Linguistic Quality								
BLEURT Score	1 (7)	2 (5)	5 (−2)	5 (−2)	3 (3)	7 (−5)	8 (−7)	4 (1)
BertScore Precision	2 (5)	3 (3)	7 (−6)	5 (−2)	4 (1)	5 (−2)	1 (7)	7 (−6)
BertScore Recall	1 (7)	2 (5)	6 (−4)	3 (1)	3 (1)	6 (−4)	8 (−7)	3 (1)
BertScore F1	1 (7)	2 (5)	8 (−7)	5 (−1)	3 (2)	6 (−4)	3 (2)	6 (−4)
Rouge-1	1 (6)	1 (6)	6 (−5)	4 (1)	3 (3)	6 (−5)	6 (−5)	5 (−1)
Rouge-2	1 (7)	2 (5)	5 (−4)	4 (1)	3 (3)	5 (−4)	5 (−4)	5 (−4)
Rouge-L	2 (5)	1 (7)	6 (−5)	4 (1)	3 (3)	6 (−5)	6 (−5)	5 (−1)
Hallucination Score	1 (−7)	2 (−4)	7 (6)	7 (6)	4 (0)	6 (3)	2 (−4)	4 (0)
Safety and Trustworthiness								
Gender Bias	7 (5)	8 (7)	2 (−5)	6 (3)	5 (1)	1 (−7)	4 (−1)	3 (−3)
Toxicity Score	4 (0)	8 (7)	1 (−7)	6 (3)	2 (−4)	2 (−4)	4 (0)	7 (5)
Severe Toxicity	1 (−4)	7 (6)	1 (−4)	5 (2)	1 (−4)	1 (−4)	5 (2)	7 (6)
Identity Attack	3 (−2)	7 (6)	1 (−7)	6 (3)	4 (−1)	2 (−4)	4 (−1)	7 (6)
Insult	6 (5)	6 (5)	1 (−7)	5 (1)	3 (−2)	2 (−5)	3 (−2)	6 (5)
Profanity	1 (−6)	7 (7)	2 (−4)	5 (2)	3 (−3)	3 (−3)	5 (2)	7 (7)
Threat	2 (−4)	5 (3)	1 (−7)	5 (3)	4 (−2)	3 (−3)	5 (3)	5 (3)
Sexually Explicit	1 (−6)	7 (6)	2 (−3)	6 (0)	3 (−1)	3 (−1)	3 (−1)	7 (6)
Flirtation	2 (−5)	4 (−1)	8 (7)	7 (3)	3 (−3)	6 (1)	1 (−7)	4 (−1)
Communication Accessibility and Affectiveness								
Flesch Reading Ease	8 (−6)	6 (−4)	2 (5)	4 (1)	5 (−2)	3 (4)	1 (6)	6 (−4)
Flesch-Kincaid Grade Level	8 (6)	6 (4)	1 (−5)	4 (−1)	5 (2)	1 (−5)	1 (−5)	6 (4)
Gunning Fog Index	8 (7)	6 (3)	1 (−7)	4 (−1)	7 (4)	2 (−5)	3 (−3)	5 (2)
Smog Index	8 (7)	7 (5)	3 (−3)	4 (0)	4 (0)	2 (−5)	1 (−7)	6 (3)
Automated Readability Index	8 (4)	5 (2)	2 (−4)	3 (−3)	6 (3)	1 (−5)	4 (0)	6 (3)
Coleman Liau Index	5 (3)	5 (3)	1 (−5)	3 (−2)	5 (3)	1 (−5)	4 (0)	5 (3)
Reflection Score	3 (3)	2 (4)	7 (−3)	1 (6)	5 (0)	3 (3)	8 (−4)	5 (0)

[a]For Linguistic Quality, metrics such as BERTScore (Precision and Recall, F1), BLEURT Score, and ROUGE (1, 2, L) indicate higher is better, while for Hallucination score indicates lower is better. For Trustworthiness metrics (eg, Toxicity), lower values are better. In Accessibility, higher values are better for Flesch Reading Ease and Reflection score, while lower values are better for Flesch-Kincaid Grade Level, Coleman-Liau Index, and Gunning

XSL•FO
RenderX

Fog Index. Each cell shows the rank (score).
[b]LLMs: large language models.

Performance of LLMs in Linguistic Quality

Quantitative Evaluation

Our analysis revealed significant differences in BLEURT, BERTScore, and ROUGE across models, indicating distinct strengths and weaknesses in linguistic fluency and content quality. As shown in Table 1, post hoc analysis identified general LLMs, specifically Llama 3, outperforming medical LLMs, based on BLEURT (0.41), BERTScore Recall (0.86), and ROUGE- 1 (0.51), indicating higher linguistic quality, fluency, and relevance. Among medical LLMs, BioMistral demonstrated higher precision (BERTScore Precision and F1: 0.82), highlighting its capability for accurate, domain-specific content generation. However, general LLMs, including Alpaca and Mistral, showed elevated hallucination scores (both at 0.57), suggesting a trade-off between fluency and factuality. Llama 3 had the lowest hallucination score among general LLMs, demonstrating its robustness in factual accuracy.

Evaluation of Qualitative Content

Assessments of the qualitative content (Table 2) revealed moderate to near-perfect interrater agreement, especially for coherence (κ_w=0.82) and accuracy (κ_w=0.60). Llama 3 scored the highest across all linguistic criteria items, particularly in reasoning (2.94) and accuracy (2.92), reflecting strong factual consistency and logical structure. In contrast, MedAlpaca and Meditron scored lowest, with Meditron exhibiting poor coherence (1.13) and excessive jargon (1.57), suggesting limitations in clarity and accessibility. Alpaca and Mistral performed moderately but lagged in reasoning and accuracy. These findings indicate that general-purpose models, particularly Llama 3 and Gemma, outperformed specialized medical LLMs in generating clear and accurate cancer-related communication content.

Table . Qualitative evaluation of general-purpose and medical large language models across linguistic quality, safety or trustworthiness, and communication or accessibility dimensions[a].

Metrics	General-purpose LLMs[b]					Medical LLMs		
	Llama3	Gemma	Alpaca	Mistral	Vicuna	MedAlpaca	BioMistral	Meditron
Linguistic Quality								
Accuracy	2.92	2.70	1.53	1.48	2.11	1.48	1.18	1.35
Coherence	2.81	2.66	1.36	1.57	1.94	1.43	1.12	1.13
Jargon	2.11	1.96	1.98	1.75	1.92	1.74	1.49	1.57
Understanding	2.94	2.67	1.60	1.66	2.04	1.58	1.18	1.40
Safety and Trustworthiness								
Harm	2.96	2.76	1.62	1.54	2.11	1.51	1.18	1.42
Trust and Confidence	2.93	2.64	1.56	1.64	2.06	1.53	1.18	1.43
Communication Accessibility and Affectiveness								
Clarity & Empathy	2.89	2.64	1.70	1.59	2.05	1.59	1.18	1.41
Compassion	2.86	2.21	1.72	1.61	2.00	1.67	1.18	1.56
Cue to Action	2.82	2.28	1.58	1.53	1.93	1.50	1.18	1.40
Domain Relevance	2.94	2.68	1.60	1.68	2.09	1.60	1.18	1.43
Usability/Acceptability	2.88	2.55	1.48	1.44	1.99	1.44	1.18	1.32

[a]All scores are mean ratings on a 1 - 3 Likert scale (1=disagree, 2=Neutral, and 3=agree).
[b]LLMs: large language models.

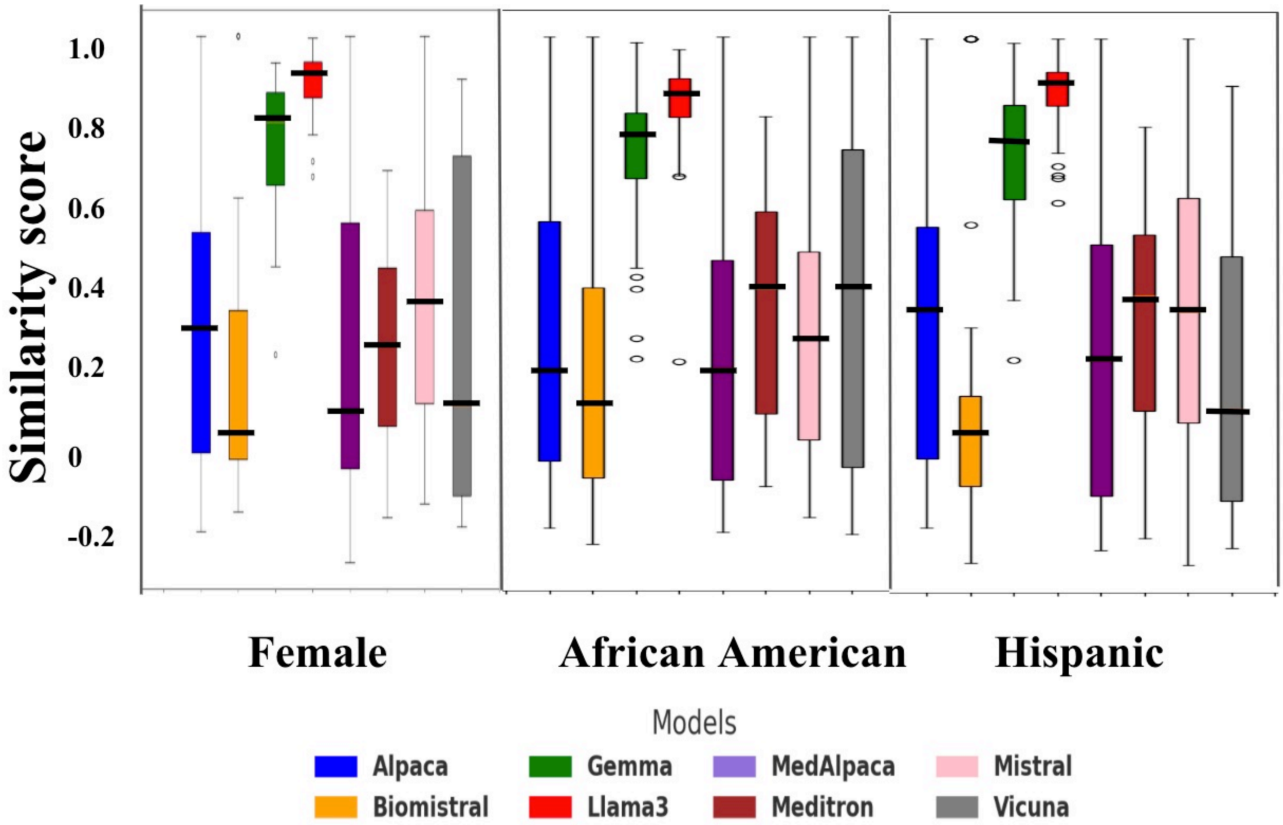
Performance of LLMs in Safety and Trustworthiness: Quantitative Evaluation

Table 1 presents toxicity and bias metrics across models. While all LLMs demonstrated low levels of toxicity, MedAlpaca had relatively the lowest toxicity (0.024), and Meditron had the highest (eg, identity attack: 0.0087). Among general-purpose models, Gemma exhibited the highest toxicity (0.038), whereas Llama 3 showed comparatively lower toxicity (0.033). To assess broader demographic biases, including race and gender, we

XSL•FO
RenderX

applied in-context impersonation for racial bias and GenBit scoring for gender bias, following prior work [25,26]. MedAlpaca showed the lowest gender bias (0.903), and Gemma the highest (1.498), followed by Llama 3 (1.43). On racial bias, Figure 3 presents similarity scores from sentenceBERT [63], showing how well LLMs maintain consistent high performance with low variability regardless of demographic context, as higher scores indicate lower bias. Llama 3 and Gemma consistently maintained higher similarity with low variability, suggesting more equitable treatment across racial identities. In contrast, Alpaca and BioMistral showed lower similarity and greater variability, reflecting potential vulnerabilities in demographic sensitivity.

Figure 3. Similarity scores between responses without context and responses with context (eg, African American, Female, and Hispanic). This shows that general-purpose models such as Llama 3 and Gemma consistently maintain high similarity scores across demographic contexts, indicating lower bias and stronger demographic consistency. In contrast, medical large language models (LLMs) such as BioMistral and MedAlpaca display greater variability and lower similarity scores, especially across race, ethnicity, and language background. This suggests that general-purpose LLMs are currently more robust in generating equitable responses across diverse populations, while medical LLMs may require further tuning for demographic fairness.



Evaluation of Qualitative Content

Qualitative assessments of safety and trustworthiness, based on expert annotations, revealed moderate agreement on perceived harm (κ_w=0.59) and trust and confidence (κ_w=0.59). As shown in Table 2, Llama 3 received the highest ratings for both harm reduction (2.96) and trustworthiness (2.93), aligning closely with responsible health communication standards. Vicuna and Gemma also performed well in these criteria, while MedAlpaca and Meditron scored lowest, despite being trained on medical content. Alpaca and Mistral showed moderate performance. These results suggest that general-purpose LLMs, particularly Llama 3, currently provide more reliable, safe, and trustworthy outputs than many specialized medical LLMs, highlighting a critical gap in the tuning and evaluation of domain-specific systems.

Performance of LLMs in Communication Accessibility and Affectiveness

Quantitative Evaluation

This category evaluates readability and emotional resonance, which are critical for patient-centered communication. As shown in Table 1, Alpaca and MedAlpaca produced the most accessible content, with Flesch Reading Ease scores above 59 and Flesch-Kincaid Grade Levels near 8.0, aligning with established guidelines for public health materials. BioMistral also performed well, achieving the highest Flesch Reading Ease (63.73) and lowest SMOG Index (5.84), although it had moderately high complexity scores on other indices. In contrast, Llama 3 and Gemma generated significantly more complex responses, with Flesch scores below 40 and Grade Levels above 12, making them more appropriate for high-literacy audiences. Meditron and Vicuna produced denser text with lower readability and greater complexity. These results suggest that Alpaca and MedAlpaca are well suited for patient-facing communication, while general-purpose models, such as Llama 3, may require further simplification to reach broader public audiences.

XSL•FO
RenderX

Evaluation of Qualitative Content

Expert ratings showed moderate to substantial agreement across clarity and empathy (κ_w=0.57), domain relevance (κ_w=0.62), and usability and applicability (κ_w=0.53). As summarized in Table 2, Llama 3 consistently outperformed across all 5 affective dimensions, including clarity and empathy (2.89), compassion (2.86), cue to action (2.82), domain relevance (2.94), and usability (2.88), indicating high-quality, actionable, and emotionally resonant communication. Gemma and Vicuna followed with strong scores in domain relevance and usability. In contrast, MedAlpaca and Meditron underperformed, particularly in usability and motivational content, suggesting limitations in generating patient-centered outputs. Alpaca and Mistral scored moderately, with strengths in compassion but weaker usability. Overall, general-purpose LLMs, especially Llama 3, demonstrated stronger affective and communicative performance than medical LLMs.

Statistical Differences and Effect Sizes Across Models

We used Welch 1-way ANOVA to test whether model performance differed significantly across all evaluation metrics (Table 3). Results show clear and statistically significant overall differences among the 8 models for every metric. Readability and content-quality measures—such as BLEURT, BERTScore, ROUGE, and classical readability indices—exhibited large effect sizes (ε^2), indicating substantial separation in linguistic quality across models. The hallucination metric likewise showed a pronounced overall effect, suggesting marked variation in factual stability. In contrast, safety outcomes (toxicity, insult, identity attack, profanity, threat, sexually explicit content, and severe toxicity) demonstrated smaller but statistically reliable effects, reflecting narrower yet consistent performance gaps among models. Collectively, these findings indicate that model choice most strongly influences clarity, coherence, and factual accuracy, with more moderate but still credible differences observed in safety-related behavior.

XSL•FO
RenderX

Table . Welch ANOVA across models for each evaluation metric[a,b].

Metrics	*F* test (*df*)	*P* value	ε²
BLEURT Score	1612.988 (7, 11831.33)	<1e-300	0.839
BertScore Precision	1091.289 (7, 12154.17)	<1e-300	0.576
BertScore Recall	438.361 (7, 12013.52)	<1e-300	0.246
BertScore F1	639.625 (7, 12024.49)	<1e-300	0.353
Rouge-1	361.108 (7, 11904.96)	<1e-300	0.205
Rouge-2	223.332 (7, 11732.70)	0.00E-02	0.13
Rouge-L	396.495 (7, 11904.96)	<1e-300	0.225
Hallucination Score	919.048 (7, 12317.09)	<1e-300	0.485
Gender Bias	17294.547 (7, 11451.45)	<1e-300	1
Toxicity Score	132.307 (7, 12270.64)	1.52E-188	0.074
Severe Toxicity	57.246 (7, 12846.45)	3.31E-81	0.031
Identity Attack	87.872 (7, 11603.12)	3.15E-125	0.052
Insult	167.374 (7, 12585.83)	1.11E-237	0.091
Profanity	66.154 (7, 12555.70)	4.07E-94	0.036
Threat	66.088 (7, 12724.36)	4.81E-94	0.036
Sexually Explicit	27.462 (7, 11935.72)	1.05E-37	0.015
Flirtation	636.676 (7, 12187.09)	<1e-300	0.347
Flesch Reading Ease	536.854 (7, 12953.81)	<1e-300	0.278
Flesch-Kincaid Grade Level	638.672 (7, 12984.04)	<1e-300	0.328
Gunning Fog Index	719.172 (7, 12807.91)	<1e-300	0.372
Smog Index	1952.647 (7, 12354.32)	<1e-300	0.955
Automated Readability Index	272.98 (7, 13078.14)	<1e-300	0.143
Coleman Liau Index	227.599 (7, 13148.51)	0.00E-02	0.119
Reflection Score	30.489 (7, 12297.43)	3.93E-42	0.017

[a]Columns report the *F* statistic, 2-sided *P* value, and epsilon-squared (ε²) effect size. Extremely small *P* values are shown in scientific notation. ε² provides the proportion of variance attributable to model differences in a 1-way (Welch) design; values closer to 1 indicate larger between-model effects. For genbit_score, ε² is capped at 1.000 for interpretability because a simple ε² approximation under Welch can slightly exceed 1 in rare cases; and the substantive conclusion (“very large effect”) is unchanged.

[b]Columns report the Welch ANOVA *F* statistic with numerator and denominator degrees of freedom within parentheses, 2-sided *P* value, and epsilon-squared effect size. The numerator degree of freedom is 7 for all tests because 8 models were compared. Denominator degree of freedom values are Welch-adjusted and therefore vary by metric. Extremely small *P* values are shown in scientific notation.

To complement the omnibus tests, we performed comparisons for every metric and report the full results. Because multiple pairwise tests increase the likelihood of false positives, we adjusted *P* values using the Benjamini-Hochberg procedure to control the false discovery rate at 0.05. After Benjamini-Hochberg adjustment, a large share of model pairs remained statistically different (average significant-pair rate 0.82, median 0.86), indicating consistent between-model separation beyond simple rank ordering. On quality metrics, leading models—particularly Llama-3 on BLEURT and BERTScore—showed meaningful effect sizes (Hedges *g*), whereas safety metrics displayed smaller but still significant gaps (adjusted *P*<.05).

Discussion

Principal Findings

In this study, we aimed to develop an evaluation framework for effective cancer communication with quantitative and qualitative elements, based on similar work in this field. Working with experts in health communication and health equity, we developed a community-centered evaluation framework that spans three main categories: (1) Linguistic Quality, (2) Safety and Trustworthiness, and (3) Communication Accessibility and Affectiveness. Our findings show that general-purpose LLMs, particularly Llama 3 and Gemma, outperformed specialized medical models in Linguistic Quality, producing more fluent and coherent responses. In contrast, medical LLMs, such as MedAlpaca and BioMistral demonstrated better communication accessibility, generating text that is easier to read at a

XSL•FO
RenderX

lower-grade level with reduced complexity. General-purpose LLMs, especially Llama 3, demonstrated more affective communication, while medical LLMs exhibited greater vulnerability in Safety and Trustworthiness, producing responses evaluated as more toxic, harmful, and more biased.

General-purpose models such as Llama 3 and Gemma outperformed medical LLMs in fluency, coherence, and factual accuracy. Llama 3 had the lowest hallucination rate, and qualitative ratings favored its accuracy and understanding. Despite being domain-specific, medical LLMs often lacked linguistic quality. Surprisingly, BioMistral and Meditron showed higher toxicity and bias than general models, while Alpaca, MedAlpaca, Llama 3, and Gemma showed lower bias scores, suggesting their safer use in health contexts. Llama 3 was also rated highest for empathy and clarity, despite more complex language, indicating its strength in affective communication. In contrast, medical LLMs such as MedAlpaca generated simpler, more readable outputs suitable for public health.

Specialized medical LLMs, although fine-tuned for health care, underperformed in safety, coherence, and affectiveness, raising concerns for clinical use. Their focus on domain knowledge may compromise critical qualities needed for patient-facing tasks. To address this, future work should embed clinical communication standards (eg, empathy and clarity) and integrate external knowledge representations to improve recall, precision, and scalability [65,66]. Hybrid neurosymbolic approaches are recommended for safer and more clinically robust outputs [67,68].

Limitations

The limitations of this study include the use of only open-source models and benchmark datasets, which may not reflect proprietary systems or real patient interactions. Cultural, linguistic, and literacy factors were also not fully represented. Our hallucination metric, based on key-term variation, serves as a practical proxy for factual inconsistency but does not replace source-grounded verification. Future work will integrate retrieval-based fact-checking and uncertainty scoring to enhance robustness. Our ethical assessment relied on proxy measures (toxicity detectors, impersonation-based fairness deltas, and hallucination or error checks) rather than real-user or clinical deployment, so residual risks (eg, nuanced stigma, privacy, or safety impacts) may not be fully captured [69]. As a next step, we plan to extend this work through simulated patient-clinician interaction studies using the same codebook and web interface developed for this project to evaluate usability, empathy, and real-world communication flow.

Conclusions

This study evaluates how LLMs communicate breast and cervical cancers information, focusing on linguistic quality, safety, and affectiveness. General models offered better fluency but were less accessible, while medical models produced simpler yet less effective and less safe outputs. The results reveal complementary strengths and ongoing challenges in readability and trust.

Acknowledgments

Generative artificial intelligence (AI) tools were used only for grammar and language editing during the preparation of this manuscript. No generative AI tool was used to create the study concept, methodology, analysis, interpretation of results, or substantive manuscript content. The authors wrote the manuscript and remain fully responsible for the accuracy, originality, and integrity of all content.

Funding

This work is funded by Microsoft Accelerating Foundation Models Research Program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of Microsoft.

Data Availability

The datasets used in this study are publicly available. Further information and materials related to the datasets may be obtained from the corresponding author upon reasonable request.

Authors' Contributions

AS contributed to the conceptualization, methodology design, formal analysis, investigation, data curation, software development, original draft writing, and visualization. VC contributed to conceptualization, methodology, investigation, formal analysis, validation, and writing – review and editing. ADR contributed conceptualization, formal analysis, and validation. UK contributed to conceptualization, methodology, investigation, resources, writing – review and editing, supervision, project administration, and funding acquisition. MYI contributed to supervision, project administration, resources, writing – review and editing, and funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Sample question from the curated dataset.
[PDF File, 86 KB - cancer_v12i1e82971_app1.pdf]

References

1. Cancer facts & figures 2023. American Cancer Society. 2023. URL: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2023/2023-cancer-facts-and-figures.pdf [accessed 2026-05-15]
2. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin 2023 Jan;73(1):17-48. [doi: 10.3322/caac.21763] [Medline: 36633525]
3. Warner ET, Tamimi RM, Hughes ME, et al. Time to diagnosis and breast cancer stage by race/ethnicity. Breast Cancer Res Treat 2012 Dec;136(3):813-821. [doi: 10.1007/s10549-012-2304-1] [Medline: 23099438]
4. Moore JX, Andrzejak SE, Jones S, Han Y. Exploring the intersectionality of race/ethnicity with rurality on breast cancer outcomes: SEER analysis, 2000-2016. Breast Cancer Res Treat 2023 Feb;197(3):633-645. [doi: 10.1007/s10549-022-06830-x] [Medline: 36520228]
5. Key statistics for cervical cancer. American Cancer Society. 2024. URL: https://www.cancer.org/cancer/types/cervical-cancer/about/key-statistics.html [accessed 2025-05-11]
6. Olusola P, Banerjee HN, Philley JV, Dasgupta S. Human papilloma virus-associated cervical cancer and health disparities. Cells 2019 Jun 21;8(6):622. [doi: 10.3390/cells8060622] [Medline: 31234354]
7. Moore de Peralta A, Holaday B, Hadoto IM. Cues to cervical cancer screening among U.S. Hispanic women. Hisp Health Care Int 2017 Mar;15(1):5-12. [doi: 10.1177/1540415316682494] [Medline: 28558513]
8. Spencer JC, Kim JJ, Tiro JA, et al. Racial and ethnic disparities in cervical cancer screening from three U.S. healthcare settings. Am J Prev Med 2023 Oct;65(4):667-677. [doi: 10.1016/j.amepre.2023.04.016] [Medline: 37146839]
9. Consedine NS, Magai C, Spiller R, Neugut AI, Conway F. Breast cancer knowledge and beliefs in subpopulations of African American and Caribbean women. Am J Health Behav 2004;28(3):260-271. [doi: 10.5993/ajhb.28.3.7] [Medline: 15152885]
10. Best AL, Vamos C, Choi SK, Thompson EL, Daley E, Friedman DB. Increasing routine cancer screening among underserved populations through effective communication strategies: application of a health literacy framework. J Canc Educ 2017 Jun;32(2):213-217. [doi: 10.1007/s13187-017-1194-7]
11. Yalamanchili A, Sengupta B, Song J, et al. Quality of large language model responses to radiation oncology patient care questions. JAMA Netw Open 2024 Apr 1;7(4):e244630. [doi: 10.1001/jamanetworkopen.2024.4630] [Medline: 38564215]
12. Musheyev D, Pan A, Gross P, et al. Readability and information quality in cancer information from a free vs paid Chatbot. JAMA Netw Open 2024 Jul 1;7(7):e2422275. [doi: 10.1001/jamanetworkopen.2024.22275] [Medline: 39058491]
13. Chen D, Huang RS, Jomy J, et al. Performance of multimodal artificial intelligence chatbots evaluated on clinical oncology cases. JAMA Netw Open 2024 Oct 1;7(10):e2437711. [doi: 10.1001/jamanetworkopen.2024.37711] [Medline: 39441598]
14. Goh E, Gallo R, Hom J, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. JAMA Netw Open 2024;7(10). [doi: 10.1001/jamanetworkopen.2024.39466]
15. Chow JCL, Li K. Developing effective frameworks for large language model-based medical chatbots: insights from radiotherapy education with ChatGPT. JMIR Cancer 2025 Feb 18;11(1):e66633. [doi: 10.2196/66633] [Medline: 39965195]
16. Chow JCL, Li K. Large language models in medical chatbots: opportunities, challenges, and the need to address AI risks. Information 2025;16(7):549. [doi: 10.3390/info16070549]
17. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016 May 24;3(1):160035. [doi: 10.1038/sdata.2016.35] [Medline: 27219127]
18. Grilo A, Marques C, Corte-Real M, Carolino E, Caetano M. Assessing the quality and reliability of ChatGPT's responses to radiotherapy-related patient queries: comparative study with GPT-3.5 and GPT-4. JMIR Cancer 2025 Apr 16;11(1):e63677. [doi: 10.2196/63677] [Medline: 40239208]
19. Zitu MM, Le TD, Duong T, Haddadan S, Garcia M, Amorrortu R, et al. Large Language Models in Cancer: Potentials, Risks, and Safeguards: Oxford University Press; 2025. [doi: 10.1093/bjrai/ubae019]
20. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health 2020 Apr 2;41(1):433-451. [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
21. Abbasian M, Khatibi E, Azimi I, et al. Foundation metrics for evaluating effectiveness of healthcare conversations powered by generative AI. NPJ Digit Med 2024 Mar 29;7(1):82. [doi: 10.1038/s41746-024-01074-z] [Medline: 38553625]
22. Zhang T, Qiu L, Guo Q, et al. Enhancing uncertainty-based hallucination detection with stronger focus. : Association for Computational Linguistics Presented at: Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing; Dec 6-10, 2023; Singapore. [doi: 10.18653/v1/2023.emnlp-main.58]
23. Perspective API. Jigsaw. 2024. URL: https://www.perspectiveapi.com [accessed 2026-05-15]
24. Erol A, Padhi T, Saha A, Kursuncu U, Aktas ME. Playing devil's advocate: unmasking toxicity and vulnerabilities in large vision-language models. arXiv. Preprint posted online on Jan 14, 2025. [doi: 10.48550/arXiv.2501.09039]
25. Huang A, Li D, Fan Z, Chen J, Zhang W, Wu W. Long-term trends in the incidence of male breast cancer and nomogram for predicting survival in male breast cancer patients: a population-based epidemiologic study. Sci Rep 2025 Jan 15;15(1):2027. [doi: 10.1038/s41598-025-85954-8] [Medline: 39814936]

XSL•FO
RenderX

26. Anderson WF, Jatoi I, Tse J, Rosenberg PS. Male breast cancer: a population-based comparison with female breast cancer. J Clin Oncol 2010 Jan 10;28(2):232-239. [doi: 10.1200/JCO.2009.23.8162] [Medline: 19996029]
27. Sengupta K, Maher R, Groves D, Olieman C. GenBiT: measure and mitigate gender bias in language datasets. Microsoft J Appl Res 2021;16:63-71 [FREE Full text]
28. Salewski L, Alaniz S, Rio-Torto I, Schulz E, Akata Z. In-context impersonation reveals large language models' strengths and biases. arXiv. Preprint posted online on May 24, 2023. [doi: 10.48550/arXiv.2305.14930]
29. Levy S, Karver TS, Adler WD, Kaufman MR, Dredze M. Evaluating biases in context-dependent health questions. arXiv. Preprint posted online on Mar 7, 2024. [doi: 10.48550/arXiv.2403.04858]
30. Poulain R, Fayyaz H, Beheshti R. Bias patterns in the application of LLMs for clinical decision support: a comprehensive study. arXiv. Preprint posted online on Apr 23, 2024. [doi: 10.48550/arXiv.2404.15149]
31. Vela MB, Erondu AI, Smith NA, Peek ME, Woodruff JN, Chin MH. Eliminating explicit and implicit biases in health care: evidence and research needs. Annu Rev Public Health 2022 Apr 5;43(1):477-501. [doi: 10.1146/annurev-publhealth-052620-103528] [Medline: 35020445]
32. Salovey P, Mayer JD. Emotional Intelligence. Imagin Cogn Pers 1990 Mar;9(3):185-211. [doi: 10.2190/DUGG-P24E-52WK-6CDG]
33. Tomkins S. Affect Imagery Consciousness: Volume I: The Positive Affects: Springer Publishing Company; 1962.
34. FLESCH R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
35. Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. Derivation of new readability formulas for navy enlisted personnel. : Naval Technical Training command. Research Branch Report; 1975.
36. Gunning R. The Technique of Clear Writing: McGraw-Hill; 1952.
37. McLaughlin GH. SMOG grading—a new readability formula. J Reading 1969;12:639-646 [FREE Full text]
38. Smith EA, Senter RJ. Automated readability index. AMRL TR 1967 May:1-14. [Medline: 5302480]
39. Coleman M, Liau TL. A computer readability formula designed for machine scoring. J Appl Psychol 1975;60(2):283-284. [doi: 10.1037/h0076540]
40. Min DJ, Pérez-Rosas V, Resnicow K, Mihalcea R. PAIR: prompt-aware margin ranking for counselor reflection scoring in motivational interviewing. 2022 Presented at: Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing; Dec 2022; Abu Dhabi, United Arab Emirates p. 148-158. [doi: 10.18653/v1/2022.emnlp-main.11]
41. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature New Biol 2023 Aug;620(7972):172-180. [doi: 10.1038/s41586-023-06291-2] [Medline: 37438534]
42. Pal A, Umapathi LK, Sankarasubbu M. MedMCQA: a large-scale multi-subject multi-choice dataset for medical domain question answering. PMLR 2022;174:248-260 [FREE Full text]
43. Jin D, Pan E, Oufattole N, Weng WH, Fang H, Szolovits P. What disease does this patient have? A large-scale open domain question answering dataset from medical exams. Appl Sci (Basel) 2021;11(14):6421. [doi: 10.3390/app11146421]
44. Jeong M, Hwang H, Yoon C, Lee T, Kang J. OLAPH: improving factuality in biomedical long-form question answering. arXiv. Preprint posted online on May 21, 2024. [doi: 10.48550/arXiv.2405.12701]
45. Agichtein E, Carmel D, Pelleg D, Pinter Y, Harman D. Overview of the TREC 2015 LiveQA track. : National Institute of Standards and Technology (NIST); 2015 Presented at: Proceedings of the Twenty-Fourth Text REtrieval Conference, TREC 2015; Nov 17-20, 2015; Gaithersburg, Maryland, USA URL: http://trec.nist.gov/pubs/trec24/papers/Overview-QA.pdf [accessed 2026-05-15]
46. Abacha AB, Mrabet Y, Sharp M, Goodwin TR, Shooshan SE, Demner-Fushman D. Bridging the gap between consumers' medication questions and trusted answers. Stud Health Technol Inform 2019 Aug 21;264:25-29. [doi: 10.3233/SHTI190176] [Medline: 31437878]
47. Manes I, Ronn N, Cohen D, Ilan Ber R, Horowitz-Kugler Z, Stanovsky G. K-QA: a real-world medical Q&A benchmark. Presented at: Proceedings of the 23rd Workshop on Biomedical Natural Language Processing; Aug 16, 2024; Bangkok, Thailand p. 277-294. [doi: 10.18653/v1/2024.bionlp-1.22]
48. Li Y, Li Z, Zhang K, Dan R, Jiang S, Zhang Y. ChatDoctor: a medical chat model fine-tuned on a Large Language Model Meta-AI (LLaMA) using medical domain knowledge. Cureus 2023 Jun;15(6):e40895. [doi: 10.7759/cureus.40895] [Medline: 37492832]
49. Taori R, Gulrajani I, Zhang T, Dubois Y, Li X, Guestrin C, et al. Alpaca: a strong, replicable instruction-following model. Stanford Center for Research on Foundation Models. 2023. URL: https://crfm.stanford.edu/2023/03/13/alpaca.html [accessed 2026-05-19]
50. Dubey A, Jauhri A, Pandey A, Kadian A, Al-Dahle A, Letman A, et al. The llama 3 herd of models. arXiv. Preprint posted online on Jul 31, 2024. [doi: 10.48550/arXiv.2407.21783]
51. Jiang AQ, Sablayrolles A, Mensch A, Bamford C, Chaplot DS, Ddl C, et al. Mistral 7B. arXiv. Preprint posted online on Oct 10, 2023. [doi: 10.48550/arXiv.2310.06825]
52. Mesnard T, Hardin C, Dadashi R, Bhupatiraju S, Pathak S, et al. Gemma: open models based on Gemini research and technology. arXiv. Preprint posted online on Mar 13, 2024. [doi: 10.48550/arXiv.2403.08295]

XSL•FO
RenderX

53. Liang P, Bommasani R, Lee T, Tsipras D, Soylu D, Yasunaga M, et al. Holistic evaluation of language models. arXiv. Preprint posted online on Nov 16, 2022. [doi: 10.48550/arXiv.2211.09110]
54. Han T, Adams LC, Papaioannou JM, Grundmann P, Oberhauser T, Löser A, et al. MedAlpaca–an open-source collection of medical conversational AI models and training data. arXiv. Preprint posted online on Apr 14, 2023. [doi: 10.48550/arXiv.2304.08247]
55. Labrak Y, Bazoge A, Morin E, Gourraud PA, Rouvier M, Dufour R. BioMistral: a collection of open-source pretrained large language models for medical domains. Presented at: Findings of the Association for Computational Linguistics ACL 2024; 2024; Bangkok, Thailand and virtual meeting p. 5848-5864. [doi: 10.18653/v1/2024.findings-acl.348]
56. Chen Z, Cano AH, Romanou A, Bonnet A, Matoba K, Salvi F, et al. Meditron-70b: scaling medical pretraining for large language models. arXiv. Preprint posted online on Nov 27, 2023. [doi: 10.48550/arXiv.2311.16079]
57. Leong HY, Gao YF, Shuai J, Zhang Y, Pamuksuz U. E icient fine-tuning of large language models for automated medical documentation. arXiv. Preprint posted online on Sep 14, 2024. [doi: 10.48550/arXiv.2409.09324]
58. Welch BL. On the comparison of several mean values: an alternative approach. Biometrika 1951 Dec;38(3/4):330. [doi: 10.2307/2332579]
59. Games PA, Howell JF. Pairwise multiple comparison procedures with unequal N's and/or variances: a Monte Carlo study. J Educat Stat 1976;1(2):113. [doi: 10.2307/1164979]
60. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. J Educ Behav Stat 1981 Jun;6(2):107-128. [doi: 10.3102/10769986006002107]
61. Wu H, Leung SO. Can Likert scales be treated as interval scales?—a simulation study. J Soc Serv Res 2017 Aug 8;43(4):527-532. [doi: 10.1080/01488376.2017.1329775]
62. Simply put: a guide for creating easy-to-understand materials. Centers for Disease Control and Prevention. 2009. URL: https://www.cdc.gov/healthliteracy/pdf/simply_put.pdf [accessed 2026-05-15]
63. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977 Jun;33(2):363-374. [doi: 10.2307/2529786] [Medline: 884196]
64. Li M, Gao Q, Yu T. Kappa statistic considerations in evaluating inter-rater reliability between two raters: which, when and context matters. BMC Cancer 2023;23(1):799. [doi: 10.1186/s12885-023-11325-z]
65. Xu R, Cui H, Yu Y, et al. Knowledge-infused prompting: assessing and advancing clinical text data generation with large language models. arXiv. Preprint posted online on Nov 1, 2023. [doi: 10.48550/arXiv.2311.00287]
66. Khandelwal V, Gaur M, Kursuncu U, Shalin VL, Sheth AP. A domain-agnostic neurosymbolic approach for big social data analysis: evaluating mental health sentiment on social media during COVID-19. : Institute of Electrical and Electronics Engineers Inc; 2024 Presented at: 2024 IEEE International Conference on Big Data, BigData 2024; Dec 15-18, 2024; Washington, USA p. 959-968. [doi: 10.1109/BigData62323.2024.10825174]
67. Meng H, Lin Z, Yang F, Xu Y, Cui L. Knowledge distillation in medical data mining: a survey. Presented at: 5th International Conference on Crowd Science and Engineering; Oct 16-18, 2021; Jinan, China p. 175-182. [doi: 10.1145/3503181.3503211]
68. Garg R, Padhi T, Jain H, Kursuncu U, Kumaraguru P. Just KIDDIN: knowledge infusion and distillation for detection of indecent memes. arXiv. Preprint posted online on Nov 19, 2024. [doi: 10.48550/arXiv.2411.12174]
69. Chow JCL, Li K. Ethical considerations in human-centered AI: advancing oncology chatbots through large language models. JMIR Bioinform Biotech 2024;5:e64406. [doi: 10.2196/64406]

Abbreviations

κ_w: weighted Cohen κ
LLM: large language model
ROUGE: Recall-Oriented Understudy for Gisting Evaluation

Edited by A Schwartz, M Balcarras, N Cahill; submitted 25.Aug.2025; peer-reviewed by JCL Chow, M Qiu, Y Guo; revised version received 27.Oct.2025; accepted 22.Dec.2025; published 26.Jun.2026.

Please cite as:
Saha A, Churchill V, Rodriguez AD, Kursuncu U, Idris MY
Large Language Models for Breast and Cervical Cancers Communication: Mixed Methods Evaluation Study Assessing Linguistic Quality, Safety, and Accessibility
JMIR Cancer 2026;12:e82971
URL: https://cancer.jmir.org/2026/1/e82971
doi:10.2196/82971

XSL•FO
RenderX

© Agnik Saha, Victoria Churchill, Anny D Rodriguez, Ugur Kursuncu, Muhammed Y Idris. Originally published in JMIR Cancer (https://cancer.jmir.org), 26.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Generative AI Chatbots as Digital Adjuncts for Sexual Health Information After Prostate Cancer in Men Who Have Sex With Men: Auto-Netnographic Study

Mats Christiansen[1,2*], RN, PGD, MMSc, PhL; Henrik Eriksson[3*], RN, RNT, PhD; Lisbeth Fagerström[2*], RN, MSc, PhD

[1]Department of Public Health and Caring Science, Uppsala University, BMC, Husargatan 3, Uppsala, Sweden

[2]Department of Natural and Health Sciences, The Faculty of Science and Engineering, Åbo Akademi University, Vaasa, Finland

[3]Section for Health Promotion and Care Sciences, University West, Trollhättan, Sweden

[*]all authors contributed equally

Corresponding Author:
Mats Christiansen, RN, PGD, MMSc, PhL
Department of Public Health and Caring Science, Uppsala University, BMC, Husargatan 3, Uppsala, Sweden

Abstract

Background: Sexual health concerns following prostate cancer treatment are common yet often insufficiently addressed in clinical practice, particularly among men who have sex with men. These individuals may face additional barriers stemming from heteronormative assumptions, limited disclosure, and a lack of culturally tailored information. As generative artificial intelligence (GenAI) chatbots become increasingly accessible, patients are using these systems to seek sensitive health information outside traditional care settings. While prior research has focused on the accuracy and safety of chatbot-generated health advice, less attention has been paid to how responses are framed and experienced in sexual minority contexts.

Objective: This study aimed to describe and compare how 4 GenAI chatbots respond to questions about sexual health following prostate cancer treatment, with a focus on the needs of a gay man, and to interpret these responses using netnographic and actor-network theory perspectives.

Methods: A qualitative exploratory study using auto-netnography was conducted. In February–March 2025, the first author interacted once with 4 widely used GenAI chatbots—ChatGPT (GPT-4o; Open AI), Claude (3.5 Sonnet; Anthropic), Copilot (GPT-4 Turbo; Microsoft), and Gemini (2.0 Flash; Google)—while assuming the role of a simulated "mock patient." Two standardized prompts were used verbatim across all platforms: an initial prompt addressing sexual health concerns after prostate cancer treatment and a supplementary prompt focusing on sexual minority–specific issues, including same-sex practices. Chatbot outputs were treated as system-generated data and analyzed qualitatively, integrating system-generated text with reflexive experiential engagement and attention to interactional framing, emotional attunement, specificity, and performative features. The analysis did not assess clinical effectiveness, safety, or generalizability.

Results: Across platforms, chatbot responses addressed treatment-related sexual health concerns using generally inclusive language, with variation in emotional tone, specificity, and cultural sensitivity. Interactional features included the scope and framing of clinical information, encouragement of dialogue, self-care advice, and explicit discussion of same-sex sexual practices. No obvious fabricated claims were identified; however, contextual inaccuracies were observed. Responses were mapped along 2 intersecting continua—logical-to-empathetic orientation and general-to-specific framing—yielding 4 interactional styles: structured overview, rational clarity, compassionate perspective, and compassionate precision. This 4-quadrant framework served as an interpretive heuristic and does not constitute an evaluation of quality or effectiveness.

Conclusions: The findings indicate that contemporary GenAI chatbots, when used as digital adjuncts, may enact communication styles that can be perceived as supportive, culturally sensitive, and LGBTQI+ (lesbian, gay, bisexual, transgender, queer, and intersex)-inclusive in specific sexual health interactions. Although these systems lack ethical consciousness and cannot replace professional care, their performative responses may complement clinical practice by facilitating reflection and access to sensitive information. The study highlights how care-like meanings may emerge through sociomaterial interactions between users and artificial intelligence systems rather than demonstrating generalized performance or clinical reliability.

(JMIR Cancer 2026;12:e81745) doi:10.2196/81745

KEYWORDS

artificial intelligence; auto-netnography; generative AI chatbots; netnography; prostate cancer; sexual minority men

XSL•FO
RenderX

Introduction

Sexual health concerns following prostate cancer treatment are common yet often insufficiently addressed in clinical practice, particularly among men who have sex with men. Sexual minority patients may face additional challenges stemming from heteronormative assumptions, limited disclosure to health care providers, and a lack of culturally tailored information and support [1-5]. As a result, men who have sex with men treated for prostate cancer may experience unmet informational and psychosocial needs related to intimacy, sexual practices, and identity-specific concerns.

Digital technologies have become increasingly prominent sources of health information, supplementing traditional clinical encounters. Alongside websites, forums, and telehealth services, generative artificial intelligence (GenAI) chatbots are now widely accessible to patients seeking health-related guidance. These conversational systems can generate responsive, personalized text and simulate dialogic engagement, making them particularly attractive for sensitive topics, such as sexual health. At the same time, concerns have been raised about the accuracy, bias, and reliability of chatbot-generated health information in both clinical and lay contexts [6-10].

Most existing research on GenAI in health contexts has focused on evaluating the correctness, safety, and technical performance of chatbot-generated outputs, often through expert benchmarking or comparison with established clinical guidelines [9,10]. While such studies are essential, they offer limited insight into how health-related responses are framed, enacted, and experienced in interaction, particularly in sensitive, identity-linked domains, such as sexual health. Less is known about how GenAI chatbots address sexuality, intimacy, and same-sex practices after prostate cancer treatment, or how interactional styles may vary across platforms when responding to identical prompts.

Qualitative and nursing-oriented research has emphasized that health communication involves not only the accuracy of information but also interactional framing, emotional attunement, and relational context [11-13]. From sociomaterial and actor–network perspectives, digital technologies can be understood as nonhuman actors that participate in enacting meanings and practices through interaction rather than merely transmitting information [14-17]. This perspective foregrounds how health-related meanings emerge relationally through specific configurations of users, technologies, and contexts.

In nursing science, particularly within the Nordic caritative caring tradition, caring is understood as an ethical and relational practice grounded in responsibility, dignity, and communion rather than solely in information provision or technical support [18,19]. This distinction offers an important reference point for interpreting artificial intelligence (AI)–mediated interactions that may resemble caring communication yet do not constitute care in a caritative sense. Building on this tradition, Andtfolk [20] has examined the possibilities and limitations of care technologies, emphasizing that digital and robotic systems may simulate aspects of caring interaction without ethical responsibility, consciousness, or the capacity for genuine caring communion.

Rather than evaluating whether GenAI chatbots provide accurate or safe medical advice, this study examines how responses from GenAI chatbots are framed, enacted, and experienced in a situated sexual health interaction with a gay man after prostate cancer treatment. By adopting an auto-netnographic approach, the study attends to both system-generated text and the researcher's reflexive, experiential engagement with the interaction [13,21]. This approach enables exploration of interactional features, such as emotional attunement, specificity, and inclusivity without making claims about clinical effectiveness or generalizability.

The aim of this study was to describe and compare how 4 GenAI chatbots respond to questions about sexual health following prostate cancer treatment, with a focus on the needs of a gay man, and to interpret these responses using netnographic and actor–network theory perspectives.

Methods

Study Design

This study used an auto-netnographic qualitative design to explore situated interactions with GenAI chatbots [15,17,21]. Auto-netnography was chosen to enable reflexive engagement with chatbot responses and to examine how interactional framing, tone, and performativity are enacted in a specific health-related inquiry. The study followed the iterative netnographic process described by Kozinets, including initiation, immersion and interaction, analytic integration, and reflexive interpretation [17]. The study was not designed to evaluate clinical effectiveness, safety, or generalizable performance of chatbots, but rather to explore how responses are produced and experienced in a particular interactional context.

This study was reported in accordance with Standards for reporting qualitative research (SRQR) and the Chatbot Health Advice Reporting Tool (CHART) reporting guideline [22,23].

Chatbot Selection and Study Context

Four widely used GenAI chatbots were included: ChatGPT (GPT-4o, free version; Open AI), Claude (Claude 3.5 Sonnet, free version; Anthropic), Copilot (GPT-4 Turbo, enterprise subscription; Microsoft), and Google Gemini (2.0 Flash, free version). These chatbots were selected for their public availability and relevance to contemporary health information-seeking. All interactions were conducted via the chatbots' web interfaces between February and March 2025, and no technical limitations were encountered during data collection.

Interactions were carried out using a simulated "mock patient" persona enacted by the first author, drawing on his lived and professional experience as a gay man and a specialist nurse. This methodological approach was used to elicit relevant chatbot responses and does not constitute data from a real patient. Accordingly, chatbot outputs were treated as system-generated data rather than human participant data.

Prompting and Interaction Procedure

Data collection began with a standardized initial prompt presented verbatim to all 4 chatbots (Textbox 1). The prompt

XSL•FO
RenderX

described a gay man in his early fifties, recently diagnosed with prostate cancer, seeking information on how different treatment options might affect sexual health, including concerns about erections, orgasm, intimacy, and sex with a male partner.

Textbox 1. Standardized initial prompt.

I'm a 53-year-old gay man, and I've been diagnosed with prostate cancer. Will be treated with either radiotherapy or surgery, both would likely include hormonal therapy as well. I'm concerned about my sex life after treatment. Can you help me think this through?

Following the initial prompt, each chatbot generated a response and, in some cases, asked brief clarification questions (eg, about specific concerns or preferences). Follow-up replies were limited to brief clarifications (eg, one-sentence responses) and did not introduce new topics beyond the prompts. These follow-up exchanges were neither standardized nor the primary focus of the analysis.

After the initial interaction, a second standardized supplementary prompt was presented verbatim across all platforms (Textbox 2). This prompt explicitly addressed sexual minority–specific considerations, including anal sex, insertive and receptive roles, partner communication, and whether health care providers typically address these topics.

Textbox 2. Standardized supplementary prompt

Are there any specific things I need to consider as a gay man regarding my sex life after treatment for prostate cancer?

The standardized prompts used across all 4 chatbots are presented verbatim.

No modifications were made to the wording of either prompt across platforms. All interactions occurred in a single session per chatbot. Chatbot responses to both standardized prompts were captured verbatim and transferred to a spreadsheet for analysis. No responses were edited, filtered, or corrected prior to analysis. The analytic focus was on how responses were framed, enacted, and performed in this specific interactional context, rather than on response stability, reproducibility, or clinical correctness.

Positionality and Reflexivity

The first author identifies as a gay man and is a specialist nurse with longstanding clinical and research experience in sexual health and prostate cancer. The coauthors served as project supervisors; HE has expertise in netnographic and qualitative research, and LF has expertise in analytic and interpretive methods. This positionality informed the selection of prompts and the analytic focus. Reflexive field notes were recorded during and after interactions, capturing immediate impressions and evolving interpretations. Preliminary analyses were discussed among all authors to support reflexivity and analytic transparency.

Analytic Approach

The analysis was qualitative and interpretive. Consistent with auto-netnographic methodology, the first author's experiential engagement with the interaction was integral to the analytic material [21]. Rather than using a formal, predefined thematic coding framework or quantification, the analysis focused on how responses from GenAI chatbots were framed, how emotional attunement and specificity were enacted, and how interactional styles varied across platforms. The analysis was informed by netnographic principles and actor–network theory, emphasizing enactment and performativity in human–technology interaction [14-17]. Auto-netnography was used to capture the lived, situated experience of interacting with GenAI chatbots, including affective responses, interpretive judgments, and reflexive meaning-making that extend beyond the textual content of chatbot outputs [15,21].

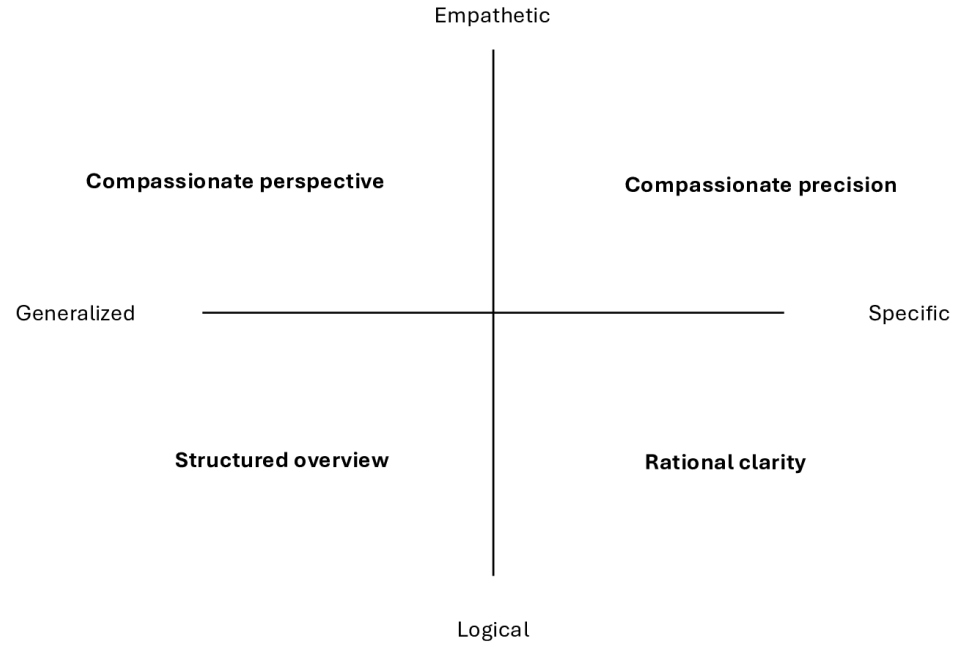
While the chatbot responses constitute system-generated text, the analytic material also included the researcher's experiential engagement with the interaction, which informed the interpretation of tone, responsiveness, and perceived relevance in context. The analysis began with a naïve reading of all prompt responses to gain an immediate overall understanding of the material. In the next step, responses were placed side by side to identify similarities and differences. As part of this comparative phase, a sentence-level content analysis was performed through line-by-line coding, enabling a more granular examination of interactional features and thematic patterns.

Initial line-by-line coding focused on interactional features, including tone, emotional attunement, specificity, and forms of guidance. These features were compared across chatbot responses to identify recurring contrasts, which were abstracted into 2 intersecting dimensions, namely empathetic–rational orientation and general–specific framing. The 4-quadrant analytic map was then developed as a heuristic to visualize and interpret how different interactional styles were enacted across platforms.

During the comparative phase, the theoretical framework and the first author's preunderstanding were mobilized to pose reflexive questions about the prompt responses, consistent with contemporary netnographic and auto-netnographic approaches to AI-mediated interaction [15,21]. Two questions guided the analysis: (1) What does this mean from a theoretical lens? and (2) How can I understand this (as a gay man)? These questions served as reflexive lenses for heuristic interpretation across both disciplinary and experiential horizons. This process revealed substantial variation in how the prompts addressed their intended recipients, in terms of both language and tone. To organize these comparative interpretations, the 4-quadrant analytic map (Figure 1) was used as an interpretive lens to present the study's findings, rather than as a classificatory or evaluative model.

XSL•FO
RenderX

Figure 1. An interpretive heuristic for describing chatbot interactional styles.



Assessments of accuracy, inclusivity, and empathy were pragmatic and reflexive. Accuracy was defined pragmatically as the absence of obvious factual errors and general alignment with established clinical knowledge familiar to the authors, rather than as a formally validated assessment. No formal scoring instruments, interrater reliability testing, or expert panel validation were used. Inclusivity was pragmatically assessed by using nonheteronormative assumptions, gender-neutral partner language, and explicit acknowledgment of same-sex practices when prompted.

Ethical Considerations

The study did not involve human participants. Chatbot outputs were treated as system-generated data. The use of a mock patient persona was a methodological strategy to elicit relevant responses, not a form of deception involving human participants. Ethical approval was obtained from the Swedish Ethical Review Authority (Dnr 2024-02924-02).

Results

The 4 GenAI-powered chatbots responded to the prompts with text addressing the specific areas outlined in the prompts (Multimedia Appendix 1). All chatbots' GenAI responses included follow-up questions or invitations to continue the dialogue. The findings include the following themes: (1) the content of the chatbot replies, (2) expressions of empathy in chatbot responses, (3) encouraging dialogue, (4) providing self-care advice, (5) discussing same-sex sexual practices, and (6) tonality and cultural sensitivity.

Content of the Chatbot Replies

Across all platforms, chatbot responses addressed treatment-related sexual health concerns using inclusive language, with variation in emotional tone and specificity. Gemini's response cited 4 published scientific papers, all of which are available with open access. Across platforms, the responses addressed commonly reported treatment-related sexual health concerns. No prominent examples of hallucinated content were identified in the analyzed outputs. However, some errors were present, such as links to support groups that mistakenly listed Swedish Hospital Services, a health service in the United States, rather than Sweden. Although occasional caveats noted that the support groups were based in the United States or Australia, it was also mentioned that some could offer online support conversations.

Building on broader discussions linking sexual health to prostate cancer treatments, a second prompt was added to examine the specific sexual issues faced by a gay man undergoing treatment (refer to Multimedia Appendix 2).

> *It's important to acknowledge that while the core medical aspects of prostate cancer treatment are the same for all men, there are specific considerations for gay men regarding their sexual health. Here's a breakdown of those key points:...* [Gemini]

XSL•FO
RenderX

Expressions of Empathy in Chatbot Responses

After the initial prompt, all chatbots responded in ways that conveyed empathy, compassion, and concern through language and tone, and an appreciation for the user's sharing of personal information. The responses used language that conveyed empathy and engagement:

> *I appreciate you sharing this with me. You're facing a big decision, and it's completely understandable to be concerned about how treatment might affect your sex life.* [ChatGPT]
>
> *I'm really sorry to hear about your diagnosis. It's completely understandable to be concerned about how treatment might affect your sex life. Let's break down some of the potential impacts and ways to manage them.* [Copilot]

Both of these excerpts included language commonly associated with empathic communication. However, some responses may appear more detached, emphasizing facts:

> *It's completely understandable that you're concerned about how prostate cancer treatment might affect your sex life. This is a common and valid concern for many men facing this diagnosis. Here's a breakdown of factors to consider, and how to approach this:…* [Gemini]

Encouraging Dialogue

The chatbots promoted interaction by asking open-ended questions to better engage the user and gather more detailed responses for consideration:

> *Would it be helpful to discuss specific strategies for maintaining intimacy during periods when hormone therapy might affect your libido or sexual function? Many men find creative ways to stay sexually active and maintain connections with partners even when dealing with these challenges.* [Claude]
>
> *Would you like to talk through the potential effects of each treatment in more detail? Or are you more interested in strategies for maintaining a fulfilling sex life post-treatment?* [ChatGPT]

In these responses, follow-up questions expanded the scope of topics addressed.

Providing Self-Care Advice

There was also advice on self-care in the responses. Several of the chatbots suggested pelvic floor exercises to improve orgasm quality:

> *Enhancing Orgasms:*
>
> *Kegel Exercises: Strengthening pelvic floor muscles may help improve orgasm intensity and control leakage.*
>
> *Mindful Arousal Techniques: Engaging in sensual touch, breathwork, and fantasy can enhance pleasure.*
>
> *Lubrication: Can help with sensation changes and dryness.* [ChatGPT]

Discussing Same-Sex Sexual Practices

Regarding same-sex sexual practices (Multimedia Appendix 2), all chatbots provided explicit descriptions of potential changes related to these practices.

> *Anal Sex:*
>
> *If you engage in receptive anal sex, this can generally be resumed after healing from either surgery or radiation (usually 4 - 6 wk for surgery, timing varies for radiation).*
>
> *If you engage in insertive anal sex, erectile challenges may affect this more significantly than some other forms of sexual activity.*
>
> *Some men find they need to adjust their preferred sexual roles based on treatment effects.* [Claude]

Tonality and Cultural Sensitivity

However, the chatbots varied in tone and vocabulary. Gemini and Copilot maintained a professional, somewhat distant tone, while ChatGPT adopted a friendlier, more collegial style. The language reflected nonheteronormative sexual practices; for example, ChatGPT mentioned "Navigating hookup culture," which allows responses sensitive to different cultural backgrounds and recognizes that not everyone is in a committed or monogamous relationship.

Claude and Copilot asked about current partners to encourage communication and support:

> *Also, do you have a partner who's part of these discussions? Partners can be valuable allies in recovery and adaptation* [Claude]
>
> *Openly discuss any changes in your sexual function or desires with your partner. This can help maintain intimacy and understanding in your relationship* [Copilot]

All the chatbots consistently used gender-neutral, inclusive language.

Care-Related Language Interpreted Through a 4-Quadrant Analytic Map

The enactment and performativity of GenAI can be mapped along two continua: (1) one from general to specific information and (2) the other from logical to empathetic responses. The 4-quadrant framework should be understood as an interpretive heuristic rather than a validated typology. It was derived from a comparative reading of chatbot responses and is intended to support interpretation of how interactional styles may vary along two intersecting dimensions, including degree of emotional attunement (logical to empathetic) and degree of specificity (general to tailored).

The interactor was prompted to provide additional information to give the chatbot broader context for engagement. In this case, the chatbot resembled an information trader, which is at the logical end of the spectrum. The compassionate aspect of GenAI places it closer to the empathetic end.

To illustrate how interactional variation was interpreted in the analysis, four analytically derived positions are described below.

XSL•FO
RenderX

1. Structured overview: Responses are characterized by logically organized general information about treatments and side effects, with limited emotional engagement or personalization.
2. Compassionate perspective: Responses acknowledge emotional concerns and the broader life context while remaining relatively general in scope.
3. Compassionate precision: Responses combine emotional attunement with tailored, context-sensitive guidance, including explicit references to same-sex practices or minority stressors.
4. Rational clarity: Responses emphasize factual explanations, references, and clinical reasoning, with minimal affective language (Figure 1).

Three of the GenAI chatbots began with compassionate responses, expressing concern and saying they were "sorry" to hear of the diagnosis. The fourth adopted a more distant or professional tone:

> *It's completely understandable that you're concerned about how prostate cancer treatment might affect your sex life. This is a common and valid concern for many men facing this diagnosis. Here's a breakdown of factors to consider, and how to approach this.* [Gemini]

In the excerpt above, the chatbot acknowledges the interactor's concern but shows less empathy. This approach is more logical and provides a structured overview, yet it is more general and does not address the interactor's specific needs.

The chatbots played a performative role in engaging with and questioning the interactor, prompting them to provide additional details. They offered information while also seeking clarification and any further necessary data. Here, the interactional pattern resembles descriptions of person-centered communication, demonstrating compassionate precision, for example, when the chatbots used inclusive language and suggested practices for different sexual positions.

> *Yes, there are several important considerations specific to gay men dealing with prostate cancer treatment:*
>
> *Anal Sex:*
>
> *If you engage in receptive anal sex, this can generally be resumed after healing from either surgery or radiation (usually 4 - 6 wk for surgery, timing varies for radiation)*
>
> *If you engage in insertive anal sex, erectile challenges may affect this more significantly than some other forms of sexual activity*
>
> *Some men find they need to adjust their preferred sexual roles based on treatment effects.* [Claude]
>
> *Your sex life will likely change, but that doesn't mean it has to end. Many men discover new ways to experience pleasure, whether through adapting physically, redefining intimacy, or exploring different forms of touch and connection. If you'd like, I can help you find LGBTQ+-friendly doctors, sex therapists, or resources specific to your area.* [ChatGPT]

Some of the chatbots attempted to provide evidence-based responses by citing open-access scientific journal papers. This approach was less compassionate and focused on presenting facts, references, and links from external sources. This illustrates rational clarity.

> *Yes, there are some specific considerations for gay men regarding their sex life after prostate cancer treatment. Here are a few key points:*
>
> *Communication with Healthcare Providers*
>
> *[Copilot]*

This compassionate perspective offered encouragement and suggestions for addressing these issues, such as:

> *Specific Considerations for Gay MenResearch indicates that gay and bisexual men might experience different challenges and lower health-related quality of life after prostate cancer treatment compared to heterosexual men4 (a provided ref.). It's important to find healthcare providers who are knowledgeable and sensitive to these differences.* [Copilot]

However, some of the care described resembled a structured summary, making it harder to distinguish differences from a compassionate perspective.

Discussion

Principal Findings

This study explored how 4 GenAI chatbots responded to a situated sexual health inquiry framed from the perspective of a gay man undergoing treatment for prostate cancer. Across platforms, chatbot responses varied in emotional attunement, degree of specificity, and attention to sexual minority–specific concerns, while consistently addressing treatment-related sexual issues and encouraging continued dialogue. Interactional features included inclusive language, provision of self-care advice, and explicit discussion of same-sex sexual practices. These variations were analytically organized using a 4-quadrant heuristic that maps emotional orientation and degree of specificity, providing an interpretive lens for understanding interactional diversity rather than for evaluating clinical performance or quality.

The findings indicate that when men who have sex with men seek health information following prostate cancer surgery, GenAI chatbots may differ not only in tone and communicative style but also in how interactional responses are organized and enacted. These differences were observed along the analytically derived dimensions of emotional attunement and informational specificity, shaping how guidance was framed and relationally positioned within the interaction. Such variation reflects different ways of presenting and contextualizing self-care–related information in sensitive health contexts, rather than stable or intrinsic properties of the systems themselves.

Building on existing research, studies on GenAI in health contexts have primarily focused on assessing the accuracy, safety, and technical performance of chatbot-generated

XSL•FO
RenderX

information [9,10]. While these evaluations are essential, they offer limited insight into how health-related responses are framed and enacted in interaction, particularly in sensitive domains, such as sexual health. Consistent with emerging qualitative and nursing-oriented research on AI-mediated communication [11,12,24,25], the present findings suggest that GenAI chatbots may simulate supportive dialogue through tone, structure, and responsiveness, despite lacking ethical agency and clinical responsibility.

Importantly, the findings do not suggest that GenAI chatbots provide care in an ethical or caritative sense. Drawing on Nordic caritative caring theory [18,19] and prior analyses of caring encounters in technologically mediated contexts [20,26,27], the observed responses are better understood as care-like enactments. In this view, caring extends beyond the delivery of correct information and encompasses relational and contextual dimensions that may be partially simulated through language and interactional form. As Andtfolk [20] has noted, digital systems may perform aspects of care without consciousness, moral responsibility, or the capacity for genuine caring communion.

From a sociomaterial and actor–network perspective, these care-like enactments emerge from specific configurations of actors, practices, and technologies rather than residing within the chatbot itself [14-17]. Chatbot responses are shaped by prompts, platform design, training data, and user expectations rather than by stable or intrinsic properties of the system. Consequently, sexual health after prostate cancer treatment is not communicated as a single, stable phenomenon but is enacted differently depending on how concerns are articulated and responded to in particular human–technology interactions.

When considered alongside nursing theories of care, this sociomaterial understanding also aligns with self-care perspectives, such as Orem's [28] theory, which emphasizes supporting individuals' capacity to reflect on and manage health-related needs. In this study, GenAI chatbots did not provide care in an ethical or caritative sense [18,19], but their care-like enactments may be understood as digitally mediated supports for reflection, sense-making, and self-management within a broader care ecology. In this way, GenAI chatbots can be conceptualized as digital adjuncts that intersect with, but do not replace, professional care.

These considerations are particularly relevant for men who have sex with men following prostate cancer treatment. The findings highlight both the potential and the limitations of GenAI chatbots as digital adjuncts in this context. Sexual minority patients have documented unmet needs in sexual health, intimacy, and disclosure after treatment [1-5]. Several chatbot responses demonstrated sensitivity to same-sex practices and sexual minority–specific concerns, aligning with frameworks for LGBTQI+ (lesbian, gay, bisexual, transgender, queer, and intersex)-inclusive health care that emphasize recognition of sexual minority identities, practices, and relational contexts [29]. However, this sensitivity was inconsistently applied across platforms, underscoring that inclusivity in AI-mediated health communication is variable rather than assured by design.

The interactional patterns observed in chatbot responses can also be interpreted through the PLISSIT (Permission, Limited Information, Specific Suggestions, and Intensive Therapy) model, which outlines progressive levels of sexual health communication [30]. Across platforms, chatbot responses consistently operated within the first three levels of the model. They invited questions and normalized concerns about sexuality (Permission), provided general information about treatment-related sexual effects (Limited Information), and, in some cases, offered tailored guidance on same-sex practices and self-care strategies (Specific Suggestions). Notably, none of the chatbots reached the level of Intensive Therapy, which would require individualized clinical assessment, ethical responsibility, and professional accountability. This pattern aligns with interpreting chatbot responses as care-like enactments rather than clinical care. Furthermore, it supports conceptualizing GenAI chatbots as digital adjuncts that may facilitate dialogue and reflection on sexual health without substituting for professional nursing or therapeutic expertise.

Strengths and Limitations

This study is based on a single, situated interaction per chatbot and does not aim to assess reproducibility, safety, or generalizability. The auto-netnographic approach involves interpretive judgment and reflexive engagement, which may introduce subjectivity [13,21]. Although no obviously fabricated claims were identified in this dataset, this absence should not be interpreted as evidence of accuracy, reliability, or safety, as prior studies have documented hallucinations and contextual errors in similar systems [9,25].

Conclusions

The findings suggest that contemporary GenAI chatbots, when used as digital adjuncts, may enact communication styles perceived as supportive, culturally sensitive, and LGBTQI+-inclusive in specific sexual health interactions following prostate cancer treatment. Although these systems lack ethical consciousness and cannot replace professional care [24], their performative responses may complement clinical practice by facilitating reflection, dialogue, and access to sensitive information within a broader nursing and self-care context. Rather than demonstrating generalized performance or clinical reliability, this study highlights how care-like meanings may emerge in specific sociomaterial interactions between users and AI systems.

Funding

This study was supported by a research award of US $800 awarded to MC by Sigma Nursing, Alpha Eta Chapter, University of California, San Francisco (UCSF). The funder had no role in the study design, data collection, analysis, or interpretation, manuscript preparation, or the decision to submit the manuscript for publication.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Prompts and initial answers on sexual health after prostate cancer treatment, specifically for men who have sex with men.
[DOCX File, 26 KB - cancer_v12i1e81745_app1.docx]

Multimedia Appendix 2
Prompts and specific answers on sexual health after prostate cancer treatment, specifically for men who have sex with men.
[DOCX File, 26 KB - cancer_v12i1e81745_app2.docx]

References

1. Rosser BRS, Polter EJ, Talley KMC, et al. Health disparities of sexual minority patients following prostate cancer treatment: results from the Restore-2 study. Front Oncol 2022;12:812117. [doi: 10.3389/fonc.2022.812117] [Medline: 35186749]
2. Wells-Prado DR, Ross MW, Rosser BRS, et al. Prostate cancer disclosure and sexual orientation: understanding outness to healthcare providers as a situational or consistent phenomenon. Patient Educ Couns 2022 Jul;105(7):2033-2037. [doi: 10.1016/j.pec.2021.11.017] [Medline: 34865891]
3. Yazdanpanah O, Benjamin DJ, Rezazadeh Kalebasty A. Prostate cancer in sexual minorities: epidemiology, screening and diagnosis, treatment, and quality of life. Cancers (Basel) 2023 May 8;15(9):2654. [doi: 10.3390/cancers15092654] [Medline: 37174119]
4. Alexis O, Garbrah W, Worsley AJ. The unmet needs of gay and bisexual men following prostate cancer treatment: an integrative literature review. J Cancer Surviv 2025 Nov 19. [doi: 10.1007/s11764-025-01934-x] [Medline: 41254405]
5. Vedovo F, Lonati C, Suardi N, et al. Sexual outcomes in men who have sex with men who underwent radical prostatectomy. Prostate Cancer Prostatic Dis 2025 Mar;28(1):70-80. [doi: 10.1038/s41391-024-00861-9] [Medline: 38918583]
6. Lee JG, Reissmann ME, Parkhomenko E, Wang DS. Evaluating the quality of online health information about prostate cancer treatment. Urol Pract 2021 May;8(3):360-366. [doi: 10.1097/UPJ.0000000000000213] [Medline: 37145667]
7. Li JZH, Kong T, Killow V, et al. Quality assessment of online resources for the most common cancers. J Cancer Educ 2023 Feb;38(1):34-41. [doi: 10.1007/s13187-021-02075-2] [Medline: 34365589]
8. Giakas JA, Zaliznyak M, Kohut-Jackson A, et al. Quality and readability of online health information on common urologic cancers: Assessing barriers to health literacy in urologic oncology. Urol Pract 2024 Jul;11(4):670-676. [doi: 10.1097/UPJ.0000000000000574] [Medline: 38899676]
9. Menz BD, Modi ND, Abuhelwa AY, et al. Generative AI chatbots for reliable cancer information: evaluating web-search, multilingual, and reference capabilities of emerging large language models. Eur J Cancer 2025 Mar 11;218:115274. [doi: 10.1016/j.ejca.2025.115274] [Medline: 39922126]
10. Altorfer FCS, Kelly MJ, Avrumova F, et al. The double-edged sword of generative AI: surpassing an expert or a deceptive “false friend”? Spine J 2025 Aug;25(8):1635-1643. [doi: 10.1016/j.spinee.2025.02.010] [Medline: 40049450]
11. Armstrong N, Powell J. Patient perspectives on health advice posted on Internet discussion boards: a qualitative study. Health Expect 2009 Sep;12(3):313-320. [doi: 10.1111/j.1369-7625.2009.00543.x] [Medline: 19555377]
12. Ronquillo CE, Peltonen LM, Pruinelli L, et al. Artificial intelligence in nursing: priorities and opportunities from an international invitational think-tank of the Nursing and Artificial Intelligence Leadership Collaborative. J Adv Nurs 2021 Sep;77(9):3707-3717. [doi: 10.1111/jan.14855] [Medline: 34003504]
13. Salzmann-Erikson M, Eriksson H. A mapping review of netnography in nursing. Qual Health Res 2023 Jul;33(8-9):701-714. [doi: 10.1177/10497323231173794] [Medline: 37192601]
14. Quinton S, Reynolds N. In: Kozinets RV, Gambetti R, editors. Netnography in Human and Non-Human Networked Sociality, 1st edition: Routledge; 2020:251-267.
15. Kozinets RV. Researching AI chatbots, platforms and the metaverse: understanding today’s netnography. In: Belk RW, Otnes C, editors. Handbook of Qualitative Research Methods in Marketing, 2nd edition: Edward Elgar Publishing; 2024:185-196. [doi: 10.4337/9781035302727.00025]
16. Lugosi P, Quinton S. More-than-human netnography. J Mark Manag 2018 Feb 12;34(3-4):287-313. [doi: 10.1080/0267257X.2018.1431303]
17. Kozinets RV, Gretzel U. Netnography evolved: new contexts, scope, procedures and sensibilities. Ann Tour Res 2024 Jan;104:103693. [doi: 10.1016/j.annals.2023.103693]
18. Eriksson K, editor. Vårdvetenskap. vetenskapen om vårdandet: om det tidlösa i tiden [Book in Swedish]. : Liber; 2018.
19. Lindström U, Nyström LL, Zetterlund JE. Theory of caritative caring. In: Alligood MR, editor. Nursing Theorists and Their Work, 9th edition: Elsevier; 2017:279-321.
20. Andtfolk M. The Possibilities for Using Humanoid Robots as a Care Resource [Dissertation]. : Åbo Akademi University; 2022 URL: https://www.doria.fi/handle/10024/185321 [accessed 2026-02-02]

XSL•FO
RenderX

21. Howard L. Auto-netnography in education: unfettered and unshackled. In: Kozinets RV, Gambetti R, editors. Netnography Unlimited: Understanding Technoculture Using Qualitative Social Media Research, 1st edition: Routledge; 2020:217-240. [doi: 10.4324/9781003001430-18]
22. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
23. CHART Collaborative, Huo B, Collins GS, et al. Reporting guideline for chatbot health advice studies: The CHART statement. JAMA Netw Open 2025 Aug 1;8(8):e2530220. [doi: 10.1001/jamanetworkopen.2025.30220] [Medline: 40747871]
24. Jackson D. Retaining human connectedness in increasingly AI-driven environments. J Adv Nurs 2025 Dec;81(12):8391-8393. [doi: 10.1111/jan.70041]
25. O'Connor S, Vercell A, Wong D, et al. The application and use of artificial intelligence in cancer nursing: a systematic review. Eur J Oncol Nurs 2024 Feb;68:102510. [doi: 10.1016/j.ejon.2024.102510] [Medline: 38310664]
26. Sunder AJ, Heden L, van Dulmen S, Acuna Mora M, Carlsson Lalloo E. The COMCARE framework for person-centred communication: a practical caring framework. Scand J Caring Sci 2025 Sep;39(3):e70074. [doi: 10.1111/scs.70074] [Medline: 40619709]
27. Tunks Leach KJ, Barker ME, Hales C, et al. Perceptions of empathic and compassionate healthcare encounters: an international exploratory qualitative descriptive study. Scand J Caring Sci 2025 Jun;39(2):e70006. [doi: 10.1111/scs.70006] [Medline: 40241327]
28. Orem DE. Nursing: Concepts of Practice, 6th edition: Mosby; 2001.
29. Roman Laporte RL, De Santis JP. Concept analysis: LGBT+ inclusive health care. ANS Adv Nurs Sci 2023;46(4):410-423. [doi: 10.1097/ANS.0000000000000467] [Medline: 36728300]
30. Annon JS. The PLISSIT model: a proposed conceptual scheme for the behavioral treatment of sexual problems. J Sex Educ Ther 1976 Apr;2(1):1-15. [doi: 10.1080/01614576.1976.11074483]

Abbreviations

AI: artificial intelligence
CHART: Chatbot Health Advice Reporting Tool
GenAI: generative artificial intelligence
LGBT/LGBT+/LGBTQ/LGBTQI: lesbian, gay, bisexual, transgender, queer, and intersex
PLISSIT: Permission, Limited Information, Specific Suggestions, and Intensive Therapy
SRQR: Standards for Reporting Qualitative Research

Edited by M Balcarras, N Cahill; submitted 04.Aug.2025; peer-reviewed by N Döring, S Chen; revised version received 07.Jan.2026; accepted 07.Jan.2026; published 09.Feb.2026.

Please cite as:
Christiansen M, Eriksson H, Fagerström L
Generative AI Chatbots as Digital Adjuncts for Sexual Health Information After Prostate Cancer in Men Who Have Sex With Men: Auto-Netnographic Study
JMIR Cancer 2026;12:e81745
URL: https://cancer.jmir.org/2026/1/e81745
doi:10.2196/81745

© Mats Christiansen, Henrik Eriksson, Lisbeth Fagerström. Originally published in JMIR Cancer (https://cancer.jmir.org), 9.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patients' Attitudes and Beliefs Toward Artificial Intelligence Use in Cancer Care: Cross-Sectional Survey Study

Marco Santos Teles[1], BS; Karolina Bryl[1], PhD; Susan Chimonas[2], PhD; Atif Khan[3], MD; Andrew S Epstein[4], MD; Bobby Daly[5], MD; Han Xiao[6], MD; Jun J Mao[1], MD

[1]Integrative Medicine Service, Memorial Sloan Kettering Cancer Center, 205 East 64th Street, New York, NY, United States

[2]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[3]Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[4]Gastrointestinal Oncology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[5]Thoracic Oncology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[6]Genitourinary Oncology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, United States

Corresponding Author:
Jun J Mao, MD
Integrative Medicine Service, Memorial Sloan Kettering Cancer Center, 205 East 64th Street, New York, NY, United States

Abstract

Background: Artificial intelligence (AI) is being rapidly integrated into oncologic care, yet little is known about how patients perceive these applications. Understanding patient perceptions is critical to ensuring AI applications align with their needs and preferences.

Objective: This study aimed to evaluate oncology patients' attitudes and beliefs on the use of AI across clinical touchpoints in cancer care.

Methods: We conducted a cross-sectional survey study with adult oncology patients from September to December 2024. The survey assessed patients' comfort with AI use across 8 clinical touchpoints of cancer care (eg, screening, diagnosis, treatment) on a 5-point Likert scale (1=very uncomfortable to 5=very comfortable). Patients also rated their concerns about AI, including potential harms related to its use (eg, medical errors, privacy breaches), on a 3-point Likert scale (1=not concerned to 3=very concerned).

Results: Of 383 patients approached, 330 (86.2% response rate) participated; 184 (55.9%) were male, 162 (49.4%) were aged 65 years or older, 35 (10.8%) were Black, 40 (12.1%) were Hispanic or Latino, and 233 (72.6%) were actively receiving cancer treatment. Patients were most comfortable with AI use in cancer screening (80.2%) and supportive care applications, including exercise (78.2%), diet (74.8%), and herbs/supplements (72.4%). Patients were least comfortable with AI use to assist with diagnosis (70.4%), symptom management (67.5%), treatment planning (64.8%), and prognosis (61.5%). Nonetheless, about half (49.7%) were at least somewhat concerned with the use of AI in cancer care, most commonly about the loss of human interaction and medical errors.

Conclusions: Although the majority of oncology patients had a favorable view of AI in cancer care, nearly half had concerns about potential harms. Incorporating patient perspectives into AI development is essential for patient-centered and high-quality cancer care.

(JMIR Cancer 2026;12:e81346) doi:10.2196/81346

KEYWORDS

patient perspectives; artificial intelligence; cancer care; digital health; patient attitudes

Introduction

Artificial intelligence (AI) is transforming oncology, offering unprecedented capabilities to analyze vast amounts of cancer treatment data and improve care [1,2]. The integration of this technology into cancer care has centered on clinical touchpoints within the "cancer continuum," such as diagnosis, prognosis, and treatment [2,3]. AI is already capable of performing tasks across these domains, including analyzing radiographic images, predicting clinical outcomes, and assisting in decision-making [4-6].

Despite these technological advances, little is known about how patients perceive the use of AI in cancer care. Prior research has focused on the views of the general population toward AI use in health care [7-9]. Although these studies suggest that patients are generally receptive to AI, data investigating views regarding its use in oncology are lacking. As the use of AI in oncology increases, understanding patient perceptions is

XSL•FO
RenderX

essential to ensure their needs and preferences are incorporated into the implementation of AI applications. Failure to incorporate these perspectives may lead to AI tools that patients are unwilling to use, ultimately eroding trust in health care, jeopardizing the clinician-patient relationship, and reducing adherence to clinical recommendations.

To address this critical need for patient-centered implementation of AI, we conducted a cross-sectional survey with adult oncology patients at a large urban academic cancer center to evaluate their attitudes and beliefs on the use of AI in clinical touchpoints of cancer care.

Methods

Questionnaire Development

The survey was designed through a multistep process in collaboration with patients. First, a literature review informed question development on AI applications in cancer care, with some questions adapted from prior AI-in-medicine surveys [7,10,11]. A patient and family advisory council at our institution provided feedback on the survey draft. We then conducted cognitive interviews with 12 oncology patients in an outpatient clinic to assess language clarity and question relevance. The survey was then revised to incorporate patient-identified topics, including AI as a physician-supervised tool (vs standalone use), its clinical applications in oncology, and potential harms. Finally, we piloted the revised survey with 21 patients to test the survey recruitment feasibility in our clinical setting.

The finalized survey (Multimedia Appendix 1) consisted of 48 items. Participants rated their comfort with AI use across 8 clinical touchpoints of cancer care: screening, diagnosis, treatment, prognosis, symptom management, diet, herbs/supplements, and exercise. They also rated their concern with 5 potential AI use harms: privacy breaches, medical errors, loss of human interaction with doctors, difficulty understanding the technology to access care, and reinforcing discrimination in health care. Question formats included multiple choice, 5-point Likert scales (1=very uncomfortable to 5=very comfortable), and 3-point Likert scales (1=not concerned to 3=very concerned). Sociodemographic, clinical, and technology use data were collected. The survey was translated into Spanish by a certified interpreter.

Survey Administration

This study used convenience clinical sampling. We distributed the survey to patients at affiliated outpatient clinics in urban academic, urban underserved, and suburban settings. Adult patients with a history of cancer were eligible to participate. Clinic staff screened patients and notified the research staff while patients waited for appointments or treatment. The research team introduced the study and provided patients with an iPad for the survey. The survey was distributed via Research Electronic Data Capture (REDCap) [12].

Data Analysis

We followed American Association for Public Opinion Research reporting guidelines. Chi-square tests were used for subgroup analysis. Statistical Package for the Social Sciences (version 28; IBM Corp) was used for statistical analysis.

Ethical Considerations

The study was reviewed by the institutional review board at Memorial Sloan Kettering Cancer Center and determined to be exempt from review under 45 Code of Federal Regulations (CFR) 46.104(d)(2)(i), which applies to research involving educational tests, survey procedures, interview procedures, or observation of public behavior when participant identities cannot readily be ascertained. The exemption determination was issued on September 5, 2024 (protocol number X24-033). Given the anonymous nature of the survey, written informed consent was not required. Participants were presented with an electronic consent statement at the start of the survey, and consent was obtained by participants selecting “I agree” before proceeding. No compensation was provided to participants.

Results

Between September and December 2024, 383 patients were approached, of whom 330 participated (86% response rate). Of these 330 participants, 184 (55.9%) were male, 162 (49.4%) were aged 65 years or older, 35 (10.8%) were Black, and 40 (12.1%) were Hispanic or Latino (Table 1). The most commonly reported cancers were prostate (114/330, 34.5%) and breast (87/330, 26.4%) cancer, and the majority of patients were actively receiving treatment (233/321, 72.6%).

XSL•FO
RenderX

Table . Participants' demographics and baseline characteristics (N=330).

Characteristic		Participants, n[a] (%)
Age (years)		
	18 - 49	58 (17.7)
	50 - 64	108 (32.9)
	≥65	162 (49.4)
Gender		
	Male	184 (55.9)
	Female	141 (42.9)
	Other	4 (1.2)
Race		
	White	234 (72.4)
	Black/African American	35 (10.8)
	Asian	25 (7.7)
	Other	29 (9)
Ethnicity		
	Hispanic	40 (12.4)
	Non-Hispanic	283 (87.6)
Education		
	High school degree or less	73 (22.6)
	College degree or more	250 (77.4)
Cancer type[b]		
	Prostate	114 (34.5)
	Breast	87 (26.4)
	Gastrointestinal	52 (15.8)
	Genitourinary	33 (10)
	Lung	22 (6.7)
	Other	66 (20)
Treatment status		
	Receiving treatment	233 (72.6)
	Not receiving treatment	88 (27.4)
Treatment modality[b]		
	Surgery	158 (47.9)
	Radiation	180 (54.5)
	Chemotherapy	209 (63.3)
	Immunotherapy	64 (19.4)
	Other	60 (18.2)
Self-rated general health		
	Excellent to good	272 (83.7)
	Fair to poor	53 (16.3)

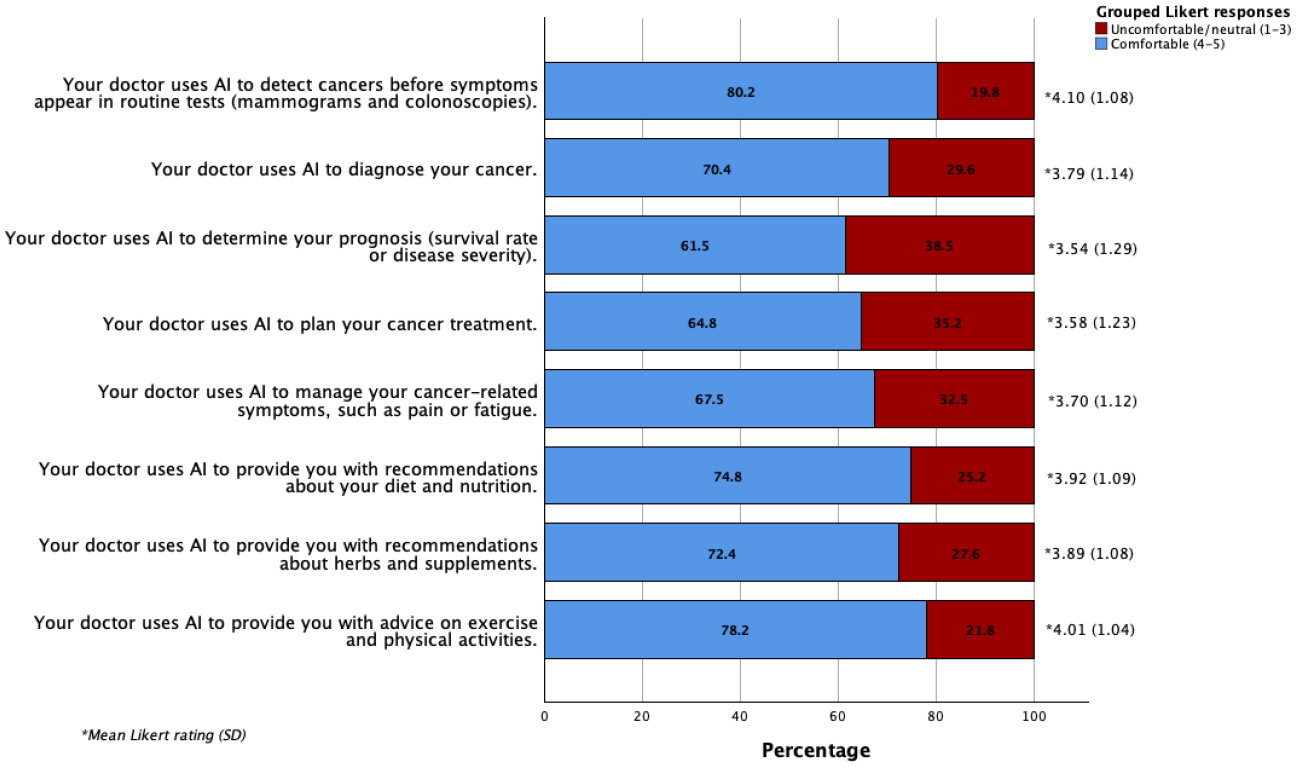
[a]Counts may not add up to 330 due to missing data (<5% missing data for each question).

[b]This is a "select all that apply" question. Patients could report multiple cancers/treatment modalities.

XSL•FO
RenderX

Most participants (300/330, 91%) reported knowing "a little" or "quite a lot" about AI, and 60 of 312 participants (19.2%) used web-based AI applications (eg, ChatGPT) at least once per week. The majority (264/328, 80.5%) believed that AI would make cancer care somewhat or much better over the next 5 years. Patients were most comfortable with AI use in cancer screening (264/329, 80.2%) and supportive care applications including recommendations for exercise (254/325, 78.2%), diet (247/330, 74.8%), and herbs/supplements (236/326, 72.4%). Patients were least comfortable with AI use to assist with diagnosis (231/328, 70.4%), symptom management (222/329, 67.5%), treatment planning (210/324, 64.8%), and prognosis (201/327, 61.5%) (Figure 1).

Figure 1. Patients' comfort level with AI use in different clinical touchpoints of cancer care. AI: artificial intelligence.



Nevertheless, 49.7% (164/330) were somewhat concerned with the use of AI in their cancer care. The most common concerns were loss of human interaction with doctors (162/330, 49.1%) and medical errors (158/330, 47.9%), followed by privacy breaches (141/330, 42.7%), difficulty understanding the technology (132/330, 40%), and reinforcing health care disparities (112/330, 33.9%).

Compared to patients who reported limited AI use, those who used it weekly were more likely to believe AI would improve cancer care (57/60, 95% vs 194/25, 77.3%; *P*=.006) and were less concerned about its use (18/60, 30% vs 135/252, 53.6%; *P*=.001). Compared to patients with less than a college degree, those with a college degree or higher were more likely to believe that AI would improve cancer care (53/73, 72.6% vs 207/248, 83.5%; *P*=.04), but not more likely to be concerned about its use (36/73, 49.3% vs 124/250, 49.6%; *P*=.97). Treatment status was not associated with differences in perceptions. Patients on active treatment were no more likely than those not on active treatment to believe that AI would positively impact cancer care (191/233, 82% vs 65/86, 75.6%; *P*=.20) or to express concerns about its use (114/233, 48.9% vs 44/88, 50%; *P*=.86).

Discussion

Principal Findings

AI holds significant potential to improve oncologic care, though successful integration of this technology must incorporate patients' perspectives, as they are the ultimate decision-makers in their care. This study highlights opportunities and challenges to optimal AI implementation in cancer care.

Although most oncology patients were receptive to AI applications, nearly half expressed concerns about potential harms, particularly loss of human interactions and medical errors. These findings align with prior studies in the general population, reflecting "cautious optimism"—patients are open to AI but emphasize the need for oversight to mitigate risks [7-9]. The most common concerns relate to previously raised ethical implications: (1) AI jeopardizing clinician-patient relationships leading to depersonalized care, and (2) overreliance on clinical support tools causing inaccurate diagnoses and recommendations [13]. These concerns echo efforts in the regulatory landscape, including the US Food and Drug Administration's evolving oversight of AI-based medical software and the EU AI Act, which designates most medical AI systems as "high-risk," requiring additional oversight measures [14]. Policymakers and developers must engage

XSL•FO
RenderX

patients during AI implementation to ensure safeguards are in place to address these concerns.

Our findings demonstrate a higher proportion of patients that were comfortable with AI than previously reported. For example, a 2019 survey of the general population found 55.4% believed AI would improve health care; in our study, 80.5% of oncology patients had this expectation [7]. Similarly, a 2022 survey found 39% of US adults were comfortable with their health care providers relying on AI for their medical care [15], while in our study, at least 60% of patients were comfortable with their doctors using AI for various clinical tasks in cancer care. These discrepancies may reflect increasing familiarity with AI in the last few years. Our findings suggest exposure to this technology may reduce patient concerns. Additionally, we observed that patients with a college degree or higher were more likely to have positive views toward the future of AI in cancer care. This is in line with another recent study that showed patients with higher education levels expressed less discomfort with AI and fewer concerns about AI technologies in clinical settings [16]. Future interventions should explore whether targeted exposures for older or less educated patients could narrow the digital divide in clinical settings.

Limitations

A limitation of this study is reliance on convenience sampling of patients at a single center in the Northeastern United States, making the sample less representative of the US cancer patient population. Furthermore, while we had a robust response rate, in-person responses may have introduced social desirability bias, as patients may have expressed more favorable views due to perceived health care expectations. The education level of our participants (77.4% with a college degree or more) was higher than the national average [17] and potentially associated with more positive views toward AI in cancer care, which may impact the generalizability of our findings.

Conclusion

Although most oncology patients had positive views of AI applications in cancer care, almost 50% had concerns, especially about loss of interaction with clinicians and potential medical errors. To optimize patient-centered care, policymakers and developers must incorporate patient input during AI implementation and address these concerns. Future research should seek to identify strategies to foster collaboration between developers and patients. Importantly, patient perceptions of AI may shift over time as public discourse and clinical exposure to AI tools evolve; therefore, a longitudinal study may better assess how these views change over time as AI becomes more integrated into oncology.

Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: refinement, correction, editing, or formatting of the manuscript to improve clarity of language. The GAI tool used was ChatGPT-4.5. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This work was supported in part by the JDJ Foundation, AKTIV Against Cancer, a National Institutes of Health (NIH)/National Cancer Institute (NCI) Cancer Center Support Grant (P30 CA008748), and the Translational and Integrative Medicine Research Fund at Memorial Sloan Kettering Cancer Center. BD is supported by a grant from the Emerson Collective and an NCI 1P50 CA271357-01.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: MST (lead), KB (equal), JJM (equal)
Data curation: MST
Formal analysis: MST (lead), KB (supporting)
Investigation: MST (lead), KB (equal)
Methodology: MST (lead), KB (equal), SC (equal), JJM (equal)
Project administration: KB
Resources: JJM
Writing: MST (lead), KB (equal), JJM (supporting), BD (supporting), AK (supporting), HX (supporting), SC (supporting)

Conflicts of Interest

JJM: grants from Tibet CheeZheng Tibetan Medicine Co Ltd and Zhongke Health International LLC. AK: stock and other ownerships with Novavax and Xtrava Health; research funding from Clovis Oncology, Merck KGaA, Varian Medical Systems.

XSL•FO
RenderX

BD: stock and other ownerships with CVS Health, Lilly, Roche/Genentech; consulting or advisory role with Varian Medical Systems and I-MAB Biopharma. ASE: honoraria from HMP Education; other relationship with UpToDate. All other authors declare no conflicts.

Multimedia Appendix 1
Survey on artificial intelligence in cancer care delivery.
[DOC File, 126 KB - cancer_v12i1e81346_app1.doc]

References

1. Kann BH, Thompson R, Thomas CR, Dicker A, Aneja S. Artificial intelligence in oncology: current applications and future directions. Oncology (Williston Park) 2019 Feb 15;33(2):46-53. [Medline: 30784028]
2. Kann BH, Hosny A, Aerts H. Artificial intelligence for clinical oncology. Cancer Cell 2021 Jul 12;39(7):916-927. [doi: 10.1016/j.ccell.2021.04.002] [Medline: 33930310]
3. Shreve JT, Khanani SA, Haddad TC. Artificial intelligence in oncology: current capabilities, future opportunities, and ethical considerations. Am Soc Clin Oncol Educ Book 2022 Apr;42:1-10. [doi: 10.1200/EDBK_350652] [Medline: 35687826]
4. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 2017 Dec 12;318(22):2199-2210. [doi: 10.1001/jama.2017.14585] [Medline: 29234806]
5. She Y, Jin Z, Wu J, et al. Development and validation of a deep learning model for non-small cell lung cancer survival. JAMA Netw Open 2020 Jun 1;3(6):e205842. [doi: 10.1001/jamanetworkopen.2020.5842] [Medline: 32492161]
6. Somashekhar SP, Sepúlveda MJ, Puglielli S, et al. Watson for Oncology and breast cancer treatment recommendations: agreement with an expert multidisciplinary tumor board. Ann Oncol 2018 Feb 1;29(2):418-423. [doi: 10.1093/annonc/mdx781] [Medline: 29324970]
7. Khullar D, Casalino LP, Qian Y, Lu Y, Krumholz HM, Aneja S. Perspectives of patients about artificial intelligence in health care. JAMA Netw Open 2022 May 2;5(5):e2210309. [doi: 10.1001/jamanetworkopen.2022.10309] [Medline: 35507346]
8. Richardson JP, Smith C, Curtis S, et al. Patient apprehensions about the use of artificial intelligence in healthcare. NPJ Digit Med 2021 Sep 21;4(1):140. [doi: 10.1038/s41746-021-00509-1] [Medline: 34548621]
9. Young AT, Amara D, Bhattacharya A, Wei ML. Patient and general public attitudes towards clinical artificial intelligence: a mixed methods systematic review. Lancet Digit Health 2021 Sep;3(9):e599-e611. [doi: 10.1016/S2589-7500(21)00132-1] [Medline: 34446266]
10. Lennartz S, Dratsch T, Zopfs D, et al. Use and control of artificial intelligence in patients across the medical workflow: single-center questionnaire study of patient perspectives. J Med Internet Res 2021 Feb 17;23(2):e24221. [doi: 10.2196/24221] [Medline: 33595451]
11. 2023 Canadian Digital Health Survey: innovative technologies and artificial intelligence. Canadian Health Infoway. 2023. URL: https://insights.infoway-inforoute.ca/2023-innovative-technologies-and-ai [accessed 2026-01-15]
12. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
13. Morley J, Machado CCV, Burr C, et al. The ethics of AI in health care: a mapping review. Soc Sci Med 2020 Sep;260:113172. [doi: 10.1016/j.socscimed.2020.113172]
14. van Kolfschooten H, van Oirschot J. The EU Artificial Intelligence Act (2024): implications for healthcare. Health Policy 2024 Nov;149:105152. [doi: 10.1016/j.healthpol.2024.105152] [Medline: 39244818]
15. Tyson A, Pasquini G, Spencer A, Funk C. 60% of americans would be uncomfortable with provider relying on AI in their own health care. Pew Research Center. URL: https://www.pewresearch.org/wp-content/uploads/sites/20/2023/02/PS_2023.02.22_AI-health_REPORT.pdf [accessed 2026-01-19]
16. Erul E, Aktekin Y, Danışman FB, et al. Perceptions, attitudes, and concerns on artificial intelligence applications in patients with cancer. Cancer Control 2025;32:10732748251343245. [doi: 10.1177/10732748251343245] [Medline: 40407404]
17. Irwin V, Wang K, Jung J, et al. Report on the condition of education 2024 (NCES 2024-144). National Center for Education Statistics. 2024. URL: https://nces.ed.gov/pubs2024/2024144.pdf [accessed 2026-01-19]

Abbreviations

AI: artificial intelligence
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

Edited by M Balcarras, N Cahill; submitted 26.Jul.2025; peer-reviewed by E Erul, Z Yu; revised version received 15.Dec.2025; accepted 22.Dec.2025; published 04.Feb.2026.

Please cite as:
Santos Teles M, Bryl K, Chimonas S, Khan A, Epstein AS, Daly B, Xiao H, Mao JJ
Patients' Attitudes and Beliefs Toward Artificial Intelligence Use in Cancer Care: Cross-Sectional Survey Study
JMIR Cancer 2026;12:e81346
URL: https://cancer.jmir.org/2026/1/e81346
doi:10.2196/81346

© Marco Santos Teles, Karolina Bryl, Susan Chimonas, Atif Khan, Andrew S Epstein, Bobby Daly, Han Xiao, Jun J Mao. Originally published in JMIR Cancer (https://cancer.jmir.org), 4.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging Innovative Electronic Health Record Data to Characterize Social Determinants of Health Among Survivors of Cancer in Persistent Poverty Areas: Cross-Sectional Study

Shiori Tanaka[1], PhD; Laura Q Rogers[1], MD; Lisa Zubkoff[1], PhD; Lori Bateman[1], PhD; Wendy Demark-Wahnefried[2], PhD; Gabrielle Rocque[1], PhD; Sejong Bae[1], PhD; Maria Pisu[1], PhD; James Booth[1], MD; Carrie R Howell[1], PhD

[1]Department of Medicine, University of Alabama at Birmingham, 1717 11th Ave South, Birmingham, AL, United States

[2]Department of Nutrition Sciences, University of Alabama at Birmingham, Webb 518, 1720 2nd Avenue South, Birmingham, AL, United States

Corresponding Author:
Carrie R Howell, PhD
Department of Medicine, University of Alabama at Birmingham, 1717 11th Ave South, Birmingham, AL, United States

Abstract

Background: Residents in persistent poverty areas experience higher cancer mortality due to social determinants of health that negatively affect multiple factors, including health behaviors.

Objective: This study aimed to characterize demographic, clinical, and social determinant of health factors among survivors of cancer in persistent poverty areas using electronic health record (EHR) data—including an embedded social risk screener and natural language processing (NLP) of social work notes—to inform community-engaged adaptation of lifestyle interventions.

Methods: EHR data from a large multispecialty group practice were extracted for patients with cancer residing in zip codes inclusive of persistent poverty areas targeted for a health behavior intervention and receiving care between January 2018 and November 2023. Self-reported social determinant of health data were obtained using the Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences (PRAPARE) and through NLP of social histories from a social work visit.

Results: We identified 2672 unique patients with cancer, of whom 578 (21.6%) had PRAPARE data and 1597 (59.8%) had social history data available for analysis. The most common cancers among survivors (n=1420, 53.1% female; n=1299, 48.6% Black individuals; mean age 65.2, SD 13.7 years) included breast (n=536, 20.1%), prostate (n=400, 15%), and lymphoid or hematopoietic (n=323, 12.1%) cancer. Among survivors in persistent poverty areas (n=509, 19%; all with a high Social Vulnerability Index), 34.6% (176/509) were single, 55.4% (282/509) had Medicare coverage (with only 73/509, 14.3% having private insurance), 36.5% (186/509) had obesity, 63.9% (325/509) had hypertension, and 31.2% (159/509) had diabetes. Of survivors in persistent poverty areas with PRAPARE data, 15.8% (19/120) lacked transportation, 4.2% (5/120) lived with housing insecurity, and 6.7% (8/120) felt unsafe where they lived.

Conclusions: Innovative EHR and NLP approaches identified several socioeconomic and safety-related challenges along with opportunities for health behavior interventions to leverage Medicare coverage and target multiple comorbidities when adapting interventions for survivors of cancer living in persistent poverty areas.

(JMIR Cancer 2026;12:e81054) doi:10.2196/81054

KEYWORDS

social determinants of health; survivors of cancer; persistent poverty; electronic health records; adapted interventions

Introduction

Persistent poverty areas, defined by the US Census Bureau, are counties or census tracts where at least 20% of the population has lived below the federal poverty level for 30 or more years. Despite medical advancements improving cancer survival [1], mortality rates remain disproportionately high in these areas [2]. Community-based research is imperative for adapting interventions to this population, ensuring effectiveness and reducing disparities [3].

Dietary and physical activity interventions help prevent cancer recurrence and mitigate comorbidities such as obesity and diabetes [4-6]. However, implementing these interventions in persistent poverty areas is challenging, with ongoing uncertainties regarding their acceptability, feasibility, and sustainability. These areas experience adverse social determinants of health (SDOH), including food insecurity, transportation barriers, and financial strains, all of which limit engagement in interventions [7]. However, these social risk data have not been systematically collected or integrated into clinical records, limiting their availability for tailoring intervention strategies [3].

XSL•FO
RenderX

Despite increasing use of electronic health records (EHRs) to improve health care quality, they remain underused for contextualizing cancer control efforts in persistent poverty areas. EHR data offer a unique opportunity to systematically characterize the demographic, clinical, and social risk profiles of local populations, providing actionable information for adapting interventions to community-specific needs. To date, only 1 study has used EHR data to characterize survivors of cancer in persistent poverty areas for future community interventions [8], but it did not capture key SDOH (eg, food insecurity) or integrate the findings into a community-engaged co-design process. With growing interest in integrating SDOH and social risk screening into health systems [9,10], richer data are becoming more readily available to characterize local populations of interest and design targeted interventions.

The Leveraging Adaptation and Multilevel Implementation Strategies to Address Unique Health Promotion Challenges Among Cancer Survivors in Persistent Poverty Areas (LEAP) study (U54CA280770) seeks to adapt evidence-based dietary and physical activity interventions in persistent poverty areas while building community capacity for implementation and sustainability. This report describes an innovative use of EHR data, including social risk data, to identify demographic, medical, and SDOH factors potentially relevant to diet and exercise intervention adaptation and delivery for survivors and cosurvivors of cancer living in persistent poverty areas. The objective of this descriptive study was to leverage EHR data—including an embedded social risk screener and natural language processing (NLP) of social work notes—to characterize demographic, clinical, and SDOH factors among survivors of cancer residing in persistent poverty areas and translate these findings into information for community-engaged adaptation of lifestyle interventions.

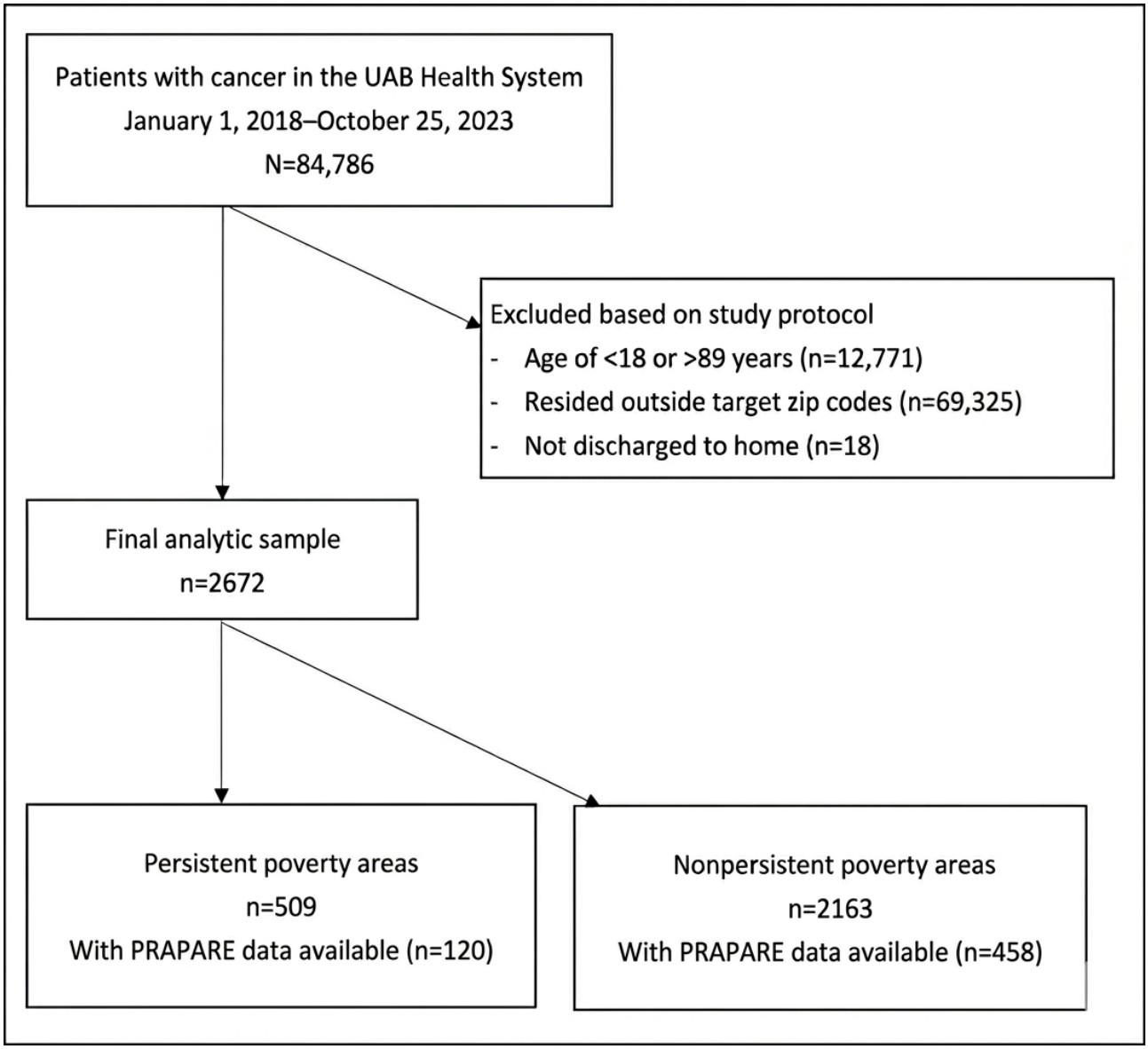
Methods

Population

We extracted EHR data from the University of Alabama at Birmingham Health System, a multispecialty group practice. The study population included patients with cancer who (1) received medical services between January 2018 and November 2023, (2) were aged 18 to 89 years, (3) were considered potential intervention participants (ie, returned home), and (4) resided in predefined communities (per the study protocol) within the Birmingham metropolitan area of Alabama that were planned locations for LEAP study adaptation and implementation (Figure 1).

XSL•FO
RenderX

Figure 1. Study flow diagram. PRAPARE: Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences; UAB: University of Alabama at Birmingham.



Because the health system's data extraction is operated by zip code, patients were initially identified using zip codes overlapping with target persistent poverty census tracts; as a result, the extracted population included residents of both persistent and nonpersistent poverty areas. Patient addresses were then geocoded to census tracts using institutional geocoding services. Persistent poverty status was determined by linking census tract codes to US Census Bureau persistent poverty designations. The Social Vulnerability Index (SVI) was linked at the census tract level using Centers for Disease Control and Prevention and Agency for Toxic Substances and Disease Registry SVI data. Consistent with the National Cancer Institute definition, survivors of cancer were defined as individuals with a documented cancer diagnosis from the time of diagnosis onward regardless of current treatment status. This definition was operationalized by identifying patients with cancer-related *International Classification of Diseases, Ninth Revision*; *International Classification of Diseases, Tenth Revision*; or Systematized Nomenclature of Medicine–Clinical Terms diagnostic codes in their EHR.

EHR data provided clinical information, including cancer type, body weight status, and comorbid conditions, as well as demographic characteristics (eg, age and race or ethnicity) and SDOH (eg, insurance and marital status). Census tract codes were used to identify persistent poverty areas among those living in low-income zip code geographic areas and were used to link the SVI, a composite neighborhood-level index assessing social vulnerabilities [11]. Additional SDOH data were available through the Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences (PRAPARE) [12], a patient-reported social risk screener integrated into the EHR system at our institution in the Social Work unit in 2020 [13]. PRAPARE includes 21 questions across 4 domains: personal characteristics, family and home, money and resources, and social and emotional health [12].

Social history notes collected via social work visits were organized in the EHR using Social Interventions Research and Evaluation Network (SIREN) [14] nomenclature. These data were analyzed using NLP to identify key SDOH themes aligned

XSL•FO
RenderX

with the Centers for Medicare and Medicaid Services [15]. We applied a rule-based NLP approach to identify events related to unmet needs from both structured questionnaire items and free-text patient responses recorded in the EHR. Text data were first preprocessed through tokenization and removal of stop words, punctuation, and noninformative characters, followed by normalization (eg, lowercasing and stemming) to standardize text across sources [16,17]. This preprocessing was applied consistently to both questionnaire prompts and patient responses.

We then conducted keyword-based text matching to identify SDOH-related events across 5 predefined domains: food insecurity, housing instability, transportation needs, interpersonal safety, and utility difficulties. The domain-specific keywords and phrases used to extract relevant events and identify unmet needs or negative responses are detailed in Table S1 in Multimedia Appendix 1. For each SDOH domain, patients were classified as experiencing unmet needs or insecurity if their most recent response indicated an adverse condition. Patients without such indicators were classified as having no identified unmet need (ie, secure) for that domain.

We summarized findings in tables for community engagement processes (eg, informing co-design and focus group guides) and intervention adaptation (eg, educational content topics, potential delivery channels, and anticipated challenges). Although hypothesis testing was not our primary focus, 2-tailed *t* tests, and chi-square or Fisher exact tests were used to estimate differences between survivors living in persistent poverty areas and those not living in these areas to examine trends in the data. Distributions of continuous variables (age and BMI) were approximately symmetric, supporting the use of *t* tests for group comparisons. Given the descriptive nature of this study and the absence of a priori hypotheses, *P* values are presented to identify potential trends rather than to establish statistical significance. No adjustment for multiple comparisons was applied, and the results should be interpreted as exploratory. All analyses were conducted using R (version 4.3.2; R Foundation for Statistical Computing).

Ethical Considerations

This study was approved by the Institutional Review Board at UAB (approval number IRB-300010921), with informed consent waived. This study was conducted in accordance with the study protocol. Only patients who met the predefined inclusion criteria were included. To protect participant privacy, we avoided reporting overly granular or potentially identifiable information in the results, including zip codes and census tracts. No compensation was provided to the participants.

Results

Among 2672 survivors, the mean age was 65.2 (SD 13.7) years, 1420 (53.1%) were female, and 1299 (48.6%) were Black individuals (Table 1). Common cancers included breast (n=536, 20.1%), prostate (n=400, 15%), and lymphoid or hematopoietic (n=323, 12.1%) cancer. A total of 19% (n=509) of the survivors lived in persistent poverty areas, with high obesity (186/509, 36.5%) and hypertension (325/509, 63.9%). These survivors were more likely to be Black individuals (297/509, 58.3% vs 1002/2163, 46.3% in nonpersistent poverty areas), insured by Medicaid (64/509, 12.6% vs 145/2163, 6.7% in nonpersistent poverty areas) or uninsured (35/509, 6.9% vs 95/2163, 4.4% in nonpersistent poverty areas), single (176/509, 34.6% vs 576/2163, 26.6% in nonpersistent poverty areas), and from neighborhoods with higher SVI (509/509, 100% vs 940/2163, 43.5% in nonpersistent poverty areas).

XSL•FO
RenderX

Table . Characteristics of survivors of cancer within a large multispecialty group practice (overall and by whether the survivor lived in a persistent poverty census tract).

Variable	Overall (N=2672)	Residing in persistent poverty areas[a] (n=509)	Residing in nonpersistent poverty areas (n=2163)	*P* value[b]
Demographics				
Age at assessment (years), mean (SD)	65.2 (13.7)	65.4 (12.8)	65.2 (13.9)	.72
Sex, n (%)				.23
Female	1420 (53.1)	259 (50.9)	1161 (53.7)	
Male	1252 (46.9)	250 (49.1)	1002 (46.3)	
Race or ethnicity, n (%)				<.001
Asian	290 (10.9)	101 (19.8)	189 (8.7)	
Black	1299 (48.6)	297 (58.3)	1002 (46.3)	
Hispanic	42 (1.6)	19 (3.7)	23 (1.1)	
White	959 (35.9)	76 (14.9)	883 (40.8)	
Other[c]	82 (3.1)	16 (3.1)	66 (3.1)	
Spoken language, n (%)				<.001
English	2467 (92.3)	445 (87.4)	2022 (93.5)	
Other or unknown	205 (7.7)	64 (12.6)	141 (6.5)	
Health insurance, n (%)				<.001
Private	645 (24.1)	73 (14.3)	572 (26.4)	
Medicaid	209 (7.8)	64 (12.6)	145 (6.7)	
Medicare	1460 (54.6)	282 (55.4)	1178 (54.5)	
Uninsured	130 (4.9)	35 (6.9)	95 (4.4)	
Other[c]	228 (8.5)	55 (10.8)	173 (8.0)	
Marital status, n (%)				<.001
Married or with a life partner	1013 (37.9)	136 (26.7)	877 (40.5)	
Single	752 (28.1)	176 (34.6)	576 (26.6)	
Other[c]	907 (33.9)	197 (38.7)	710 (32.8)	
Clinical measures				
BMI (kg/m^2)[d], mean (SD)	29.4 (9.51)	29.1 (9.00)	29.5 (9.63)	.39
BMI category (kg/m^2), n (%)				.77
<25	791 (29.6)	153 (30.1)	638 (29.5)	
25 to <30	705 (26.4)	127 (25.0)	578 (26.7)	
≥30	1019 (38.1)	186 (36.5)	833 (38.5)	
Comorbidity[e], n (%)				
Diabetes mellitus	858 (32.1)	159 (31.2)	699 (32.3)	.68
Hypertension	1714 (64.1)	325 (63.9)	1389 (64.2)	.92
Hyperlipidemia	1106 (41.4)	208 (40.9)	897 (41.5)	.84
Coronary artery disease	519 (19.4)	103 (20.2)	416 (19.2)	.65
Overall comorbidity	1877 (70.2)	349 (68.6)	1528 (70.6)	.39
Cancer site[e], n (%)				.01
Breast	536 (20.1)	108 (21.2)	428 (19.8)	

XSL•FO
RenderX

Variable	Overall (N=2672)	Residing in persistent poverty areas[a] (n=509)	Residing in nonpersistent poverty areas (n=2163)	*P* value[b]
Prostate	400 (15.0)	87 (17.1)	313 (14.5)	
Lymphoid or hematopoietic	323 (12.1)	37 (7.3)	286 (13.2)	
Lung	151 (5.7)	31 (6.1)	120 (5.5)	
Colorectal	139 (5.2)	30 (5.9)	109 (5.0)	
Skin	81 (3.0)	5 (1.0)	76 (3.5)	
Bones or soft tissue	61 (2.3)	13 (2.6)	48 (2.2)	
Gastrointestinal	218 (8.2)	38 (7.5)	180 (8.3)	
Genitourinary other than prostate	275 (10.3)	65 (12.8)	210 (9.7)	
Head and neck	142 (5.3)	32 (6.3)	110 (5.1)	
Brain or central nervous system	38 (1.4)	4 (0.8)	34 (1.6)	
Other	308 (11.5)	59 (11.6)	249 (11.5)	
Social Vulnerability Index[f], n (%)				<.001
Low (0 to <0.33)	429 (16.1)	0 (0.0)	429 (19.8)	
Moderate (0.33 to <0.66)	606 (22.7)	0 (0.0)	606 (28.0)	
High (≥0.66)	1449 (54.2)	509 (100.0)	940 (43.5)	

[a]Persistent poverty areas were defined as census tracts in which 20% or more of the population has lived below the poverty level for at least 30 years.

[b]Continuous and categorical variables were compared using a *t* test and chi-square test or Fisher exact test, respectively.

[c]"Other" includes "other," "multiple," and "decline to answer" for race and ethnicity; "other" and "missing" for health insurance; and "separated," "divorced," "widowed," and "unknown" for marital status.

[d]A total of 5.9% (157/2672) of patients had missing data.

[e]Diagnoses were based on the *International Classification of Diseases, Ninth Revision*; *International Classification of Diseases, Tenth Revision*; or Systematized Nomenclature of Medicine codes.

[f]The Social Vulnerability Index (SVI) summarizes the extent to which a community is socially vulnerable to disaster. It was developed by the Centers for Disease Control and Prevention. SVI values range from 0 (least vulnerable) to 1 (most vulnerable). Data were missing for 7.0% (188/2672) of the participants, all from nonpersistent poverty areas.

Richer SDOH data were available for 2 subsets of patients: those with PRAPARE data and those with SIREN data. Of the survivors with PRAPARE data available (578/2672, 21.6%), survivors from persistent poverty areas (120/578, 20.8%) reported lack of transportation (19/120, 15.8%), 4.2% (5/120) reported housing insecurity, and 6.7% (8/120) reported feeling unsafe where they lived (Table 2). While no single social risk differed significantly between the 2 groups, survivors in persistent poverty areas reported a higher mean number of social risks when compared with those in nonpersistent poverty areas (5.28, SD 1.84 vs 4.55, SD 1.92). Of those with SIREN data available through NLP (1597/2672, 59.8%), 13.6% (41/302) of survivors in persistent poverty areas experienced food insecurity (Table S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Prevalence of social determinants of health extracted from the electronic health record (EHR) using Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences (PRAPARE) questions: survivors of cancer overall and by residence in persistent poverty areas vs nonpersistent poverty areas[a]

PRAPARE tool responses	Overall (n=578)	Residing in persistent poverty areas[b] (n=120)	Residing in nonpersistent poverty areas (n=458)	*P* value[c]
Personal characteristics, n (%)				
Migrant work in the last 2 years				>.99
Yes	5 (0.9)	1 (0.8)	4 (0.9)	
No	520 (90.0)	108 (90.0)	412 (90.0)	
Missing or refused to answer	53 (9.2)	11 (9.2)	42 (9.2)	
Veteran status				.07
Yes	42 (7.3)	14 (11.7)	28 (6.1)	
No	477 (82.5)	96 (80.0)	381 (83.2)	
Missing or refused to answer	59 (10.2)	10 (8.3)	49 (10.7)	
Family and home, n (%)				
Housing status				.72
Lack of housing	12 (2.1)	3 (2.5)	9 (2.0)	
Have housing	550 (95.2)	115 (95.8)	435 (95.0)	
Missing or refused to answer	16 (2.8)	2 (1.7)	14 (3.1)	
Housing security				>.99
Worried	24 (4.2)	5 (4.2)	19 (4.1)	
Not worried	523 (90.5)	108 (90.0)	415 (90.6)	
Missing or refused to answer	31 (5.4)	7 (5.8)	24 (5.2)	
Money and resources, n (%)				
Highest educational grade completed				>.99
High school education or lower	262 (45.3)	56 (46.7)	206 (45.0)	
Higher than a high school education	129 (22.3)	27 (22.5)	102 (22.3)	
Missing or refused to answer	187 (32.4)	37 (30.8)	150 (32.8)	
Current work				.09
Unemployed	232 (40.1)	50 (41.7)	182 (39.7)	
Part time or other	26 (4.5)	6 (5.0)	20 (4.4)	
Unemployed but not seeking work[d]	212 (36.7)	49 (40.8)	163 (35.6)	
Full-time work	46 (8.0)	3 (2.5)	43 (9.4)	
Missing or refused to answer	62 (10.7)	12 (10.0)	50 (10.9)	
Income between 100% and 200% of the FPL[e]	26 (4.5)	6 (5.0)	20 (4.4)	
Income more than 200% of the FPL	20 (3.5)	2 (1.7)	18 (3.9)	
Reported resource needs				.93

XSL•FO
RenderX

PRAPARE tool responses	Overall (n=578)	Residing in persistent poverty areas[b] (n=120)	Residing in nonpersistent poverty areas (n=458)	*P* value[c]
Medicine or health care needs	82 (14.2)	17 (14.2)	65 (14.2)	
Resource needs[f]	30 (5.2)	6 (5.0)	24 (5.2)	
No needs	408 (70.6)	83 (69.2)	325 (71.0)	
Missing or refused to answer	58 (10.0)	14 (11.7)	44 (9.6)	
Transportation				.16
Lacking transportation	68 (11.8)	19 (15.8)	49 (10.7)	
Having transportation	476 (82.4)	94 (78.3)	382 (83.4)	
Missing or refused to answer	34 (5.9)	7 (5.8)	27 (5.9)	
Social and emotional health, n (%)				
Social integration and support				.32
≤5 times a week	275 (47.6)	61 (50.8)	214 (46.7)	
>5 times a week	220 (38.1)	40 (33.3)	180 (39.3)	
Missing or refused to answer	83 (14.4)	19 (15.8)	64 (14.0)	
Stress level rating				.62
High and medium-high stress[g]	198 (34.3)	39 (32.5)	159 (34.7)	
Less stress	291 (50.3)	64 (53.3)	227 (49.6)	
Missing or refused to answer	89 (15.4)	17 (14.2)	72 (15.7)	
Feeling safe where they lived				.10
No or unsure	21 (3.6)	8 (6.7)	13 (2.8)	
Yes	429 (74.2)	89 (74.2)	340 (74.2)	
Missing or refused to answer	128 (22.1)	23 (19.2)	105 (22.9)	
Afraid of ex-partner in the last year				>.99
Yes or unsure	5 (0.9)	1 (0.8)	4 (0.9)	
No	410 (70.9)	92 (76.7)	318 (69.4)	
Missing or refused to answer	163 (28.2)	27 (22.5)	136 (29.7)	
Risks per patient[h]				
0 to 4, n (%)	162 (28.0)	19 (15.8)	143 (31.2)	<.001
≥4, n (%)	416 (72.0)	101 (84.2)	315 (68.8)	

XSL•FO
RenderX

PRAPARE tool responses	Overall (n=578)	Residing in persistent poverty areas[b] (n=120)	Residing in nonpersistent poverty areas (n=458)	*P* value[c]
Mean (SD)	4.69 (1.92)	5.28 (1.84)	4.53 (1.92)	<.001

[a]Missing data and refusals to answer were included in the calculation, but the data are not shown in the table. Ethnicity, race, and spoken language under "Personal characteristics," as well as health insurance under "Money and resources" in the PRAPARE tool overlap with the EHR data presented in Table 1. Therefore, these variables were not reported in this table; however, their distribution in the subsample was comparable to that of the full sample.

[b]Persistent poverty areas were defined as census tracts in which more than 20% of the population has lived below the poverty level for at least 30 years.

[c]Continuous and categorical variables were compared using a *t* test and chi-square test or Fisher exact test, respectively.

[d]Patients reported Family and Medical Leave Act and those who were disabled or retired were categorized as "Unemployed but not seeking work."

[e]FPL: federal poverty level. The FPL was US $36,450 for a single individual in 2024.

[f]Resource needs included food, housing, transportation, childcare, utilities, phone, and others apart from medical or health care needs.

[g]High and medium-high stress included the response categories of "quite a bit," "very much," and "somewhat" stressed.

[h]Out of 22 negative events on the PRAPARE tool.

We collated the descriptive findings and created a summarization table to share with community stakeholders and academic investigators (Table 3). We linked our findings to a "lesson learned" goal that we then translated into a "callout" that was incorporated into our ongoing community engagement communications being used for LEAP study adaptation and implementation planning. For instance, we found that 64.1% (1714/2672) of the survivors in our sample had hypertension, leading to a callout that emphasized that survivors of cancer often experience other chronic conditions.

Table . Translation of electronic health record data into "lessons learned" shared with community stakeholders and academic investigators. These "lessons learned" reflect descriptive prevalence within the persistent poverty population and are intended for community engagement communication; they do not represent statistically significant differences between populations.

Social determinants of health for the patients with cancer in the health system living in persistent poverty[a] census tracts	Goal for the "lesson learned" designed for lay public communication	"Lesson learned" related to survivors of cancer in the community shared to guide priorities for intervention adaptations
A total of 62% of patients with cancer living in persistent poverty census tracts are overweight or obese.	Raise awareness of the need for diet and exercise while avoiding a focus on weight alone.	Nearly 2 in 3 survivors of cancer in your community are at greater health risk due to excess body weight.
A total of 31% of patients with cancer have diabetes, and 64% have hypertension (41% have hyperlipidemia, and 20% have coronary artery disease).	Emphasize that cancer often occurs in the context of other medical conditions.	Cancer is not alone (2 out of 3 survivors of cancer may have another medical condition that healthy diet and exercise could help with).
A total of 55% of patients with cancer living in persistent poverty census tracts have Medicare, suggesting possible opportunities to leverage Medicare benefits related to nutrition and exercise counseling and programs.	Raise awareness of the need to help connect individuals with resources already available to them.	One in 2 survivors of cancer may have insurance coverage for diet and exercise education.
There are 25 different cancer types; prostate cancer was the second most prevalent cancer type in our study (17%), only slightly lower than breast cancer (21%).	Encourage people to think about a broad spectrum of cancer types and the importance of focusing on programs addressing the diverse needs of both male and female survivors of cancer.	Cancer has multiple faces in your community—men, women, and over 24 different cancer types.
The Social Vulnerability Index[b] was high, indicating that the study areas were characterized by census-based indicators of low socioeconomic status.	Raise awareness of the major contributions to the Social Vulnerability Index for survivors of cancer specifically.	Survivors of cancer may face barriers due to older age, disability, discrimination, and limited financial resources.
Of the 2672 survivors of cancer with electronic health record data, 509 lived in persistent poverty areas.	The goal was to raise awareness of the prevalence of living in a persistent poverty area among health system patients.	One in 5 patients with cancer at the UAB[c] Health System faces significant financial struggles.

[a]Persistent poverty areas were defined as census tracts in which more than 20% of the population has lived below the poverty level for at least 30 years.

[b]The Social Vulnerability Index (SVI) summarizes the extent to which a community is socially vulnerable to disaster. It was developed by the Centers for Disease Control and Prevention. SVI values range from 0 (least vulnerable) to 1 (most vulnerable).

[c]UAB: University of Alabama at Birmingham.

XSL•FO
RenderX

Discussion

Principal Findings

This study sought to characterize demographic, clinical, and SDOH factors among survivors of cancer in persistent poverty areas using EHR data to inform community-engaged adaptation of lifestyle modification interventions. We found that survivors of cancer in persistent poverty areas within our health system catchment were predominantly Black individuals and older adults with Medicare coverage and experienced high rates of obesity and chronic conditions, including hypertension and diabetes. Among those with social risk screening data, notable proportions reported transportation barriers, health care needs, food instability, fewer social interactions, and stress, consistent with the findings of a previous study [18].

Our data also highlight the burden of social risks in persistent poverty areas, consistent with previous research [2,8,19]. Health disparities among survivors of cancer are driven by a complex interplay of SDOH [20], which also serve as primary barriers to adequate health care access and physician communication [21-24]. Understanding these needs—as we did in this study—in planning tailored interventions will be imperative for sustainability and improved outcomes.

Survivors of cancer in persistent poverty areas often face multiple challenges of cancer survivorship, poverty, and clusters of problems that collectively affect their health. Adapting interventions to address the unique barriers that these populations face is important. In this study, we aggregated data on SDOH to guide the discussions with the planning, implementation, and evaluation working groups (community stakeholders and survivors and cosurvivors of cancer) convened as part of the ongoing LEAP study community engagement processes. This approach aligns with evidence-based cancer control strategies in persistent poverty areas [8]. Our findings support a comprehensive approach to promoting healthy eating and physical activity that addresses clusters of social risks rather than focusing on a single factor (eg, telehealth options to minimize transportation barriers, meal programs or home gardening to address food insecurity, and peer support groups to reduce social isolation).

The growing integration of SDOH screening into EHR systems represents a paradigm shift in how health care systems can support cancer control efforts in underserved communities [25]. Our approach aligns with emerging recommendations from national organizations emphasizing the importance of systematic SDOH data collection and documentation to inform care delivery and population health management [22]. The scoping review by Li et al [26] demonstrated that EHR-embedded SDOH data can be effectively linked to health outcomes and used to design targeted interventions, although standardization of data collection methods remains an ongoing challenge. Future efforts should focus on expanding SDOH screening beyond social work encounters to capture a more representative picture of social risks among survivors of cancer.

Community-based participatory research approaches have demonstrated effectiveness in optimizing health behaviors among survivors of cancer, particularly when interventions are tailored to local community resources and cultural contexts [3]. A recent systematic review found that most community-based interventions for survivors of cancer focused on physical activity and nutrition, with 88% showing improvement in at least one outcome measure; however, the review also highlighted a concerning lack of diversity in participant populations [27]. Our study provides foundational data to inform such community-based participatory research efforts by characterizing the specific demographic, clinical, and social risk profiles of survivors of cancer in our target communities—information essential for adopting evidence-based interventions to address local needs. The Active Living After Cancer program, for example, demonstrated that community-delivered physical activity interventions can successfully improve outcomes among medically underserved survivors of cancer and their caregivers when implemented through trusted community partnerships [28].

Notably, more than half (282/509, 55.4%) of the survivors of cancer in persistent poverty areas in our sample had Medicare coverage, suggesting opportunities to leverage existing Medicare benefits for lifestyle interventions. While Medicare currently covers medical nutrition therapy only for diabetes and chronic kidney disease, recent legislative efforts such as the Medical Nutrition Therapy Act for dietitian-delivered nutrition counseling will be important for this population [29]. In the meantime, health care systems serving persistent poverty areas can explore partnerships with community-based programs and leverage existing Medicare preventive services to address modifiable health behaviors among survivors of cancer.

The strengths of this study include using real-world health system data and NLP to capture both medical and nonmedical factors influencing health outcomes, enabling a more informed approach to addressing disparities prior to intervention development. Furthermore, the aggregated SDOH data were "translated" into a format for nonscience community-based partners to inform the ongoing community-engaged "co-design" of the planned LEAP intervention.

Limitations included the bias in the study sample of survivors of cancer, which was limited to the project target areas defined by low socioeconomic status indicators, and the PRAPARE- and NLP-derived data were also limited to individuals with the most acute social needs (eg, referral to social work). For this reason, we did not conduct formal hypothesis testing and did not focus on those results for interpretation in this report, and the findings may not be generalizable to other medical practice catchments. Our sample reflects survivors of cancer actively engaged in care at our health system, which may underrepresent patients with rapidly fatal cancers and those who have disengaged from care. This "survivor bias" should be considered when interpreting cancer type distributions; however, for the purpose of informing intervention adaptation, characterizing the population currently available for enrollment is appropriate.

Missing data were present for several variables, particularly PRAPARE (2094/2672, 78.4% missing) and NLP-derived SDOH (1076/2672, 40.3% missing) data, which reflects the clinical reality that these data are collected only for patients

XSL•FO
RenderX

referred to social work services. This represents a form of data missing not at random, as patients with more acute social needs are more likely to have social work encounters. This limitation should be considered when interpreting prevalence estimates, which likely represent lower bounds for the broader population.

Despite these limitations, this is the first study to use an EHR-embedded social risk screener to provide insightful information for formative work among survivors of cancer in persistent poverty areas, offering valuable insights into their unique challenges and underscoring the importance of adapting interventions to address their specific social and health needs.

Conclusions

This study demonstrates that leveraging SDOH data embedded in health system EHRs provides a feasible approach for characterizing populations of survivors of cancer in persistent poverty areas to inform community-based intervention adaptation. The social risk profiles identified highlight the importance of tailoring lifestyle interventions to address transportation, food access, and safety concerns alongside chronic disease management.

Acknowledgments

The authors would express their sincere appreciation to the research staff members who supported this study and their community partners who tirelessly work to improve the health and well-being of their communities.

Funding

This article was partially supported by the National Institutes of Health (NIH) National Cancer Institute (NCI; P30CA013148, U54CA280770, U54CA118938). This research project is supported by U54CA280770 (Project 2 LEAP). Additional funding was provided by the National Institute on Minority Health and Health Disparities (Howell - 1K01 MD0172706) and American Heart Association Grant # 931540 / Carrie R. Howell / 2022. The content is solely the responsibility of the authors and does not necessarily represent the official view of the NCI, NIHMD or others supporting this work.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to institutional policies regarding HIPAA (Health Insurance Portability and Accountability Act) protected medical files.

Authors' Contributions

Conceptualization: CRH, LQR, LZ
Data curation: ST, CRH, SB
Formal analysis: ST, CRH
Methodology: ST, CRH, LQR, LZ, SB
Supervision: CRH, LQR, LZ, SB, MP
Writing—original draft: ST, CRH, LQR, LZ
Writing—review and editing: All authors

Conflicts of Interest

None declared.

Multimedia Appendix 1
Natural language processing keywords used and social driver prevalence.
[DOCX File, 29 KB - cancer_v12i1e81054_app1.docx]

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024;74(1):12-49. [doi: 10.3322/caac.21820] [Medline: 38230766]
2. Moss JL, Pinto CN, Srinivasan S, Cronin KA, Croyle RT. Persistent poverty and cancer mortality rates: an analysis of county-level poverty designations. Cancer Epidemiol Biomarkers Prev 2020 Oct;29(10):1949-1954. [doi: 10.1158/1055-9965.EPI-20-0007] [Medline: 32998949]
3. Kale S, Hirani S, Vardhan S, et al. Addressing cancer disparities through community engagement: lessons and best practices. Cureus 2023 Aug 14;15(8):e43445. [doi: 10.7759/cureus.43445] [Medline: 37711952]
4. Gan T, Cheng HL, Tse MM. Feasibility, acceptability, and effects of behavior change interventions for improving multiple dietary behaviors among cancer survivors: a systematic review. Support Care Cancer 2022 Mar;30(3):2877-2889. [doi: 10.1007/s00520-021-06582-2] [Medline: 34581862]

XSL•FO
RenderX

5. Wang S, Yang T, Qiang W, Zhao Z, Shen A, Zhang F. Benefits of weight loss programs for breast cancer survivors: a systematic reviews and meta-analysis of randomized controlled trials. Support Care Cancer 2022 May;30(5):3745-3760. [doi: 10.1007/s00520-021-06739-z] [Medline: 35022884]
6. Mbous YP, Patel J, Kelly KM. A systematic review and meta-analysis of physical activity interventions among colorectal cancer survivors. Transl Behav Med 2020 Oct 12;10(5):1134-1143. [doi: 10.1093/tbm/ibz176] [Medline: 33044539]
7. Benson C, Bishaw A, Glassman B. Persistent poverty in counties and census tracts. : United States Census Bureau; 2023 URL: https://www.census.gov/content/dam/Census/library/publications/2023/acs/acs-51%20persistent%20poverty.pdf [accessed 2026-03-18]
8. Raber M, Rechis R, LaRue DM, et al. Enhancing the utilization of healthy living interventions among cancer survivors in historically underserved populations and communities. Cancer Causes Control 2023 Aug;34(8):635-645. [doi: 10.1007/s10552-023-01701-2] [Medline: 37160832]
9. Gruß I, Bunce A, Davis J, Dambrun K, Cottrell E, Gold R. Initiating and implementing social determinants of health data collection in community health centers. Popul Health Manag 2021 Feb;24(1):52-58. [doi: 10.1089/pop.2019.0205] [Medline: 32119804]
10. Vilendrer S, Thomas SC, Belnap T, et al. Screening for social determinants of health during primary care and emergency department encounters. JAMA Netw Open 2023 Dec 1;6(12):e2348646. [doi: 10.1001/jamanetworkopen.2023.48646] [Medline: 38113046]
11. Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. A social vulnerability index for disaster management. J Homeland Secur Emerg Manage 2011;8(1). [doi: 10.2202/1547-7355.1792]
12. Weir RC, Proser M, Jester M, Li V, Hood-Ronick CM, Gurewich D. Collecting social determinants of health data in the clinical setting: findings from national PRAPARE implementation. J Health Care Poor Underserved 2020;31(2):1018-1035. [doi: 10.1353/hpu.2020.0075] [Medline: 33410822]
13. Howell CR, Bradley H, Zhang L, et al. Real-world integration of the protocol for responding to and assessing patients' assets, risks, and experiences tool to assess social determinants of health in the electronic medical record at an academic medical center. Digit Health 2023;9:20552076231176652. [doi: 10.1177/20552076231176652] [Medline: 37252259]
14. Gottlieb L, Fichtenberg C, Adler N. Introducing the Social Interventions Research and Evaluation Network: SIREN; 2017.
15. Billioux A, Verlander K, Anthony S, Alley D. Standardized screening for health-related social needs in clinical settings: the accountable health communities screening tool. National Academy of Medicine. 2017. URL: https://nam.edu/perspectives/standardized-screening-for-health-related-social-needs-in-clinical-settings-the-accountable-health-communities-screening-tool/ [accessed 2026-03-18]
16. Patra BG, Sharma MM, Vekaria V, et al. Extracting social determinants of health from electronic health records using natural language processing: a systematic review. J Am Med Inform Assoc 2021 Nov 25;28(12):2716-2727. [doi: 10.1093/jamia/ocab170] [Medline: 34613399]
17. Lybarger K, Dobbins NJ, Long R, et al. Leveraging natural language processing to augment structured social determinants of health data in the electronic health record. J Am Med Inform Assoc 2023 Jul 19;30(8):1389-1397. [doi: 10.1093/jamia/ocad073] [Medline: 37130345]
18. Hastert TA, McDougall JA, Robinson JR, et al. Age at diagnosis and social risks among Black cancer survivors: results from the Detroit Research on Cancer Survivors cohort. Cancer 2024 Jun 1;130(11):2060-2073. [doi: 10.1002/cncr.35212] [Medline: 38280205]
19. Hallgren E, Yeary KH, DelNero P, et al. Barriers, facilitators, and priority needs related to cancer prevention, control, and research in rural, persistent poverty areas. Cancer Causes Control 2023 Dec;34(12):1145-1155. [doi: 10.1007/s10552-023-01756-1] [Medline: 37526781]
20. Lara-Morales A, Soto-Ruiz N, Agudelo-Suárez AA, García-Vivar C. Social determinants of health in post-treatment cancer survivors: scoping review. Eur J Oncol Nurs 2024 Jun;70:102614. [doi: 10.1016/j.ejon.2024.102614] [Medline: 38795448]
21. Palmer NR, Geiger AM, Felder TM, Lu L, Case LD, Weaver KE. Racial/ethnic disparities in health care receipt among male cancer survivors. Am J Public Health 2013 Jul;103(7):1306-1313. [doi: 10.2105/AJPH.2012.301096] [Medline: 23678936]
22. Tucker-Seeley R, Abu-Khalaf M, Bona K, et al. Social determinants of health and cancer care: an ASCO policy statement. JCO Oncol Pract 2024 May;20(5):621-630. [doi: 10.1200/OP.23.00810] [Medline: 38386945]
23. Huang H, Wei T, Huang Y, et al. Association between social determinants of health and survival among the US cancer survivors population. BMC Med 2024 Aug 26;22(1):343. [doi: 10.1186/s12916-024-03563-0] [Medline: 39183305]
24. DiMartino LD, Birken SA, Mayer DK. The relationship between cancer survivors' socioeconomic status and reports of follow-up care discussions with providers. J Cancer Educ 2017 Dec;32(4):749-755. [doi: 10.1007/s13187-016-1024-3] [Medline: 27006193]
25. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: a systematic review. J Am Med Inform Assoc 2020 Nov 1;27(11):1764-1773. [doi: 10.1093/jamia/ocaa143] [Medline: 33202021]

XSL•FO
RenderX

26. Li C, Mowery DL, Ma X, et al. Realizing the potential of social determinants data in EHR systems: a scoping review of approaches for screening, linkage, extraction, analysis, and interventions. J Clin Transl Sci 2024 Oct 10;8(1):e147. [doi: 10.1017/cts.2024.571] [Medline: 39478779]
27. Seven M, Marie Moraitis A. Community-based interventions designed to optimize health behaviors among cancer survivors: an integrative systematic review. Support Care Cancer 2022 Oct;30(10):8405-8415. [doi: 10.1007/s00520-022-07097-0] [Medline: 35524870]
28. Mama SK, Mitchell SJ, Tracy PV, et al. Expanding active living after cancer to underserved cancer survivors and their caregivers. J Natl Cancer Inst 2024 Aug 1;116(8):1333-1342. [doi: 10.1093/jnci/djae097] [Medline: 38688563]
29. Cancer in Medicare: an American Cancer Society Cancer Action Network chartbook. American Cancer Society Cancer Action Network. 2024. URL: https://www.fightcancer.org/sites/default/files/national_documents/acscan-medicare-chartbook.pdf [accessed 2026-03-18]

Abbreviations:

EHR: electronic health record
LEAP: Leveraging Adaptation and Multilevel Implementation Strategies to Address Unique Health Promotion Challenges Among Cancer Survivors in Persistent Poverty Areas
NLP: natural language processing
PRAPARE: Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences
SDOH: social determinants of health
SIREN: Social Interventions Research and Evaluation Network
SVI: Social Vulnerability Index

Edited by M Balcarras; submitted 21.Jul.2025; peer-reviewed by NR Geyer, P Wong, S Goyal; revised version received 27.Feb.2026; accepted 02.Mar.2026; published 13.Apr.2026.

Please cite as:
Tanaka S, Rogers LQ, Zubkoff L, Bateman L, Demark-Wahnefried W, Rocque G, Bae S, Pisu M, Booth J, Howell CR
Leveraging Innovative Electronic Health Record Data to Characterize Social Determinants of Health Among Survivors of Cancer in Persistent Poverty Areas: Cross-Sectional Study
JMIR Cancer 2026;12:e81054
URL: https://cancer.jmir.org/2026/1/e81054
doi:10.2196/81054

© Shiori Tanaka, Laura Q Rogers, Lisa Zubkoff, Lori Bateman, Wendy Demark-Wahnefried, Gabrielle Rocque, Sejong Bae, Maria Pisu, James Booth, Carrie R Howell. Originally published in JMIR Cancer (https://cancer.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Short Paper

The Role of Resilience in Coping With Future Uncertainty Among People With Brain Tumors: Cross-Sectional Study

Li-Ting Huang Longcoy[1], PhD, DrPH; Shu-Yuan Liang[2], PhD; Ardith Z Doorenbos[1], PhD

[1]College of Nursing, University of Illinois Chicago, Chicago, IL, United States

[2]National Taipei University of Nursing and Health Sciences, Taipei City, Taiwan

Corresponding Author:
Shu-Yuan Liang, PhD
National Taipei University of Nursing and Health Sciences
No. 365 Ming-Te Road
Peitou District
Taipei City, 112
Taiwan
Phone: 886 2 28227101 ext 3105
Email: shuyuan@ntunhs.edu.tw

Abstract

Background: Adults with brain tumors learn to navigate unpredictable physical and psychological symptoms along with the possibilities of tumor recurrence. As a result, they tend to become resilient to confronting profound uncertainty and actively employ coping strategies. Yet, the impact of resilience on coping strategies among people with brain tumors has not been fully explored.

Objective: This study aimed to examine the effects of resilience on the association between future uncertainty and two distinct types of coping strategies (problem-focused coping and emotion-focused coping) among people with brain tumors in Taiwan.

Methods: A parent study recruited 95 adults with brain tumors undergoing at least 1 month of chemotherapy or radiotherapy at a veterans general hospital in northern Taiwan. We assessed resilience, future uncertainty, and coping strategies via a secondary analysis of data from the parent study collected using the Chinese versions of the Resilience Scale, the European Organization for Research and Treatment of Cancer's Quality of Life Questionnaire for brain cancer, and the revised Ways of Coping Checklist. Simple mediation models were conducted to examine the role of resilience between future uncertainty and the two types of coping strategies.

Results: Most participants demonstrated low resilience and responded to stress with both problem- and emotion-focused coping strategies. Simple mediation analyses showed a statistically significant association between an increase in resilience and adoption of each type of coping strategy. In addition, resilience was a statistically significant mediator in the association between future uncertainty and both problem- and emotion-focused coping strategies.

Conclusions: Brain tumor disease trajectories require people to effectively adopt both problem- and emotion-focused coping strategies to confront uncertainty. Health care providers play a crucial role in evaluating and fostering their patients' resilience to promote adaptability through effective coping strategies.

(JMIR Cancer 2026;12:e71674) doi:10.2196/71674

KEYWORDS

resilience; problem-focused coping; emotion-focused coping; adaptation; uncertainty in illness theory

Introduction

Given the nature of brain tumors, uncertainty may persist throughout the cancer trajectory. In Taiwan, brain tumors account for 1.2% of cancer cases; however, the mortality rate has been rising for 2 decades [1]. The lack of clarity about prognoses leads people with brain tumors to experience ongoing uncertainty along with unpredictable symptom patterns and increases in functional dependency [2]. As a result, these sources of uncertainty may negatively affect their ability to cope with cancer [3].

According to Mishel's uncertainty in illness theory [4], an individual responds to uncertainty by assessing whether it presents a danger or an opportunity. When there is a potential

XSL•FO
RenderX

positive outcome from the uncertainty, the individual may use strategies such as selective ignoring to maintain uncertainty as a source of hope. A danger appraisal may prompt the individual to take action to mitigate the uncertainty [4,5]. For people with brain tumors, the ongoing perception of uncertainty highlights the need for effective coping strategies to support successful adaptation; this is a process that may be strengthened through resilience. The known positive association between resilience and cancer coping [6,7] suggests that cultivating resilience may enhance cognitive appraisal processes and thus increase the adoption of effective coping strategies.

Resilience is defined as the process by which an individual can, in the face of stress, flexibly utilize available external resources (eg, social support from others), internal resources (eg, individual strengths and skills), and existential resources (eg, practices of meaning-making and expressing gratitude) [8]. A qualitative study of people with brain tumors identified the need for resilience, including living in the moment and finding joy, as coping mechanisms for navigating uncertainty [9]. Other studies have shown that resilience mediates the association between symptom distress and quality of life and that resilience has effects on reducing psychological distress [10]. These findings highlight that resilience can serve as a protective factor to empower people to demonstrate flexibility in overcoming difficulties through various resources and coping strategies.

There has been little research on the impact of resilience on how individuals with brain tumors cope with uncertainty. Thus, this study, guided by the uncertainty in illness theory, investigates the effects of resilience on the association between uncertainty and individual coping strategies, including problem-focused coping and emotion-focused coping, among people with brain tumors.

Methods

Sample and Setting

This secondary analysis used data from 95 people with primary brain tumors, which was sufficient to reach the desired power of .80 to examine the effect of future uncertainty on coping strategies through resilience. Participants were recruited during an outpatient visit to a veterans general hospital in Taiwan. Eligibility criteria were (1) patients aged 20 years or older and (2) patients diagnosed with a benign or malignant primary brain tumor and receiving at least 1 month of treatment. People with a diagnosed mental illness or unable to communicate were not eligible to participate.

Ethical Considerations

The parent study was approved by the Taipei Veterans General Hospital (2014-09-007AC) institutional review board (for more details about the parent study, please see [11,12]). Before consenting, participants were informed that they could withdraw from the study at any time without penalty, and that this would not affect their treatment. In addition, no identifiable health information was collected or included in the statistical or data management programs. No compensation was provided to participants for completing the survey.

Instruments

Coping Strategies

We used the Chinese version of the Ways of Coping Checklist–Revised to assess participants' coping strategies. The Ways of Coping Checklist–Revised uses 42 items to assess 5 factors: problem-focused coping, seeking social support, self-blame, distancing, and wishful thinking [13]. We categorized the first 2 factors as problem-focused coping and the other 3 factors as emotion-focused coping [13]. All items use a 5-point Likert scale (0 to 4); higher scores indicate more frequent adoption of a strategy. The McDonald ω for the internal consistency of the total scores of problem-focused and emotion-focused coping were 0.86 and 0.83, respectively.

Future Uncertainty

Participants' future uncertainty was measured using the subscale of the European Organization for Research and Treatment of Cancer's Quality of Life Questionnaire [14]. The 4 items that measure future uncertainty are as follows: (1) Did you feel uncertain about the future? (2) Did you feel you had setbacks in your condition? (3) Were you concerned about disruption of family life? and (4) Did your outlook on the future worsen? All items were measured on a scale of 1 to 4; higher scores represent worse symptoms. The McDonald ω for the total score was 0.80.

Resilience

Participants' resilience was assessed using the Chinese version of the Resilience Scale, with 25 items in 5 domains of resilience, including meaningful life, perseverance, self-reliance, equanimity, and existential aloneness. All items use a 7-point Likert scale ranging from 1 to 7, for a total score ranging from 25 to 175 (≥146=moderately high resilience; 131-145=moderate resilience; ≤130=low resilience) [15]. The McDonald ω for the total score was 0.97.

Data Analysis

To examine how resilience affects the effect of future uncertainty on coping, we conducted 2 simple mediation models, with resilience as the mediator and future uncertainty as the predictor of the two different outcome variables: problem-focused coping and emotion-focused coping. Each model consisted of one direct effect of future uncertainty (X) as a predictor of coping strategies (Y) and one indirect effect of future uncertainty (X) on coping strategies (Y) through resilience (M). Inferences about direct and indirect effects were estimated by bootstrapping, which generated 95% percentile bootstrap CIs [16].

Results

Participants adopted both problem-focused coping strategies (mean 58.06, SD 11.10) and emotion-focused coping strategies (mean 43.96, SD 12.06). Half (48/95, 51%) of the participants reported low resilience, and a quarter each reported moderate (24/95, 25%) and moderately high (23/95, 24%) resilience.

Two simple mediation models examined the role of resilience in the different types of coping strategies. The model for resilience as a mediator between future uncertainty and

XSL•FO
RenderX

problem-focused coping resulted in a standardized coefficient of total effect of −0.202 (Figure 1). When the total effect of future uncertainty was partitioned into direct and indirect effects, the indirect effect of future uncertainty on problem-focused coping through resilience (Figure 1, path A × path B) was statistically different from zero (β=−0.268, SE 0.068, 95% CI −0.406 to −0.142).

The model with emotion-focused coping as the outcome resulted in a total effect of future uncertainty that was statistically significant (Figure 2). The direct effect of future uncertainty on emotion-focused coping was 0.370. The indirect effect of uncertainty through resilience was significantly different from zero (β=−0.133, SE 0.053, 95% CI −0.249 to −0.041).

Figure 1. Simple mediation model estimation results. C' is the direct effect of future uncertainty on problem-focused coping; C is the total effect of future uncertainty on problem-focused coping. *P<.05; **P<.01.

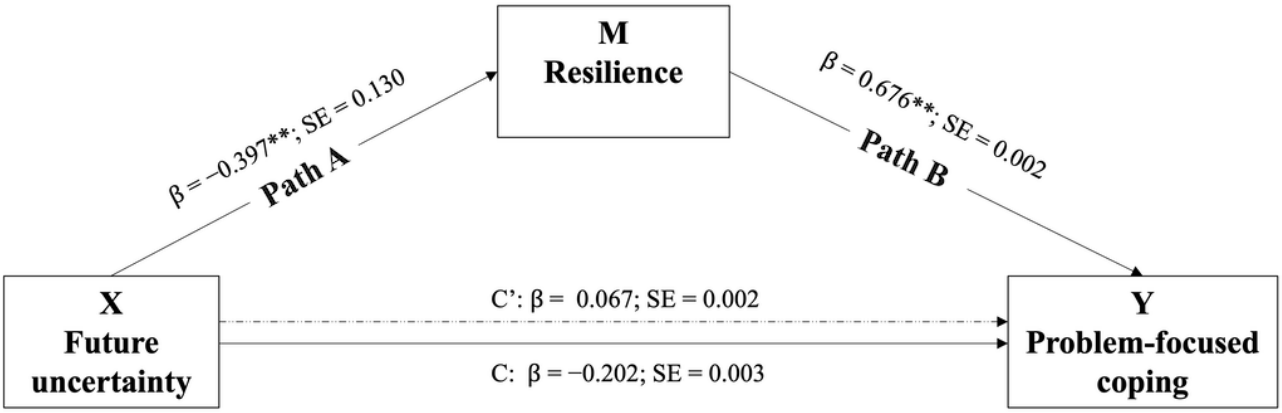
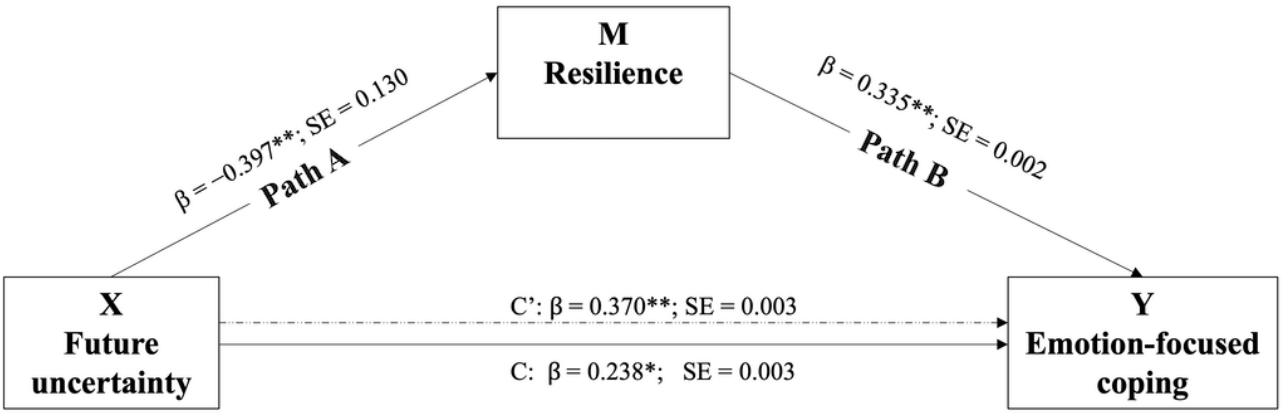


Figure 2. Simple mediation model estimation results. C' is the direct effect of future uncertainty on emotion-focused coping; C is the total effect of future uncertainty on emotion-focused coping. *P<.05; **P<.01.



Discussion

Principal Findings

This study investigates the influence of resilience on adopting 2 distinct coping strategies amidst the uncertainty associated with brain tumors. We found that resilience positively facilitated the adoption of both problem-focused and emotion-focused coping strategies and a mediating role for resilience between uncertainty and individual coping strategies. In particular, the mediating effects of resilience appeared to be more pronounced for problem-focused coping, as the direct effect of uncertainty on problem-focused coping became nonsignificant when resilience was included. Indeed, focus group discussions among people with brain tumors revealed that constant uncertainty led to fear, despair, and adjustment difficulties. Over time, they learned to adopt problem-focused coping strategies such as planning end-of-life care to reduce caregiver burden [9]. In other words, resilience is a process of realization through awareness of accurate appraisals, available resources, and the flexible use of coping strategies [8]. When the cause of uncertainty can be managed, awareness of available informational and instrumental social support through resilience may increase the likelihood of adopting problem-focused coping strategies with greater confidence [7,9,17].

Unclear symptom patterns and prognoses for brain tumors may prompt people to adopt emotion-focused coping strategies. People with brain tumors expressed the need to maintain hope as part of their coping strategies and to live in the present moment to mitigate the impact of psychological distress [9]. In fact, cultivating hope is integral to developing resilience [18], and it has been identified as a predictor of resilience among individuals with cancer [6]. Interestingly, our findings showed that the reduction in resilience resulting from increased uncertainty weakened the overall effect of uncertainty on emotion-focused coping. One possible explanation is that in coming to terms with their condition, people with brain tumors may ultimately cultivate hope, which in turn strengthens resilience and enables them to flexibly adopt either

XSL•FO
RenderX

problem-focused or emotion-focused coping strategies in response to uncertainty.

Limitations

This study captured participants' resilience levels at only one point in time. Our goal was to examine the impact of resilience while recruiting a sufficient sample for analysis within a feasible recruitment period. The study's sample size may limit further investigation. Although our sample shows slightly lower resilience, it remains generally comparable to other samples with brain tumors [19,20].

Clinical Implications

Health care providers can help patients with brain tumors develop resilience by recognizing and applying their external, internal, and existential resources to address uncertainty. With such assistance, the process of building resilience can be accelerated within 12 weeks [21]. Additionally, health care providers can provide meaning-making or gratitude exercises to increase resilience and facilitate successful adaptation [22].

Conclusion

This study highlights the importance of developing resilience to help people with brain tumors accurately appraise and flexibly apply effective coping strategies. The findings provide potential targets for a resilience-building intervention to reduce the impact of uncertainty.

Funding

When writing this manuscript, LTHL was supported by the National Institute of Nursing Research of the National Institutes of Health under award K99NR020358. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The data collected during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

LTHL wrote the manuscript and analyzed the data. LTHL, AZD, and SYL are responsible for the study design. All authors reviewed and edited the manuscript.

Conflicts of Interest

None declared.

References

1. Cancer registry annual report, 2022 Taiwan. Health Promotion Administration, Ministry of Health and Welfare, Taiwan. URL: https://www.hpa.gov.tw/Pages/List.aspx?nodeid=269 [accessed 2025-12-17]
2. Cahill J, LoBiondo-Wood G, Bergstrom N, Armstrong T. Brain tumor symptoms as antecedents to uncertainty: an integrative review. J Nurs Scholarsh 2012 Jun;44(2):145-155. [doi: 10.1111/j.1547-5069.2012.01445.x] [Medline: 22564342]
3. Hong S, Lee J, Lee J, Chang JH, Park CG, Kim TH. The mediating effect of uncertainty in illness on cancer coping in patients with primary malignant brain tumors. Cancer Nurs 2023;46(6):467-476. [doi: 10.1097/NCC.0000000000001177] [Medline: 36480344]
4. Mishel MH. Uncertainty in illness. Image J Nurs Sch 1988;20(4):225-232. [doi: 10.1111/j.1547-5069.1988.tb00082.x] [Medline: 3203947]
5. Zhang Y. Uncertainty in illness: theory review, application, and extension. Oncol Nurs Forum 2017 Nov 01;44(6):645-649. [doi: 10.1188/17.ONF.645-649] [Medline: 29052657]
6. Dong M, Nie A, Liu C, Zheng J. Psychological resilience of patients with bladder cancer after radical cystectomy and urinary diversion: a cross-sectional study. Psychooncology 2022 Jan;31(1):21-29. [doi: 10.1002/pon.5770] [Medline: 34344061]
7. Chen S, Wu S, Wang T, Rosenberg J, Lu Y, Liang S. Factors influencing the coping strategies of liver cancer patients undergoing transarterial chemoembolization. Int J Nurs Pract 2022 Aug;28(4):e13033. [doi: 10.1111/ijn.13033] [Medline: 34913227]
8. Rosenberg AR. Seeking professional resilience. Pediatrics 2018 Mar;141(3):e2017238 [FREE Full text] [doi: 10.1542/peds.2017-2388] [Medline: 29437139]
9. Loughan AR, Reid M, Willis KD, Barrett S, Lo K. The emotional journey of neuro-oncology: primary brain tumor patients share their experience during this life-threatening disease. Neurooncol Pract 2023 Feb;10(1):71-78 [FREE Full text] [doi: 10.1093/nop/npac067] [Medline: 36659970]

XSL•FO
RenderX

10. Ye ZJ, Qiu HZ, Li PF, Liang MZ, Zhu YF, Zeng Z, et al. Predicting changes in quality of life and emotional distress in Chinese patients with lung, gastric, and colon-rectal cancer diagnoses: the role of psychological resilience. Psychooncology 2017 Jun;26(6):829-835. [doi: 10.1002/pon.4237] [Medline: 27479936]
11. Liang SY, Liu HC, Lu YY, Wu SF, Chien CH, Tsay SL. The influence of resilience on the coping strategies in patients with primary brain tumors. Asian Nurs Res (Korean Soc Nurs Sci) 2020 Feb;14(1):50-55 [FREE Full text] [doi: 10.1016/j.anr.2020.01.005] [Medline: 31996321]
12. Pan CJ, Liu HC, Liang SY, Liu CY, Wu WW, Cheng SF. Resilience and coping strategies influencing the quality of life in patients with brain tumor. Clin Nurs Res 2019 Jan;28(1):107-124. [doi: 10.1177/1054773817714562] [Medline: 28658969]
13. Sheu S. Uncertainty and anxiety in patients with initial attack of myocardial infarction: the effect of coping methods. Hu Li Yan Jiu 2001 Apr;9(2):159-171. [Medline: 11548461]
14. Taphoorn MJB, Claassens L, Aaronson NK, Coens C, Mauer M, Osoba D, EORTC Quality of Life Group,Brain Cancer, NCICRadiotherapy Groups. An international validation study of the EORTC brain cancer module (EORTC QLQ-BN20) for assessing health-related quality of life and symptoms in brain cancer patients. Eur J Cancer 2010 Apr;46(6):1033-1040. [doi: 10.1016/j.ejca.2010.01.012] [Medline: 20181476]
15. Wagnild GM. The Resilience Scale User's Guide: For the U.S. English Version of The Resilience Scale TM and the 14-Item Resilience Scale TM (RS-14 TM). Montana, USA: The Resilience Center; 2011.
16. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. New York: The Guilford Press; 2013.
17. Sihvola S, Kuosmanen L, Kvist T. Resilience and related factors in colorectal cancer patients: a systematic review. Eur J Oncol Nurs 2022 Feb;56:102079 [FREE Full text] [doi: 10.1016/j.ejon.2021.102079] [Medline: 34844135]
18. Fredrickson BL. Chapter 1: positive emotions broaden and build. In: Advances in Experimental Social Psychology. Amsterdam, Netherlands: Academic Press; 2013:1-53.
19. Yao Q, Deng S, Liu L, Luo Y. Mediating effects of psychological resilience on the relationship between social support and body image dissatisfaction among patients with primary brain tumors in China. Psychol Health Med 2025 Jun;30(5):1043-1055 [FREE Full text] [doi: 10.1080/13548506.2024.2447007] [Medline: 39758004]
20. Jing Y, Liu C, Wang L, Liu X, Wang L. Association between psychological resilience and symptom burden in postoperative patients with brain gliomas and its influencing factors. Am J Cancer Res 2025;15(2):652-660. [doi: 10.62347/QTZW9567] [Medline: 40084376]
21. Rosenberg AR, Baker KS, Syrjala KL, Back AL, Wolfe J. Promoting resilience among parents and caregivers of children with cancer. J Palliat Med 2013 Jun;16(6):645-652 [FREE Full text] [doi: 10.1089/jpm.2012.0494] [Medline: 23646887]
22. Casellas-Grau A, Font A, Vives J. Positive psychology interventions in breast cancer. A systematic review. Psychooncology 2014 Jan;23(1):9-19. [doi: 10.1002/pon.3353] [Medline: 23897834]

Edited by J Lau; submitted 23.Jan.2025; peer-reviewed by SL Coven, A Zucca; comments to author 20.Jun.2025; revised version received 21.Nov.2025; accepted 23.Nov.2025; published 06.Jan.2026.

Please cite as:
Longcoy LTH, Liang SY, Doorenbos AZ
The Role of Resilience in Coping With Future Uncertainty Among People With Brain Tumors: Cross-Sectional Study
JMIR Cancer 2026;12:e71674
URL: https://cancer.jmir.org/2026/1/e71674
doi:10.2196/71674
PMID:

©Li-Ting Huang Longcoy, Shu-Yuan Liang, Ardith Z Doorenbos. Originally published in JMIR Cancer (https://cancer.jmir.org), 06.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital and Paper-Based Hospital Workflows and 60-Day Mortality in Acute Leukemia: Retrospective Natural Experiment

Christian Omar Ramos Peñafiel[1*], MSc, PhD; Álvaro Cabrera García[2*], MD; Carolina Balderas Delgado[2*], MD; Rocío Luna Tentle[2*], MD; Andrea Milán Salvatierra[3*], MD; Nora Godínez Cubillo[4*], MD; Gustavo Acosta Altamirano[4*], MD; Rosa Sánchez Conejo[4*], MD; Adán Germán Gallardo Rodríguez[5*], MSc, PhD

[1]Department of Hematology, Hospital General de México, Mexico City, Mexico

[2]Department of Hematology, Hospital Regional de Alta Especialidad de Ixtapaluca, Ixtapaluca, State of Mexico, Mexico

[3]Department of Hematology, Hospital Juárez de México, Mexico City, Mexico

[4]Department of Management and Direction, Hospital Regional de Alta Especialidad de Ixtapaluca, Ixtapaluca, State of Mexico, Mexico

[5]School of Sport Sciences, Universidad Anáhuac México, Huixquilucan, State of Mexico, Mexico

[*]all authors contributed equally

Corresponding Author:
Adán Germán Gallardo Rodríguez, MSc, PhD
School of Sport Sciences
Universidad Anáhuac México
Av. Universidad Anáhuac #46
Locality Lomas Anahuac
Huixquilucan, State of Mexico, 52786
Mexico
Phone: 52 5584005072
Email: nutriologo.agallardo8@gmail.com

Abstract

Background: Electronic medical record (EMR) systems have been associated with better clinical workflows and fewer documentation errors. However, evidence regarding their effect on time-sensitive leukemia care in public hospitals in Latin America remains limited.

Objective: This study aimed to compare 60-day mortality and urgent supportive-care processes between 2 tertiary public hospitals in Mexico with different documentation models (integrated EMR vs traditional physical records [TPR]), under the hypothesis that digital workflows may facilitate more timely treatment.

Methods: We conducted a retrospective natural experiment including 274 patients with newly diagnosed acute leukemia treated between February 2023 and April 2025. Clinical characteristics, treatment intensity, complications during induction, antibiotic administration times, and survival outcomes were abstracted from finalized institutional records that are routinely reviewed at discharge by the institutional medical record committee as part of standard quality-assurance procedures. The primary outcome was 60-day mortality. Secondary outcomes included time to first antibiotic dose after recognition of febrile neutropenia, treatment-related complications, and the number of operational steps required for urgent care processes. A subgroup analysis of 70 patients with complete timing documentation was performed for antibiotic administration. Comparisons between hospitals were performed using univariate tests, Kaplan-Meier 60-day survival curves, and multivariable logistic regression.

Results: Of the 274 included patients, 104 (38%) were treated at the hospital using an integrated EMR, and 170 (62%) at the hospital using TPR. Sixty-day mortality was lower in the EMR hospital (6/104, 5.8%) than in the TPR hospital (61/170, 35.9%; $P<.001$). In the subgroup with complete timing data, the mean time from febrile episode recognition to first antibiotic administration was shorter in the hospital using EMR than in the hospital using TPR (54, SD 18.4 minutes; $P<.001$). Although the EMR hospital used a higher proportion of high-intensity regimens, patients in that hospital had better 60-day outcomes. In multivariable analysis, hospital type remained independently associated with 60-day mortality (odds ratio 0.11, 95% CI 0.05-0.26; $P<.001$), whereas kidney injury and hepatotoxicity were associated with worse outcomes.

Conclusions: In this natural experiment, the hospital using an integrated EMR had a lower 60-day mortality and shorter time to antibiotic administration than the hospital using TPR. These findings are hypothesis-generating and suggest that digital

XSL•FO
RenderX

workflows may contribute to more timely urgent supportive care, but they should be interpreted with caution, given the retrospective 2-center design and the potential for residual confounding.

(JMIR Cancer 2026;12:e71306) doi:10.2196/71306

KEYWORDS

electronic medical records; traditional physical records; acute leukemia; 60-day mortality; Latin America

Introduction

Electronic medical records (EMRs) are digital patient records that integrate clinical history, laboratory results, diagnoses, treatments, and medication orders. Compared with traditional physical records (TPR), EMRs may improve traceability, facilitate communication among physicians, nurses, pharmacists, and ancillary services, and reduce delays in the execution of urgent clinical decisions [1,2].

Over the last 2 decades, digitalization has transformed hospital care in many high-income settings. In the United States, widespread adoption accelerated after the Health Information Technology for Economic and Clinical Health (HITECH) Act, and similar transitions have occurred in several European countries [3]. In oncology, digital tools have been associated with better coordination of care, improved remote monitoring, fewer unplanned admissions, and more timely responses to complications [4].

Despite these advances, Latin America continues to face major barriers in cancer care, including fragmented health systems, limited access to innovative therapies, administrative delays, and marked social inequities [5,6]. These limitations are especially relevant in hematologic malignancies, where delays in diagnosis, treatment initiation, and supportive care can have an immediate impact on survival [7].

In acute leukemias, early mortality remains a major challenge in the region. Patients with acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) are highly susceptible to infectious and hemorrhagic complications during the first weeks of treatment [8-12]. In Mexico, delays from symptom onset to treatment initiation have been associated with poor outcomes, and infectious complications are a leading cause of early death [13]. In acute promyelocytic leukemia, hemorrhage remains an additional major contributor to early mortality [14].

A substantial proportion of these delays occur not only at the diagnostic stage but also during hospital care, particularly when urgent interventions, such as antibiotic administration or blood component delivery, depend on multiple manual steps [15]. In many public referral hospitals in Mexico, these processes still rely on paper charts, handwritten requests, and fragmented communication pathways.

Therefore, we treated the contrast between these 2 hospitals—Hospital Regional de Alta Especialidad de Ixtapaluca and Hospital Juárez de México—as a natural experiment. Our aim was to compare 60-day mortality and time-sensitive supportive-care processes between a fully digital public hospital and a hospital using TPR, while also exploring clinical predictors of mortality. We hypothesized that EMR-supported workflows would be associated with faster delivery of urgent supportive care and lower 60-day mortality in patients with newly diagnosed acute leukemia.

Methods

Setting

We conducted a retrospective natural experiment, a type of observational study, of patients with newly diagnosed acute leukemia treated between February 2023 and April 2025 at 2 tertiary public hospitals in Mexico. The comparison included Hospital Juárez de México, which uses a TPR model with paper-based requests and manual delivery processes, and Hospital Regional de Alta Especialidad de Ixtapaluca, which operates with a fully digital infrastructure, including EMR-based prescribing, real-time medication management, and electronic request processing. Acute leukemia diagnosis was established using the standard institutional diagnostic workup, including morphological evaluation of peripheral blood and/or bone marrow together with immunophenotyping by flow cytometry. The French-American-British classification, when documented in the record, was used only as a descriptive morphological framework.

Participants

Eligible patients were consecutive adolescents and adults with a confirmed diagnosis of acute leukemia who received inpatient induction or initial inpatient antileukemic treatment during the study period. Patients were included regardless of leukemia subtype if the diagnosis and inpatient treatment course were documented in the institutional record. Patients without vital-status information through day 60 or with missing primary outcome data were excluded from the primary analysis.

Data Collection: Data Sources and Variables

Data were retrospectively abstracted from finalized institutional records at each center using a predefined study dataset. Data abstraction was performed by trained investigators involved in the study using standardized extraction forms developed for the project.

At both hospitals, the medical record of each hospitalized patient is routinely reviewed at discharge by the institutional medical record committee as part of standard administrative and quality-assurance procedures; thus, the study used records that had already undergone routine institutional review. In the TPR hospital, study variables were obtained from paper charts, handwritten medical and nursing notes, medication sheets, and discharge documentation. In the EMR hospital, the same variables were obtained from the digital platform used for clinical documentation, medication tracking, and time-stamped orders. Collected variables included age, sex, leukemia subtype,

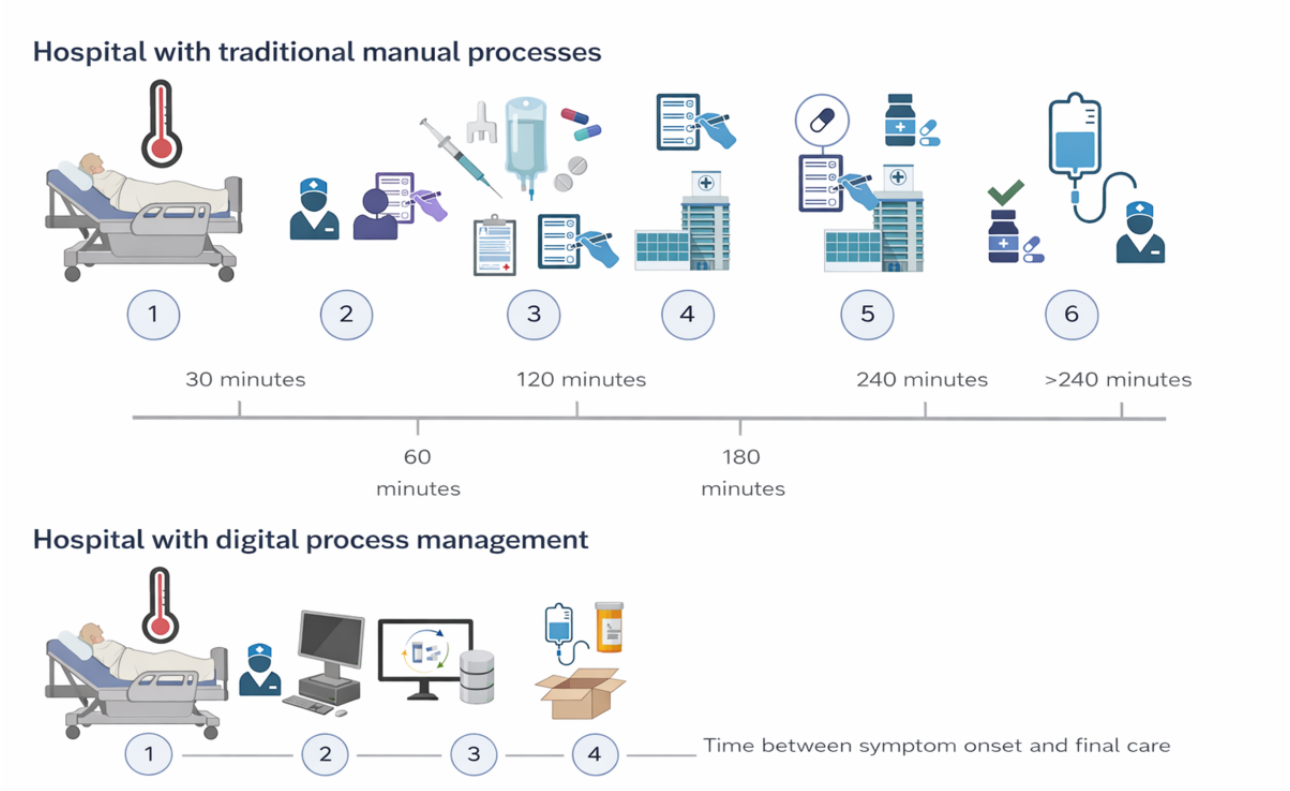
XSL•FO
RenderX

baseline white blood cell count, prior infection, overweight status (BMI ≥25 kg/m^2), treatment intensity, treatment-related complications, organ toxicities, and vital status at 60 days. No formal duplicate abstraction or interrater reliability exercise was performed for the research dataset, which should be considered when interpreting the results.

Outcomes and Treatment Classification

The primary outcome was 60-day mortality, defined as death from any cause within 60 days of hospital admission. Secondary outcomes were time from recognition of febrile neutropenia to first antibiotic administration, treatment-related complications during induction, and the number of procedural steps required for urgent antibiotic or blood component delivery. Because our objective was not to assess case-detection capacity but rather to evaluate whether different documentation environments could more efficiently translate clinical decisions into timely pharmacological interventions, workflow analyses focused on execution steps and time stamps after clinical recognition of febrile neutropenia. To characterize workflow differences, we recorded the number of operational steps required for emergent care processes in each hospital. For antibiotic timing analyses, we evaluated a subgroup of 70 patients with complete documentation of febrile episode recognition, antibiotic prescription, and administration time. In the TPR hospital, timing data were obtained from handwritten nursing notes and medication records; in the EMR hospital, they were extracted from electronic time stamps. Figure 1 summarizes the process flow in both institutions. Treatment regimens followed local institutional practice. For ALL, the most frequently used regimens were Hyper-CVAD and CALGB 10403 [16,17]. For AML, the standard 7+3 regimen was the preferred induction approach. For analysis, chemotherapy was categorized as low-, intermediate-, or high-intensity. High-intensity regimens included fludarabine, cytarabine, and granulocyte colony-stimulating factor (FLAG), whereas lower-intensity regimens included azacitidine and low-dose cytarabine [18].

Figure 1. Differences between a hospital that uses traditional physical records (TPRs) and one that use electronic medical records (EMRs) in the duration of antibiotic use and blood-component administration. The sequence of events in the TPR hospital is as follows—step 1: symptom identification; step 2: patient or relatives personally notify health care staff; step 3: symptom registration by nursing or medical personnel; step 4: manual medication request, gathering authorization from supervisors or responsible staff, and verifying medication availability; step 5: approval by senior staff or infectious disease specialists and forwarding the request to the pharmacy; step 6: the pharmacy receives the request, validates prescriptions, dispenses medications, and waits for personnel transport; and step 7: antibiotics administration to the patient. EMR hospital—step 1: symptom identification via automated alerts to nursing staff; step 2: cause identification and digital antibiotic requests with real-time pharmacy validation; and step 3: medication transport via capsule system or designated transport staff, and antibiotics administration to the patient.



Statistical Analysis

Because this was a retrospective study that included all consecutive eligible patients treated during the study period, a prospective sample size calculation was not performed. However, considering the observed mortality rates between hospitals, the available sample size provided adequate statistical power to detect clinically meaningful differences between groups. Assuming an early mortality of approximately 30% in the TPR hospital and 10% in the EMR hospital, the sample of 274 patients would provide more than 80% statistical power to detect an absolute difference of at least 20 percentage points between groups at a 2-sided alpha level of .05.

XSL•FO
RenderX

Descriptive statistics were used to summarize baseline characteristics and outcomes. Categorical variables were compared using chi-square or Fisher exact tests, as appropriate, and continuous variables were compared using 2-tailed Student *t* tests or Mann-Whitney *U* tests according to distribution. Missing data were not imputed. Analyses were performed using complete-case information for each variable, and patients without 60-day follow-up were excluded from the primary outcome analysis. Sixty-day survival was analyzed using Kaplan-Meier curves and log-rank tests. To explore factors associated with 60-day mortality in this natural experiment, we fitted a multivariable logistic regression model. Covariates were selected a priori on the basis of clinical relevance, prior literature on early mortality in acute leukemia, baseline imbalances between hospitals, and the need to preserve parsimony relative to the number of deaths observed. The final model included hospital type (EMR vs TPR) as the main exposure of interest, together with age, leukemia type, white blood cell count at diagnosis, prior infection, overweight status, kidney injury, and hepatotoxicity. Odds ratios (ORs) with 95% CIs were reported. A 2-sided *P* value <.05 was considered statistically significant. All analyses were performed using SPSS (version 25.0; IBM Corp).

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and the institutional procedures of the participating centers. At both hospitals, medical records are routinely reviewed by the institutional medical record committee at discharge as part of standard care and quality control processes. The study used deidentified information derived from these finalized records. The protocol was approved by the Biosafety, Ethics, and Research Committee of the Hospital Regional de Alta Especialidad de Ixtapaluca (protocol HRAEI-CBSI-09-24-2025). Because this was a retrospective analysis of anonymized clinical records, informed consent was waived. All data were handled in a deidentified manner, and no personally identifiable information was accessed or included in the research dataset, in accordance with applicable data protection regulations. No financial or nonfinancial compensation was provided to participants, as this study was based exclusively on retrospective review of existing clinical records.

Results

Patient Characteristics

A total of 274 patients were included in the study: 170 (62%) in the TPR group and 104 (38%) in the EMR group. Sex distribution was balanced overall, and no significant difference was observed between hospitals (male sex: 53/104, 51% for EMRs vs 84/170, 49.4% for TPRs; *P*=.45).

The overall mean age was 35 years (range 16-87 years). Patients with AML were older than those with ALL (41 vs 32 years; *P*=.04). ALL was the most frequent diagnosis (183/274, 66.8%), followed by AML (n=46, 16.8%), acute promyelocytic leukemia (n=24, 8.8%), and other myeloid subtypes (n=21, 7.7%).

Regarding comorbidities, 36 (13.1%) patients had diabetes, 18 (6.6%) had hypertension, and 53 (19.3%) were classified as overweight or obese. Secondary leukemia was documented in 9 cases (3.3%), all corresponding to myeloid disease.

Baseline clinical and demographic characteristics by hospital are summarized in Table 1.

XSL•FO
RenderX

Table 1. Baseline demographic and clinical characteristics of patients[a].

	Electronic medical record group (n=104)	Traditional physical records group (n=170)	*P* value
Sex, n (%)			.45
Male	53 (51)	84 (49.4)	
Female	51 (49)	86 (50.6)	
Age (years), median (IQR; range)	29.00 (21.00-41.75; 16.00-87.00)	35.00 (17.00-69.00)	.13
Age (y), n (%)			.07
<35	63 (60.6)	86 (50.6)	
>35	41 (39.4)	84 (49.4)	
White blood cell count at diagnosis (<30×10^3 vs >30×10^3), n (%)			<.001
<30×10^3	79 (76.0)	94 (55.3)	
>30×10^3	25 (24.0)	76 (44.7)	
White blood cell count at diagnosis (<100×10^3 vs >100×10^3) , n (%)			.005
<100×10^3	91 (87.5)	126 (74.1)	
>100×10^3	13 (12.5)	44 (25.9)	
French-American-British classification, n (%)			__[b]
Acute lymphoblastic leukemia	81 (77.9)	102 (60.0)	
Acute myeloid leukemia	15 (14.4)	31 (18.2)	
Acute promyelocytic leukemia	3 (2.9)	21 (12.4)	
M6 and M7	2 (1.9)	13 (7.6)	
Biphenotypic	3 (2.9)	3 (1.8)	
Secondary leukemia, n (%)			.47
Present	4 (3.8)	5 (2.9)	
Absent	100 (96.2)	165 (97.1)	
***BCR::ABL1* mutation, n (%)**			.40
Present	8 (7.7)	16 (9.4)	
Absent	96 (92.3)	154 (90.6)	
Central nervous system infiltration , n (%)			.06
Present	3 (2.9)	14 (8.2)	
Absent	101 (97.1)	156 (91.8)	
Diabetes, n (%)			.25
Present	16 (15.4)	20 (11.8)	
Absent	88 (84.6)	150 (88.2)	
Hypertension, n (%)			.56
Present	7 (6.7)	11 (6.5)	
Absent	97 (93.3)	159 (93.5)	
Overweight, n (%)			<.001
Present	45 (43.3)	8 (4.7)	
Absent	59 (56.7)	162 (95.3)	
Previous infection, n (%)			.002
Present	41 (39.4)	38 (22.4)	
Absent	63 (60.6)	132 (77.6)	

XSL•FO
RenderX

	Electronic medical record group (n=104)	Traditional physical records group (n=170)	*P* value
Cytogenetic, n (%)			—
Without karyotype	2 (1.9)	36 (21.2)	
Low risk	48 (46.2)	67 (39.4)	
Medium risk	06 (5.8)	12 (7.1)	
High risk	48 (46.2)	55 (32.4)	
Elapsed days between symptoms to diagnosis, n (median; range)	30.00 (17.75-60.00; 0.00-180.00)	30.00 (21.00-60.00; 0.00-360.00)	.17
Elapsed days between diagnostic to chemotherapy treatment, n (median; range)	6.00 (3.00-9.00: 1.00-29.00)	3.00 (1.00-6.00; 1.00-15.00)	<.001
Chemotherapy regimen, n (%)			—
Low intensity	1 (1.0)	9 (5.3)	
Medium intensity	26 (25.0)	128 (75.3)	
High intensity	77 (74)	33 (19.4)	

[a]The Wilcoxon signed-rank test was used to compare medians of quantitative nonparametric variables for independent groups, and the chi-square test was used for categorical variables. A *P* value of <.05 was considered statistically significant.

[b]Not applicable.

Baseline Characteristics Before Chemotherapy

Before treatment initiation, among 274 patients, 71 (25.9%) presented with fever and required antibiotic therapy. At admission, the average white blood cell count was 58×10^3/µL (range 0.3-462×10^3/µL), without a significant difference between lymphoid and myeloid leukemias (*P*=.35).

Cytogenetic data were available for 159 (58%) patients, with no significant difference between hospitals (*P*=.31). Among 183 patients with ALL, *BCR::ABL1* was detected in 24 (13.1%), and central nervous system infiltration was documented in 17 (9.3%).

Detailed baseline characteristics stratified by hospital are shown in Table 1.

Treatment Patterns by Hospital

Induction treatment was categorized among 274 patients as low intensity in 10 (3.6%), intermediate intensity in 154 (56.2%), and high intensity in 110 (40.1%). The median time from diagnosis to chemotherapy initiation was 5 (range 1-29; IQR 4) days. High-intensity regimens were more frequently used in the EMR hospital (77/104, 74%) than in the TPR hospital (33/170, 19.4%), whereas intermediate-intensity regimens were more common in the TPR hospital (TPR hospital: 128/170, 75.3% vs EMR hospital, 26/104, 25%). Low-intensity treatment was uncommon in both centers and was mainly reserved for patients who were frail or those with major comorbidities.

Among patients with ALL, most received intermediate-intensity regimens, whereas patients with AML were predominantly treated with standard induction approaches, such as 7+3. Treatment characteristics during induction are summarized in Table 2.

XSL•FO
RenderX

Table 2. General characteristics and complications during chemotherapy[a].

	Electronic medical record group (n=104)	Traditional physical record group (n=170)	*P* value
Fever during chemotherapy, n (%)			.001
Present	98 (94.2)	136 (80.0)	
Absent	6 (5.8)	34 (20.0)	
Neutropenic colitis, n (%)			<.001
Confirmed	53 (51.0)	37 (21.8)	
Absent	51 (49.0)	133 (78.2)	
Lung pneumonia, n (%)			.07
Confirmed	31 (29.8)	35 (20.6)	
Absent	73 (70.2)	135 (79.4)	
Systemic candidemia, n (%)			.23
Confirmed	16 (15.4)	20 (11.8)	
Absent	88 (84.6)	150 (88.2)	
Major hemorrhage, n (%)			.28
Presence	27 (26.0)	50 (29.4)	
Absent	77 (74.0)	120 (70.6)	
Tumoral lysis syndrome, n (%)			.56
Present	13 (12.5)	19 (11.2)	
Absent	91 (87.5)	151 (88.2)	
Acute kidney failure, n (%)			.02
Present	15 (14.4)	44 (25.9)	
Absent	89 (85.6)	126 (74.1)	
Acute liver failure, n (%)			.03
Present	19 (18.3)	47 (27.6)	
Absent	85 (81.7)	123 (72.4)	
Pancreatitis, n (%)			.03
Present	0 (0.0)	7 (4.1)	
Absent	104 (100.0)	163 (95.9)	
Toxic megacolon, n (%)			.11
Present	1 (0.9)	8 (4.7)	
Absent	103 (99.1)	162 (95.3)	
Neutropenia (days elapsed), median (range)	23 (18-28)	22 (10.75-30.25)	.42
Deep neutropenia (days elapsed), median (range)	9 (0-27)	6 (0-37)	.14
Positive cultures, n (%)			<.001
Present	67 (64.4)	64 (37.6)	
Absent	37 (35.6)	106 (62.4)	
Gram-negative bacteria, n (%)			<.001
Present	49 (47.1)	40 (23.5)	
Absent	55 (52.9)	130 (76.5)	
Gram-positive bacteria, n (%)			.002
Present	33 (31.7)	27 (15.9)	
Absent	71 (68.3)	143 (84.1)	

XSL•FO
RenderX

	Electronic medical record group (n=104)	Traditional physical record group (n=170)	*P* value
Fungal isolation, n (%)			<.001
Present	14 (13.5)	06 (3.5)	
Absent	90 (86.5)	164 (96.5)	
Clostridium difficile **, n (%)**			<.001
Present	16 (15.4)	5 (2.9)	
Absent	88 (84.6)	165 (97.1)	
More than 1 organism in culture, n (%)			<.001
Present	31 (29.8)	11 (6.5)	
Absent	73 (70.2)	159 (93.5)	
60-day survival , n (%)			<.001
Live	98 (94.2)	109 (64.1)	
Death	6 (5.8)	61 (35.9)	

[a]The Wilcoxon signed-rank test was used to compare medians of quantitative nonparametric variables for independent groups, and the chi-square test was used for categorical variables. A *P* value <.05 was considered statistically significant.

Treatment-Related Complications by Hospital

All patients underwent induction treatment. During this phase, among 274 patients, tumor lysis syndrome occurred in 32 (11.7%), kidney injury in 59 (21.5%), hepatotoxicity in 65 (23.7%), and pancreatitis in 7 (2.6%) patients.

Febrile neutropenia was the most frequent adverse event during induction therapy. It occurred in 98 (94.2%) of 104 patients in the EMR hospital and in 136 (80%) of 170 patients in the TPR hospital (*P*=.001).

Other infectious complications included neutropenic colitis in 90 (32.8%), positive cultures in 131 (47.8%), gram-negative isolation in 89 (32.5%), and gram-positive isolation in 60 (21.9%) patients. Microbiological isolation was reported more often in the EMR hospital.

Major hemorrhage was documented in 77 (28.1%) patients, with no significant difference between hospitals (EMR: 27/104, 26%; TPR: 50/170, 29.4%; *P*=.28).

Sixty-Day Mortality

Sixty-day mortality for the overall cohort was 24.5% (67/274). Mortality was significantly lower in the EMR hospital (6/104, 5.8%) than in the TPR hospital (61/170, 35.9%; *P*<.001).

Although infectious complications during chemotherapy were more frequently documented in the EMR hospital, severe organ toxicities, such as kidney injury and liver injury, were more common in the TPR hospital.

Mortality was higher among patients with AML (11/46, 23.9%) than among those with ALL (35/183, 19.1%). The leading causes of death were septic shock (41/274, 15%), hemorrhage (20/274, 7.3%), and fulminant liver failure (3/274, 1.1%).

Most deaths from septic shock occurred in patients with ALL, whereas hemorrhagic deaths were concentrated in myeloid leukemias, particularly acute promyelocytic leukemia.

In the multivariable logistic regression analysis, hospital type remained independently associated with 60-day mortality. Care at the EMR hospital was associated with significantly lower odds of death compared with the TPR hospital (OR 0.11, 95% CI 0.05-0.26; *P*<.001). In contrast, kidney injury (OR 1.68, 95% CI 1.29-2.20) and hepatotoxicity (OR 1.49, 95% CI 1.14-1.78) were independently associated with higher odds of 60-day mortality.

Workflow Measures and Time to Antibiotic Administration

The median duration of neutropenia was 22 (15-30) days overall, and severe neutropenia lasted a median of 7 (3-12) days. Despite a longer duration of neutropenia in the EMR group, only 6 (9%) of the 67 deaths occurred in that hospital.

The TPR hospital required more procedural steps for urgent care processes than the EMR hospital. For antibiotic administration, the TPR hospital required a median of 5 (range 3-7) steps versus 2 steps in the EMR hospital. A similar pattern was observed for blood component requests.

In the subgroup of 70 patients with complete timing documentation, the mean time from recognition of febrile neutropenia to first antibiotic administration was significantly longer in the TPR (272 minutes) hospital than in the EMR hospital (54 minutes; *P*<.001).

Secondary Survival Analyses and Multivariable Model

Kaplan-Meier analysis of 60-day survival showed poorer survival in patients with AML than in those with ALL. As a secondary analysis, survival also differed according to chemotherapy intensity, which was relevant because treatment intensity varied substantially between hospitals. Survival was also lower in patients treated at the TPR hospital than in those treated at the EMR hospital (Figures 2 and 3).

In multivariable logistic regression, hepatotoxicity, kidney injury, overweight status, age older than 35 years, and AML

XSL•FO
RenderX

were associated with higher odds of death. By contrast, high-intensity treatment and care at the EMR hospital were associated with lower odds of 60-day mortality. Figure 4 summarizes the ORs for the main variables included in the model.

Overall, these results support the hypothesis that a hospital environment using digitalized workflows may enable faster execution of urgent supportive-care processes and may be associated with lower 60-day mortality; however, the observational design does not allow a causal interpretation.

Figure 2. Kaplan-Meier curve comparing 60-day survival according to chemotherapy intensity (secondary analysis).

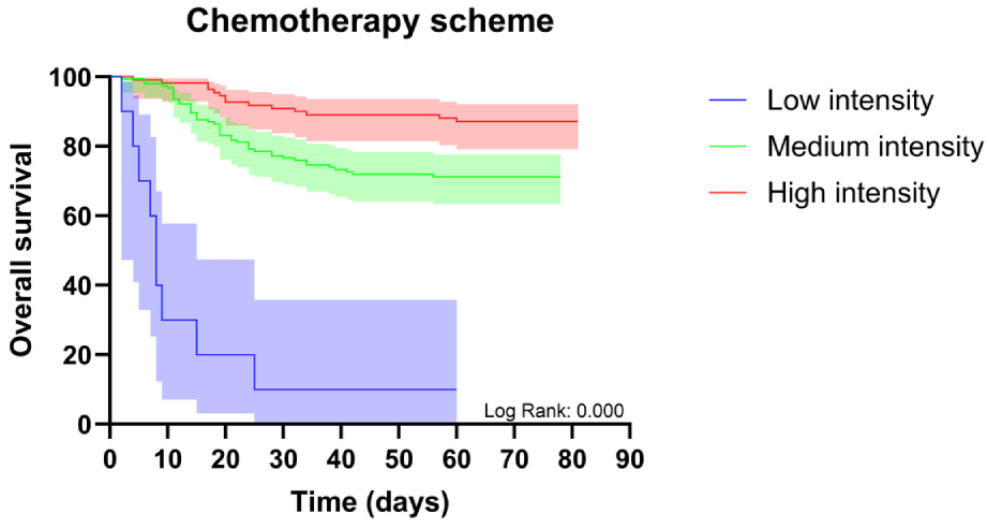
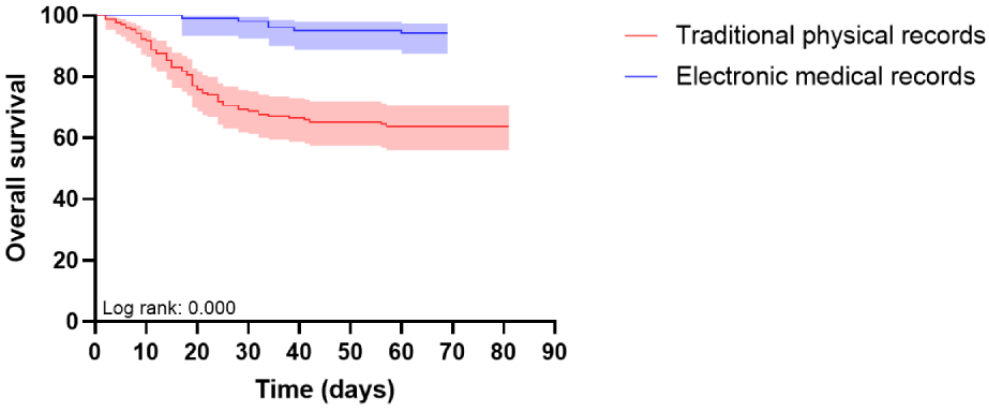
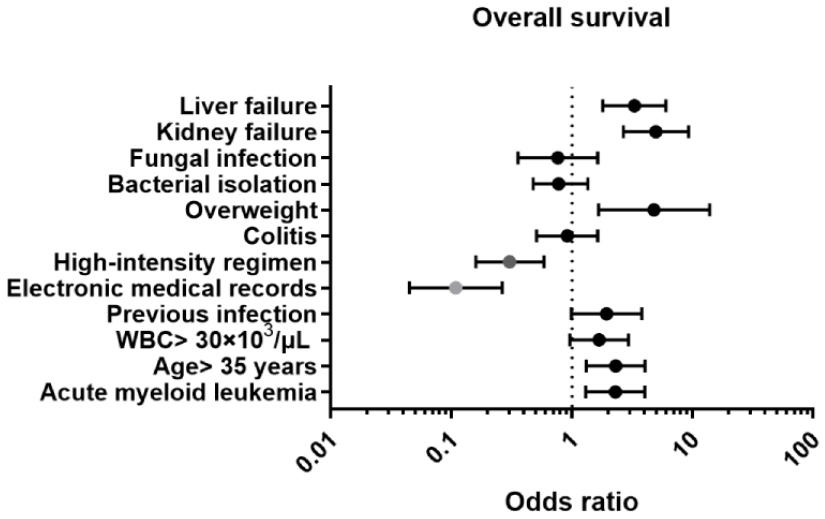


Figure 3. Kaplan-Meier curve comparing 60-day survival between the 2 hospital groups.



XSL•FO
RenderX

Figure 4. Forest plot of the odds ratios for the main factors associated with 60-day mortality. WBC: white blood cell.



Discussion

Principal Findings

In this retrospective natural experiment, 2 tertiary public hospitals caring for patients with newly diagnosed acute leukemia within the same national setting but with different documentation and medication-management environments were compared. The hospital using an integrated EMR had lower 60-day mortality, shorter times to antibiotic administration in patients with febrile neutropenia, and fewer procedural steps for urgent supportive care than the hospital using TPR. These findings are consistent with the hypothesis that digitalized workflows may improve the timeliness of critical inpatient actions during the first weeks of leukemia treatment, a period in which infectious and hemorrhagic complications strongly influence 60-day mortality.

Several considerations temper the interpretation of these findings. First, this was not an experimental comparison, and the exposure of interest operated at the hospital level rather than the patient level. The 2 hospitals differed in case mix, treatment patterns, and potentially in staffing, clinical experience, referral dynamics, waiting times, bed availability, socioeconomic context, and supportive-care resources. Residual confounding is therefore likely. Second, the study relied on routine clinical and administrative records, which may differ in completeness between paper-based and digital systems. The EMR center may have captured adverse events more completely because of more traceable documentation, whereas some complications may have been underrecorded in paper charts. Third, although all inpatient records are routinely reviewed at discharge by the institutional medical record committee, this quality control process does not eliminate the possibility of information bias in the research dataset. Accordingly, our results should be interpreted as an observed between-hospital association rather than proof of a direct causal effect of the EMR itself.

Despite therapeutic advances, outcomes for acute leukemia in Latin America remain closely tied to health-system performance, including timely diagnosis, rapid treatment initiation, reliable transfusion support, and prompt management of febrile neutropenia [19-24]. In that context, digital tools may matter less as isolated technologies than as enablers of coordination; they can shorten communication chains, document time stamps, reduce transcription delays, and facilitate faster validation of urgent orders.

This interpretation is supported by the marked difference in time to first antibiotic administration observed in our subgroup analysis. For patients with febrile neutropenia, even a modest delay can worsen clinical deterioration, particularly during induction therapy. Therefore, the shorter times observed in the EMR hospital are clinically plausible and may represent one mechanism linking digital workflow to lower 60-day mortality [22-25].

Digital systems may also strengthen accountability and continuous quality improvement. Because records are structured, time-stamped, and easier to audit, EMRs can help identify operational bottlenecks that remain largely invisible in paper-based systems. In hematology-oncology programs, where care depends on repeated coordination among physicians, nurses, pharmacists, blood bank staff, and laboratory services, this additional traceability may be especially valuable [26-29].

At the same time, the use of an integrated EMR should not be viewed as the only explanation for the observed differences. Digitalization likely functions as one component of a broader institutional environment that includes organizational culture, training, responsiveness of ancillary services, and medication availability. Therefore, the lower mortality observed in the EMR hospital may reflect both the digital platform itself and the systems-level practices that accompany its use.

The broader regional literature supports this interpretation. Studies from Latin America consistently show that standardized pathways, timely supportive care, and better-organized referral and treatment networks can improve survival in hematologic malignancies [30-32]. Our results add to that literature by suggesting that hospital digitalization may be one practical strategy for strengthening these pathways in public-sector

XSL•FO
RenderX

settings, but they do not establish that digitalization alone explains the observed mortality difference.

Rather than framing EMRs only as documentation tools, our findings support evaluating them as part of institutional redesign aimed at reducing preventable deaths within 60 days of admission. In acute leukemia, the clinically relevant question is not only whether data are stored electronically but also whether digital systems shorten the interval between clinical recognition, order entry, validation, and treatment delivery.

From a policy perspective, this distinction is important. Public hospitals operating with limited resources may derive meaningful benefit from digital tools when digitalization is linked to workflow simplification, clinical accountability, and rapid-response processes. Therefore, investments in EMRs are likely to be most effective when accompanied by staff training, supportive-care protocols, antimicrobial stewardship, and transfusion logistics optimization.

A major strength of this study is that it compares 2 real-world tertiary public hospitals caring for patients with acute leukemia within the same national health context but with clearly different documentation and medication-management models. The analysis also incorporated clinically relevant covariates associated with 60-day mortality and examined not only outcome differences but also a plausible process-of-care mechanism, namely time to antibiotic administration.

This study also has important limitations. Its retrospective design introduces the possibility of selection bias and information bias. Because only 2 hospitals were included, the analysis is best interpreted as a natural experiment and has limited generalizability. Residual confounding is likely, because unmeasured differences between centers—including staffing, referral complexity, waiting times, socioeconomic context, supportive-care capacity, and other institutional characteristics—may have influenced mortality. Although medical records at discharge are routinely reviewed by the institutional medical record committee, documentation quality may still differ between paper and digital systems. No formal duplicate abstraction or interrater reliability assessment was performed for the research dataset. Missing data were handled through complete-case analysis rather than imputation, and the antibiotic timing analysis was restricted to a subgroup with complete documentation, which may have introduced additional selection bias.

Future Directions

Future research should validate these findings in prospective multicenter cohorts, incorporate standardized abstraction procedures across institutions, and include more detailed measures of workflow, staffing, supportive-care capacity, and case severity. Studies evaluating cause-specific mortality, time to transfusion support, and cost-effectiveness would also help clarify which components of digital implementation have the greatest clinical impact.

Further work should also explore how digital systems interact with staff training, antimicrobial stewardship, transfusion logistics, and clinical decision support in patients with hematologic malignancies. Such studies would help distinguish the specific contribution of the EMR platform from the contribution of the broader institutional practices that accompany digital care.

Conclusions

In this retrospective natural experiment, the hospital using an integrated EMR had a lower 60-day mortality and faster antibiotic administration than the hospital using TPR in patients with acute leukemia. These results support the hypothesis that digitalized workflows may improve the timeliness of urgent supportive care during induction therapy.

However, the findings should be interpreted with caution because the study was observational and compared 2 nonequivalent institutions using routine clinical records. Prospective multicenter studies are needed to confirm whether the observed between-hospital differences persist after more detailed adjustment for case mix, workflow, and institutional characteristics, and to identify which components of digital care are most relevant for improving outcomes.

Acknowledgments

During the preparation of this paper, the authors used ChatGPT (GPT-5.3; OpenAI) to check spelling and grammar. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This research received no external funding.

Data Availability

Data are available from the corresponding author on reasonable request, subject to institutional privacy and confidentiality restrictions.

Authors' Contributions

Conceptualization: AGGR, CORP
Data curation: AGGR, CORP
Formal analysis: AGGR, GAA

XSL•FO
RenderX

Funding acquisition: CORP, RSC, ACG
Investigation: RSC, CORP
Methodology: ACG, CORP
Project administration: NGC, CORP
Resources: CORP, AMS, RSC
Software: CBD
Supervision: AMS
Validation: CORP, AMS, NGC
Visualization: CBD, GAA, RLT
Writing—original draft: ACG, CORP
Writing—review and editing: AMS, RSC, RLT
All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Zheng H, Jiang S. Frequent and diverse use of electronic health records in the United States: a trend analysis of national surveys. Digit Health 2022 Jul 06;8:20552076221112840 [FREE Full text] [doi: 10.1177/20552076221112840] [Medline: 35832476]
2. Tapuria A, Porat T, Kalra D, Dsouza G, Xiaohui S, Curcin V. Impact of patient access to their electronic health record: systematic review. Inform Health Soc Care 2021 Jun 02;46(2):192-204 [FREE Full text] [doi: 10.1080/17538157.2021.1879810] [Medline: 33840342]
3. Uslu A, Stausberg J. Value of the electronic medical record for hospital care: update from the literature. J Med Internet Res 2021 Dec 23;23(12):e26323 [FREE Full text] [doi: 10.2196/26323] [Medline: 34941544]
4. Aapro M, Bossi P, Dasari A, Fallowfield L, Gascón P, Geller M, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct;28(10):4589-4612 [FREE Full text] [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
5. Barrios CH, Werutsky G, Mohar A, Ferrigno AS, Müller BG, Bychkovsky BL, et al. Cancer control in Latin America and the Caribbean: recent advances and opportunities to move forward. Lancet Oncol 2021 Nov;22(11):e474-e487. [doi: 10.1016/S1470-2045(21)00492-7] [Medline: 34735817]
6. Buteau AC, Castelo-Loureiro A, Barragan-Carrillo R, Bejarano S, Kihn-Alarcón AJ, Soto-Perez-de-Celis E. Disparities in cancer control in Central America and the Caribbean. Hematol Oncol Clin North Am 2024 Feb;38(1):35-53. [doi: 10.1016/j.hoc.2023.07.007] [Medline: 37597998]
7. Turner BE, Steinberg JR, Weeks BT, Rodriguez F, Cullen MR. Race/ethnicity reporting and representation in US clinical trials: a cohort study. Lancet Reg Health Am 2022 Jul;11:1-12 [FREE Full text] [doi: 10.1016/j.lana.2022.100252] [Medline: 35875251]
8. Li L, Wang Y. Recent updates for antibody therapy for acute lymphoblastic leukemia. Exp Hematol Oncol 2020 Nov 27;9(1):33 [FREE Full text] [doi: 10.1186/s40164-020-00189-9] [Medline: 33292550]
9. Short NJ, Kantarjian H. Using immunotherapy and novel trial designs to optimise front-line therapy in adult acute lymphoblastic leukaemia: breaking with the traditions of the past. Lancet Haematol 2023 May;10(5):e382-e388. [doi: 10.1016/S2352-3026(23)00064-9] [Medline: 37003279]
10. Quiroz E, Aldoss I, Pullarkat V, Rego E, Marcucci G, Douer D. The emerging story of acute lymphoblastic leukemia among the Latin American population - biological and clinical implications. Blood Rev 2019 Jan;33:98-105. [doi: 10.1016/j.blre.2018.08.002] [Medline: 30126753]
11. Corrêa de Araujo Koury L, Ganser A, Berliner N, Rego EM. Treating acute promyelocytic leukaemia in Latin America: lessons from the International Consortium on Acute Leukaemia experience. Br J Haematol 2017 Jun;177(6):979-983. [doi: 10.1111/bjh.14589] [Medline: 28466545]
12. Hansen BA, Wendelbo Ø, Bruserud Ø, Hemsing AL, Mosevoll KA, Reikvam H. Febrile neutropenia in acute leukemia. Epidemiology, etiology, pathophysiology and treatment. Mediterr J Hematol Infect Dis 2020;12(1):e2020009 [FREE Full text] [doi: 10.4084/MJHID.2020.009] [Medline: 31934319]
13. Lopez-Garcia YK, Valdez-Carrizales M, Nuñez-Zuno JA, Apodaca-Chávez E, Rangel-Patiño J, Demichelis-Gómez R. Are delays in diagnosis and treatment of acute leukemia in a middle-income country associated with poor outcomes? A retrospective cohort study. Hematol Transfus Cell Ther 2024;46(4):366-373 [FREE Full text] [doi: 10.1016/j.htct.2023.05.010] [Medline: 37442650]
14. Stahl M, Tallman MS. Acute promyelocytic leukemia (APL): remaining challenges towards a cure for all. Leuk Lymphoma 2019 Dec;60(13):3107-3115 [FREE Full text] [doi: 10.1080/10428194.2019.1613540] [Medline: 31842650]
15. Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. Mhealth 2017;3:38 [FREE Full text] [doi: 10.21037/mhealth.2017.08.07] [Medline: 29184890]

XSL•FO
RenderX

16. Kantarjian H, Thomas D, O'Brien S, Cortes J, Giles F, Jeha S, et al. Long-term follow-up results of hyperfractionated cyclophosphamide, vincristine, doxorubicin, and dexamethasone (Hyper-CVAD), a dose-intensive regimen, in adult acute lymphocytic leukemia. Cancer 2004 Dec 15;101(12):2788-2801 [FREE Full text] [doi: 10.1002/cncr.20668] [Medline: 15481055]
17. Stock W, Luger SM, Advani AS, Yin J, Harvey RC, Mullighan CG, et al. A pediatric regimen for older adolescents and young adults with acute lymphoblastic leukemia: results of CALGB 10403. Blood 2019 Apr 04;133(14):1548-1559 [FREE Full text] [doi: 10.1182/blood-2018-10-881961] [Medline: 30658992]
18. Roman Diaz JL, Vazquez Martinez M, Khimani F. New approaches for the treatment of AML beyond the 7+3 regimen: current concepts and new approaches. Cancers (Basel) 2024 Feb 05;16(3):677 [FREE Full text] [doi: 10.3390/cancers16030677] [Medline: 38339429]
19. Rivera-Luna R, Perez-Vera P, Galvan-Diaz C, Velasco-Hidalgo L, Olaya-Vargas A, Cardenas-Cardos R, et al. Triple-hit explanation for the worse prognosis of pediatric acute lymphoblastic leukemia among Mexican and Hispanic children. Front Oncol 2022;12:1072811 [FREE Full text] [doi: 10.3389/fonc.2022.1072811] [Medline: 36531028]
20. Rego EM, Jácomo RH. Epidemiology and treatment of acute promyelocytic leukemia in Latin America. Mediterr J Hematol Infect Dis 2011;3(1):e2011049 [FREE Full text] [doi: 10.4084/MJHID.2011.049] [Medline: 22110899]
21. Martínez-Anaya D, Moreno-Lorenzana D, Reyes-León A, Juárez-Figueroa U, Dean M, Aguilar-Hernández MM, et al. Characterization of philadelphia-like pre-B acute lymphoblastic leukemia: experiences in Mexican pediatric patients. Int J Mol Sci 2022 Aug 24;23(17):9587 [FREE Full text] [doi: 10.3390/ijms23179587] [Medline: 36076986]
22. O'Horo JC, Marcelin JR, Abu Saleh OM, Barwise AK, Odean PM, Rivera CG, et al. Standardizing febrile neutropenia management: antimicrobial stewardship in the hematologic malignancy population. J Oncol Pract 2019 Sep;15(9):e843-e848 [FREE Full text] [doi: 10.1200/JOP.18.00775] [Medline: 31322989]
23. Keck JM, Wingler MJ, Cretella DA, Vijayvargiya P, Wagner JL, Barber KE, et al. Approach to fever in patients with neutropenia: a review of diagnosis and management. Ther Adv Infect Dis 2022;9:20499361221138346 [FREE Full text] [doi: 10.1177/20499361221138346] [Medline: 36451936]
24. Ornelas-Sánchez M, Nuño-Vázquez L, Loera-Reyna A, Torres-Reyes D, Rivera-Gómez R, Sánchez A, et al. The "Golden Hour": a capacity-building initiative to decrease life-threating complications related to neutropenic fever in patients with hematologic malignancies in low- and middle-income countries. Blood Adv 2018 Nov 30;2(Suppl 1):63-66 [FREE Full text] [doi: 10.1182/bloodadvances.2018GS112240] [Medline: 30504206]
25. Albagmi S. The effectiveness of EMR implementation regarding reducing documentation errors and waiting time for patients in outpatient clinics: a systematic review. F1000Res 2021;10:514 [FREE Full text] [doi: 10.12688/f1000research.45039.2] [Medline: 35035887]
26. Poissant L, Pereira J, Tamblyn R, Kawasumi Y. The impact of electronic health records on time efficiency of physicians and nurses: a systematic review. J Am Med Inform Assoc 2005;12(5):505-516 [FREE Full text] [doi: 10.1197/jamia.M1700] [Medline: 15905487]
27. Llera AS. A fresh perspective on Latin America cancer care: uncovering hidden messages in unconventional data sources. Lancet Reg Health Am 2023 Aug;24:100559 [FREE Full text] [doi: 10.1016/j.lana.2023.100559] [Medline: 37533928]
28. Khela H, Khalil J, Daxon N, Neilson Z, Shahrokhi T, Chung P, et al. Real world challenges in maintaining data integrity in electronic health records in a cancer program. Tech Innov Patient Support Radiat Oncol 2024 Mar;29:100233 [FREE Full text] [doi: 10.1016/j.tipsro.2023.100233] [Medline: 38293266]
29. Weis JM, Levy PC. Copy, paste, and cloned notes in electronic health records: prevalence, benefits, risks, and best practice recommendations. Chest 2014 Mar 01;145(3):632-638. [doi: 10.1378/chest.13-0886] [Medline: 24590024]
30. Hungria V, Gaiolla R, Galvez K, Remaggi G, Schutz N, Bittencourt R, et al. Health care systems as determinants of outcomes in multiple myeloma: final results from the Latin American MYLACRE study. Blood Adv 2025 Mar 25;9(6):1293-1302 [FREE Full text] [doi: 10.1182/bloodadvances.2024013838] [Medline: 39657126]
31. Torres-Roman JS, Valcarcel B, Guerra-Canchari P, Santos CA, Barbosa IR, La Vecchia C, et al. Leukemia mortality in children from Latin America: trends and predictions to 2030. BMC Pediatr 2020 Nov 07;20(1):511 [FREE Full text] [doi: 10.1186/s12887-020-02408-y] [Medline: 33160309]
32. Farias MA, Badino M, Marti M, Báscolo E, García Saisó S, D'Agostino M. Digital transformation as a strategy to strengthen essential public health functions in the Americas [Article in Spanish]. Rev Panam Salud Publica 2023 Nov 15;47:e150 [FREE Full text] [doi: 10.26633/RPSP.2023.150] [Medline: 38024443]

Abbreviations

ALL: acute lymphoblastic leukemia
AML: acute myeloid leukemia
EMR: electronic medical record
FLAG: fludarabine, cytarabine, granulocyte colony-stimulating factor
HITECH: Health Information Technology for Economic and Clinical Health
OR: odds ratio

XSL•FO
RenderX

TPR: traditional physical record

Edited by M Balcarras; submitted 14.Jan.2025; peer-reviewed by YL Wang, P Sripornsawan; comments to author 12.May.2025; accepted 12.Mar.2026; published 17.Apr.2026.

Please cite as:
Ramos Peñafiel CO, Cabrera García Á, Balderas Delgado C, Luna Tentle R, Milán Salvatierra A, Godínez Cubillo N, Acosta Altamirano G, Sánchez Conejo R, Gallardo Rodríguez AG
Digital and Paper-Based Hospital Workflows and 60-Day Mortality in Acute Leukemia: Retrospective Natural Experiment
JMIR Cancer 2026;12:e71306
URL: https://cancer.jmir.org/2026/1/e71306
doi:10.2196/71306
PMID:

©Christian Omar Ramos Peñafiel, Álvaro Cabrera García, Carolina Balderas Delgado, Rocío Luna Tentle, Andrea Milán Salvatierra, Nora Godínez Cubillo, Gustavo Acosta Altamirano, Rosa Sánchez Conejo, Adán Germán Gallardo Rodríguez. Originally published in JMIR Cancer (https://cancer.jmir.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessing the Accuracy, Completeness, and Reference Quality of GPT-4 and Google for Gynecologic Cancer Information: Comparative Quantitative Study

Alexandra H Smick[1], MD; Martins Ayoola[1], MD; Rajeshree Rajpara[2], BS; Isabel Lazo[3], MD; Dana M Chase[1], MD

[1]Gynecologic Oncology, University of California, Los Angeles, Los Angeles, CA, United States
[2]Undergraduate College, University of California, Los Angeles, Los Angeles, CA, United States
[3]Gynecologic Oncology, Olive View-UCLA Medical Center, Sylmar, CA, United States

Corresponding Author:
Alexandra H Smick, MD
Gynecologic Oncology
University of California, Los Angeles
100 Medical Plaza Suite 383
Los Angeles, CA, 90095
United States
Phone: 1 310 794 7274
Email: asmick14@gmail.com

Abstract

Background: Patients with newly diagnosed gynecologic cancers often seek information online, but the quality of available resources may be inconsistent. Although GPT-4 may offer an alternative to traditional internet search engines, its performance remains largely under-studied in gynecologic oncology.

Objective: This study aimed to compare the completeness, accuracy, and reference quality of responses generated by GPT-4 with those generated by Google in clinical scenarios involving a new diagnosis of a gynecologic cancer.

Methods: Clinical scenarios representing early- and advanced-stage endometrial, ovarian, and cervical cancers were developed by gynecologic oncologists using publicly available patient education materials. Each scenario included 4 standardized questions addressing etiology, prognosis, treatment, and treatment efficacy. GPT-4 and Google were queried for each question, with new sessions for GPT-4 and private browsing for Google to minimize bias. Responses were independently evaluated by 4 gynecologic oncology experts who were blinded to each other's ratings. Accuracy was scored on a 6-point Likert scale; completeness and reference quality were scored on 3-point scales. Reference quality was categorized as low (commercial), medium (institutional or government), or high (peer reviewed). Optional free-text reviewer comments were collected and summarized descriptively to provide context for the quantitative findings. Descriptive statistics and Wilcoxon signed-rank tests were used for analysis.

Results: Across 6 clinical scenarios and 21 standardized questions (N=84 total responses), GPT-4 outperformed Google across all evaluated domains. The median accuracy score was 6.00 (IQR 5.00-6.00) for GPT-4 and 5.00 (IQR 4.00-6.00) for Google (P=.04). The median completeness score was 3.00 (IQR 3.00-3.00) for GPT-4 and 2.00 (IQR 1.00-3.00) for Google (P=.009). Reference quality was also higher for GPT-4, with a median score of 3.00 (IQR 3.00-3.00) compared to 2.00 (IQR 2.00-2.00) for Google (P=.009). Reviewer comments noted that GPT-4 provided more accurate, comprehensive, and personalized responses, while Google returned less detailed content from general consumer health websites rather than peer-reviewed sources.

Conclusions: GPT-4 may serve as a reliable and high-quality resource for information in gynecologic oncology. Compared to Google, it delivered more accurate and complete content, with higher-quality references. Further research is needed to assess the readability and accessibility of GPT-4–generated content across diverse patient populations.

(JMIR Cancer 2026;12:e76471) doi:10.2196/76471

KEYWORDS

artificial intelligence; AI; gynecologic cancer; patient education as topic; large language models; gynecologic neoplasms; health literacy; cancer communication; patient information materials

XSL•FO
RenderX

Introduction

Approximately 100,000 individuals are diagnosed with gynecologic malignancy each year in the United States [1]. Following a new cancer diagnosis, patients frequently seek supplemental educational materials to better understand their condition and treatment options. Traditional resources, such as those from professional societies and academic institutions, are often comprehensive but presented in generalized formats, requiring patients to interpret the content in the context of their own diagnosis and clinical stage [2].

AI, particularly large language models (LLMs) such as GPT-4, has gained significant attention in medicine due to its ability to generate rapid, detailed responses based on vast amounts of textual data. ChatGPT, developed by OpenAI, is powered by the GPT-4 language model and accessible to the public through the ChatGPT Plus subscription. In contrast, Google is a search engine that retrieves existing content from across the internet using proprietary algorithms [3,4]. As of mid-2024, Google began integrating its own LLM, Gemini, into certain aspects of its platform. However, the standard search interface continues to primarily direct users to external web pages rather than generating original text. Previous research has explored the utility of LLMs for clinical tasks such as staging lung cancer, summarizing histopathology reports, and supporting diagnostic decision-making [5,6]. Although early findings are promising, concerns remain regarding the accuracy, reliability, and lack of citation transparency of AI-generated content, reinforcing the need for expert oversight.

As GPT-4 becomes increasingly accessible to the public, patients are likely to turn to it, alongside traditional internet search engines, for information about their diagnosis [2]. Prior studies in oncology have evaluated the quality and readability of GPT-4's responses to patient questions, highlighting variability in content accuracy and source quality [7,8]. Other investigations have emphasized the potential of AI to improve health literacy by simplifying medical content, yet its role in reliably supporting patient education remains unknown [9]. Within gynecologic oncology specifically, research on AI-based educational tools is limited. Early studies have examined GPT-4's ability to provide guideline-concordant recommendations for ovarian cancer diagnosis and management, but no prior work has systematically evaluated its performance in delivering high-quality informational content to newly diagnosed patients across multiple gynecologic cancers [10,11].

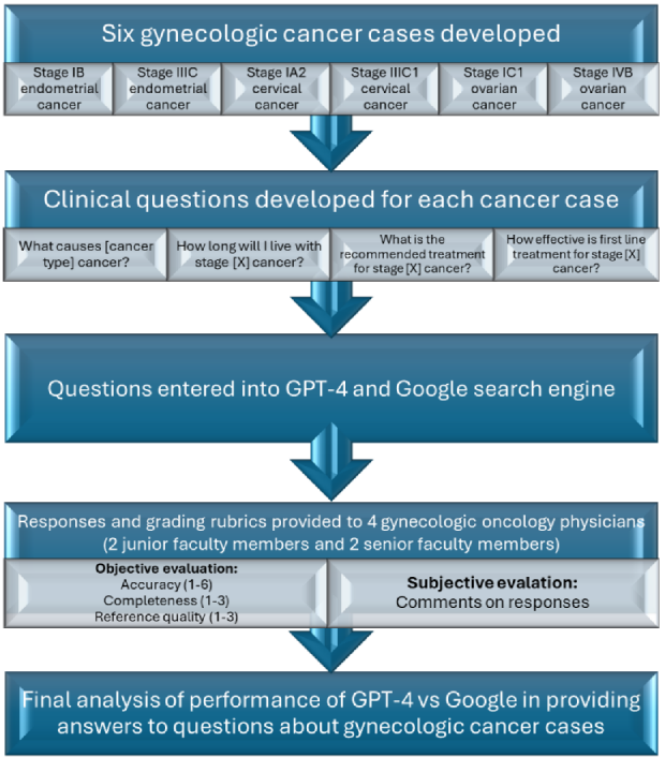
This study aimed to assess GPT-4's effectiveness in generating patient-directed informational content in the context of newly diagnosed gynecologic cancers. We compared GPT-4 responses with Google search results across standardized clinical scenarios, focusing on 3 key domains: accuracy, completeness, and reference quality. Our findings contribute to the growing body of literature evaluating AI's role in digital health and its potential integration into patient-centered cancer education.

Methods

Study Design

This was a comparative, cross-sectional evaluation of an AI-based LLM (GPT-4; OpenAI) and a standardized internet search engine (Google; Alphabet Inc) using predefined gynecologic oncology clinical scenarios and patient-centered questions (Figure 1). Methods were structured and reported in accordance with the CHART (Chatbot Health Advice Reporting Tool) guideline for chatbot health advice studies [12].

Figure 1. Overview of the study design and evaluation process.



XSL•FO
RenderX

Ethical Considerations

Ethics approval was not required for this study because it did not involve human participants, patient data, or animal subjects. The study evaluated outputs from publicly accessible AI tools and internet search engines and therefore did not constitute human subjects research.

Models

ChatGPT responses were generated using the GPT-4 model via the paid ChatGPT Plus subscription (OpenAI [13]). Google searches were conducted using the Chrome browser.

Prompt Engineering

Six clinical scenarios representing common presentations of gynecologic cancers were developed by a panel of gynecologic oncology providers. For each cancer type (ovarian, endometrial, and cervical), 7 patient-centered questions were created based on commonly asked topics in clinical practice and information available from professional society websites [14], yielding a total of 21 unique questions (Textbox 1). Questions were entered verbatim into both platforms, with each question generating 1 response from GPT-4 and 1 response from Google. A new ChatGPT session was initiated prior to each question to prevent learning across questions or contextual carryover.

Accuracy was assessed using a 6-point Likert scale: 1="completely incorrect," 2="more incorrect than correct," 3="approximately equal incorrect and correct," 4="more correct than incorrect," 5="nearly all correct," and 6="correct." Completeness was assessed using a 3-point Likert scale: 1="incomplete," addressing some aspects of the question but with significant parts missing or incomplete; 2="adequate," addressing all aspects of the question and providing the minimum information required to be complete; and 3="comprehensive," addressing all aspects of the question and providing additional information or context beyond what was expected. Reference quality was assessed using a 3-point Likert scale: 1="bottom tier," including any .com sources such as WebMD; 2="middle tier," including any .gov or .org sources such as the American Cancer Society and National Institutes of Health; and 3="top tier," including peer-reviewed sources or randomized controlled trials such as the National Comprehensive Cancer Network and UpToDate.

Textbox 1. Clinical questions about diagnosis, treatment options, and survival outcomes for common gynecologic malignancies.

1. What causes endometrial cancer?
2. How long will I live with stage IB endometrial cancer?
3. What is the recommended treatment for stage IB endometrial cancer?
4. How effective is first-line treatment for stage IB endometrial cancer?
5. How long will I live with stage IIIC endometrial cancer?
6. What is the recommended treatment for stage IIIC endometrial cancer?
7. How effective is first-line treatment for stage IIIC endometrial cancer?
8. What causes cervical cancer?
9. How long will I live with stage IA2 cervical cancer?
10. What is the recommended treatment for stage IA2 cervical cancer?
11. How effective is first-line treatment for stage IA2 cervical cancer?
12. How long will I live with stage IIIC1 cervical cancer?
13. What is the recommended treatment for stage IIIC1 cervical cancer?
14. How effective is first-line treatment for stage IIIC1 cervical cancer?
15. What causes ovarian cancer?
16. How long will I live with stage IC1 ovarian cancer?
17. What is the recommended treatment for stage IC1 ovarian cancer?
18. How effective is first-line treatment for stage IC1 ovarian cancer?
19. How long will I live with stage IVB ovarian cancer?
20. What is the recommended treatment for stage IVB ovarian cancer?
21. How effective is first-line treatment for stage IVB ovarian cancer?

Query Strategy

Google searches were conducted in private browsing mode with the browser cache cleared. For each query, the top-ranked nonsponsored web page was recorded to allow direct comparison with GPT-4's single-response output (Multimedia Appendix 1). This approach reflects typical user behavior and aligns with prior studies demonstrating reliance on the first search results [15].

XSL•FO
RenderX

Performance Evaluation

Responses from both platforms were evaluated for accuracy, completeness, and reference quality by 4 gynecologic oncology physicians from 2 institutions (1 academic tertiary referral center and 1 county-based safety-net hospital). Reviewers represented a range of clinical experience, with 2 physicians having <5 years of clinical experience and 2 having >5 years of clinical experience. Reviewers scored responses independently and were blinded to each other's evaluations. Reviewers were also invited to provide optional free-text comments within the standardized evaluation form. These qualitative comments were summarized descriptively. Comments were reviewed in their entirety and grouped into thematic categories reflecting perceptions of accuracy, completeness, reference quality, clarity, and patient comprehensibility. A total of 76 optional comments were submitted across all reviewers. Representative comments were selected to illustrate common themes across platforms and are presented in Textbox 2. This descriptive approach aligns with conventional qualitative content analysis methods used to contextualize the quantitative findings [16]. Accuracy was defined as concordance with National Comprehensive Cancer Network guidelines and current literature and scored using a 6-point Likert scale. Completeness was scored using a 3-point Likert scale based on the inclusion of major concepts relevant to each question. Reference quality was categorized as low, medium, or high and scored using a 3-point Likert scale.

Textbox 2. Representative reviewer comments on completeness, accuracy, and reference quality for Google and GPT-4.

Google

- Technically all the information was correct, but it did not answer the specific question. They would likely need to answer this question with another source.
- The question is not answered, only treatments are mentioned. It also makes it sound like the treatments may not be that effective.
- The information looks very accurate, but there is a lot of set up with prognosis and you must scroll down to get your answer.
- Having to weed through the information to get the answer, the information is not as clear and is a little confusing.
- Some information is outdated.
- Very general information which makes it easy for a person with average literacy to understand.
- Answers to survival questions are given but there is no information about the cancer stage.
- There is no mention of treatment effectiveness or immunotherapy.

GPT-4

- Accurate information is given that is straight to the point and concise.
- Complete and thorough answers were given that were concise with high quality resources.
- ChatGPT does a nice job of putting information into context for the patient.
- ChatGPT is very comprehensive, although it does not address immunotherapy as an option.
- There is specific data of effectiveness given with very thorough data on treatment options.
- There is very good information provided on surgery and chemotherapy (including details about bevacizumab and PARPi) and also discusses palliative care.
- Information is concise and complete. It talked about what the stage meant.
- My only concern is the patient's reading level and ability to understand the language in CGPT versus the broad, simple language used by Google.

Overall

- I thought I was going to come in biased against ChatGPT. As I looked through the answers, I think ChatGPT did much better answering questions than a random Google search with surprisingly high quality, nuanced answers.
- Google searches with answers from Mayo, COH, ACS were good for basic questions, like what is this cancer, what is the stage; although it gave you all the information on staging and you must scroll to your particular area of interest.
- ChatGPT did a better job at individualizing the answer. It also did better at prognosis, effectiveness questions. It seemed to do it with nuance and "compassion." But it was also direct, and I think that can be helpful for physicians when it is difficult to give an accurate prognosis.

Sample Size

A total of 21 standardized patient-centered questions were evaluated. Each question generated 1 response from GPT-4 and 1 response from Google, resulting in 21 responses per platform. All responses were independently evaluated by 4 gynecologic oncology physicians, yielding 84 physician evaluations per platform. The number of questions was selected to capture key informational domains, including etiology, prognosis, treatment, and treatment efficacy, while maintaining feasibility for blinded expert review. This approach is consistent with prior studies evaluating the quality of AI-generated and online health information [8,10,11,17-19].

XSL•FO
RenderX

Data Analysis

Statistical analyses were conducted using JMP Pro (version 16.0.0; SAS Institute Inc). Descriptive statistics were calculated for all outcomes, and paired comparisons between platforms were performed using Wilcoxon signed-rank tests. Interrater reliability was assessed using intraclass correlation coefficients (ICCs) derived from a 2-way random-effects model with single measures and consistency (ICC[2,1]). Reliability thresholds were interpreted using established criteria [20].

Results

Accuracy was graded by 4 physicians on a Likert scale of 1 to 6, with 1="completely incorrect" and 6="completely correct." The median score was 5.00 (IQR 4.00-6.00) for Google compared to 6.00 (IQR 5.00-6.00) for GPT-4 (*P*=.04). For Google, the distribution by accuracy score from 1 to 6 was 6% (5/84), 2.4% (2/84), 8.3% (7/84), 16.7% (14/84), 8.3% (7/84), and 58.3% (49/84), respectively. For GPT-4, the distribution from 1 to 6 was 0% (0/84), 0% (0/84), 2.4% (2/84), 10.7% (9/84), 17.8% (15/84), and 69% (58/84), respectively. A score of 5 or 6 (the most accurate) accounted for 66.6% (56/84) of responses for Google compared to 86.9% (73/84) for GPT-4 (Table 1).

Completeness was then graded by 4 physicians on a Likert scale of 1 to 3, with 1="incomplete," addressing only some aspects of the questions, and 3="comprehensive," addressing all aspects of the questions. The median score was 2.00 (IQR 1.00-3.00) for Google compared to 3.00 (IQR 3.00-3.00) for GPT-4 (*P*=.009). For Google, the distribution by completeness score from 1 to 3 was 36.9% (31/84), 28.6% (24/84), and 34.5% (29/84), respectively. For GPT-4, the distribution from 1 to 3 was 1.2% (1/84), 17.8% (15/84), and 81% (68/84), respectively. A score of 2 or 3 (mostly complete) accounted for 63.1% (53/84) of responses for Google compared to 98.8% (83/84) for GPT-4 (Table 2).

Finally, reference quality was graded by 4 physicians on a 3-point Likert scale, with 1="bottom tier (.com)" and 3="top tier (peer reviewed)." The median score was 2.00 (IQR 2.00-2.00) for Google compared to 3.00 (IQR 3.00-3.00) for GPT-4 (*P*=.009). For Google, the distribution by reference quality score from 1 to 3 was 10.7% (9/84), 77.4% (65/84), and 11.9% (10/84), respectively. For GPT-4, the distribution from 1 to 3 was 0% (0/84), 22.6% (19/84), and 77.4% (65/84), respectively. A score of 2 or 3 (highest quality references) accounted for 89.3% (75/84) of responses for Google compared to 100% (84/84) for GPT-4 (Table 3).

Reviewers then provided comments on their perceptions of the accuracy, completeness, and reference quality of GPT-4 compared to Google. A total of 76 optional comments were submitted across all reviewers. Representative comments illustrating the major themes identified across reviewers are presented in Textbox 2. Overall, reviewers noted that Google source responses had clear and precise language but often did not completely answer the question and did not provide enough information on prognosis or treatment options. For GPT-4, reviewers generally commented that answers were very complete and comprehensive, especially regarding disease stage and prognosis, but might be more challenging for patients to understand.

To assess the consistency of reviewer scoring, we calculated interrater reliability using ICCs based on a 2-way random-effects model evaluating consistency. Completeness scores for Google demonstrated moderate consistency (ICC 0.595, 95% CI 0.389-0.782), while all other domains showed poor consistency (ICC<0.50), except for GPT-4 reference quality, which demonstrated excellent interrater reliability (ICC 0.935, 95% CI 0.879-0.970). Full results are shown in Table 4.

Table 1. Comparison of accuracy scores for clinical scenarios between GPT-4 and Google.

Measure	Google (n=84)	GPT-4 (n=84)
Accuracy score 1, n (%)	5 (6)	0 (0)
Accuracy score 2, n (%)	2 (2.4)	0 (0)
Accuracy score 3, n (%)	7 (8.3)	2 (2.4)
Accuracy score 4, n (%)	14 (16.7)	9 (10.7)
Accuracy score 5, n (%)	7 (8.3)	15 (17.8)
Accuracy score 6, n (%)	49 (58.3)	58 (69)
Mean (SD)[a]	4.94 (1.51)	5.54 (0.78)
Median (IQR)[a]	6.00 (4.00-6.00)	6.00 (5.00-6.00)

[a]*P*=.04.

XSL•FO
RenderX

Table 2. Comparison of completeness scores for clinical scenarios between GPT-4 and Google.

Measure	Google (n=84)	GPT-4 (n=84)
Completeness score 1, n (%)	31 (36.9)	1 (1.2)
Completeness score 2, n (%)	24 (28.6)	15 (17.8)
Completeness score 3, n (%)	29 (34.5)	68 (81)
Mean (SD)[a]	1.98 (0.85)	2.80 (0.43)
Median (IQR)[a]	2.00 (1.00-3.00)	3.00 (3.00-3.00)

[a]*P*=.009.

Table 3. Comparison of reference quality scores for clinical scenarios between GPT-4 and Google.

Measure	Google (n=84)	GPT-4 (n=84)
Reference quality score 1, n (%)	9 (10.7)	0 (0)
Reference quality score 2, n (%)	65 (77.4)	19 (22.6)
Reference quality score 3, n (%)	10 (11.9)	65 (77.4)
Mean (SD)[a]	2.01 (0.48)	2.77 (0.42)
Median (IQR)[a]	2.00 (2.00-2.00)	3.00 (3.00-3.00)

[a]*P*=.009.

Table 4. Interrater reliability by source and scoring domain.

Source and scoring domains	Interclass correlation[a] (95% CI)	Interpretation[b]
Google		
Accuracy	0.075 (−0.087 to 0.327)	Poor
Completeness	0.595 (0.389 to 0.782)	Moderate
Reference quality	0.016 (−0.128 to 0.255)	Poor
GPT-4		
Accuracy	0.111 (−0.061 to 0369)	Poor
Completeness	0 (−0.138 to 0.234)	Poor
Reference quality	0.935 (0.879 to 0.970)	Excellent

[a]Intraclass correlation coefficients were calculated using a 2-way random-effects model with single measures (ICC[2,1]).

[b]Poor<0.5, moderate=0.5-0.75, good=0.75-0.90, and excellent>0.90.

Discussion

Principal Results

This study found that GPT-4 performed better than Google sources when providing high-quality responses to clinical scenarios involving a new diagnosis of gynecologic cancer. Physician reviewers rated GPT-4 significantly higher than Google in terms of accuracy, completeness, and reference quality. Reviewers were surprised by the quality of the GPT-4 answers and references and saw the potential for GPT-4 to better address patient-specific questions regarding diagnosis, prognosis, and treatment options with both precise and comprehensive information. The most common shortcoming identified in the Google-sourced content was the provision of general information that was not specific to the patient's cancer diagnosis, stage, or treatment options and therefore was difficult to individually interpret.

Comparison With Prior Work

Although a growing number of studies have evaluated GPT-4 for cancer-related patient education, relatively few have focused specifically on gynecologic oncology. Initial research on AI in medicine focused on its diagnostic role and potential for clinical decision-making. For example, in otorhinolaryngology, GPT-4 was compared to UpToDate in terms of usefulness and reliability in diagnosing certain common clinical scenarios. UpToDate was found to be more accurate than GPT-4 (*P*=.009), and the authors concluded that improvements were needed in the usefulness and reliability of GPT-4's evidence-based knowledge [21]. Similarly, in various clinical cases of thyroid cancer, GPT-4 was found to be reasonably accurate (approximately 76.7%) when providing information about thyroid cancer, but recommendations about evaluation, treatment, and follow-up were deemed general and inadequate [22].

XSL•FO
RenderX

More recently, studies in gynecologic oncology have demonstrated that GPT-4 generally provides guideline-concordant recommendations for gynecologic cancer diagnosis and management, while emphasizing the continued importance of physician oversight for complex clinical decision-making [10,11,23]. Chou et al [17] further reported that GPT-4 generated patient-directed responses regarding ovarian cancer that were comparable to those provided by gynecologic oncology clinicians, supporting its potential role as a patient education tool.

As AI is increasingly explored as a tool for patient-facing communication, several prior studies have evaluated the use of GPT-4 to generate patient information. One study used various LLM platforms to generate patient education materials for both common and rare dermatologic conditions and found that these platforms were able to create accessible, understandable content at an appropriate reading level [18]. Another study focused on the ability of GPT-4 to provide patients with information about screening mammography but documented concerns about the understandability and actionability of the answers, specifically from a health literacy standpoint [19]. Although AI seems to offer an opportunity to enhance patients' access to health information, there continues to be concern about the possibility of misinformation and the quality of references.

Our findings build upon this growing body of literature. Compared to top-ranked Google search results, GPT-4 consistently outperformed in terms of accuracy, completeness, and reference quality. These findings are consistent with prior studies demonstrating that GPT-4 generally provides accurate, comprehensive, and guideline-concordant information across a variety of cancer-related clinical scenarios [10,11,17,22]. Consistent with the findings of Chou et al [17], we found that GPT-4 generated comprehensive, patient-directed responses that were highly rated by gynecologic oncology physicians. Similarly, our findings align with those of Piazza et al [10] and Finch et al [11], who reported strong agreement between ChatGPT-generated recommendations and established ovarian cancer guidelines. However, although previous studies primarily compared ChatGPT with clinical guidelines or clinician responses, our study directly compared GPT-4 with a conventional Google search strategy, reflecting a common approach used by patients when seeking information after a new cancer diagnosis. Reviewers noted that GPT-4 responses were generally more comprehensive and better aligned with the clinical context, often including relevant treatment information and prognostic details that were lacking in standard search results. Although previous concerns have been raised about the citation quality of previous GPT-4 models, our findings suggest that, in this study, the reference quality was significantly higher than that of Google search results [3,4,8,19,22]. Overall, these results support the potential of GPT-4 to generate personalized, high-quality responses that may enhance patient understanding, support shared decision-making, and improve access to reliable health information.

Strengths and Limitations

Strengths of this study include the novel design used to investigate the role of AI in generating informational content related to gynecologic malignancies. By querying GPT-4 and Google with multiple patient scenarios, results were more generalizable to patients with different types and stages of gynecologic cancers. The use of gynecologic oncology physicians, who were blinded to each other's assessments, enabled the clinical evaluation of response quality by experts who routinely guide patients through diagnosis and treatment. Evaluation of accuracy, completeness, and reference quality offered a comprehensive assessment of the content delivered by each platform. Importantly, we also assessed interrater reliability, which is often omitted in similar studies. Although agreement varied by domain, the inclusion of this analysis strengthens the rigor of our evaluation.

Some limitations of this study include a relatively small number of reviewers, a subjective grading scale, and no direct assessment of the readability of the sources used. Although reviewers were blinded to each other's ratings, they were aware of the source of each response (GPT-4 or Google), which may have introduced bias. Additionally, all 4 reviewers were from only 2 institutions, which raises the possibility that assessments may have been influenced by local practice patterns or institutional standards, potentially limiting generalizability. Interrater reliability was variable across domains, which likely reflects the inherently subjective nature of evaluating informational content, particularly when responses were largely accurate and complete, resulting in clustering of scores at the higher end of the grading scales. In this context, small differences in individual scoring behavior may disproportionately affect interrater reliability metrics despite overall agreement in comparative performance between platforms. Another important consideration is that the phrasing of questions may influence the quality and clarity of responses generated by LLMs. In this study, the scenarios were designed to reflect the types of questions a well-informed patient might ask after receiving a new cancer diagnosis. In real-world settings, however, patients may not know what questions to ask, may lack the background to phrase them effectively, or may not have the clinical knowledge needed to interpret detailed responses. GPT-4 does not tailor responses based on a user's health literacy level, and this study did not formally assess the readability or accessibility of the responses. This represents an important limitation, as prior research has shown that the effectiveness of patient education materials depends not only on content accuracy but also on whether they are comprehensible and actionable. Additionally, by evaluating only the top-ranked Google result for each query, the study may not fully capture the breadth of information users might access through additional searches.

Since completing this study, Google has expanded the integration of LLM-generated summaries within its search platform. This underscores the rapidly evolving landscape of online health information and highlights the need for continued evaluation of AI-assisted search tools. Although such integrations may narrow performance gaps between traditional search engines and stand-alone LLMs, our findings remain relevant by illustrating the potential advantages of conversational, context-aware responses over static web page retrieval. Future studies should directly compare AI-generated

XSL•FO
RenderX

summaries embedded within search engines with stand-alone LLMs, with a focus on reference transparency, patient personalization, and health literacy.

Conclusions

The results of this study suggest that GPT-4 may be a useful tool for delivering accurate, comprehensive, and well-referenced informational content in gynecologic oncology when compared with top-ranked Google search results within the context of standardized clinical scenarios. However, these findings should be interpreted cautiously given several important methodological limitations, including incomplete blinding of reviewers to response source, variable interrater reliability across evaluation domains, and the restriction of the Google comparator to a single top-ranked result per query. Accordingly, this study should be viewed as a preliminary, hypothesis-generating evaluation rather than a definitive assessment of comparative performance. Although LLMs such as GPT-4 show promise in generating context-aware responses tailored to specific clinical scenarios, further research is needed to clarify their appropriate role in supporting patient education and health care professional–patient communication in gynecologic oncology.

Acknowledgments

The authors declare that generative AI (GPT-4; OpenAI) was used as described in the Methods section.

Funding

The authors declare that no financial support was received for this work.

Authors' Contributions

Conceptualization: AHS, IL, DMC
Data collection: AHS, MA, IL, RR, DMC
Statistical analysis: AHS, MA
Writing—original manuscript: AHS, RR
Writing—review and editing: MA, IL, DMC

Conflicts of Interest

DMC has served on speakers' bureaus for Corcept, AstraZeneca, GSK, AbbVie, and Pfizer; has served as a consultant or advisor for Genmab, Corcept, AstraZeneca, GSK, AbbVie, and Pfizer; and has received research funding from GSK and Merck.

Multimedia Appendix 1
Gynecologic oncology case questions, source materials, and physician grading rubrics used to evaluate Google and GPT-4 responses.
[DOCX File , 44 KB - cancer_v12i1e76471_app1.docx]

References

1. SEER*Stat databases: SEER November 2023 submission. National Cancer Institute. 2023. URL: https://www.seer.cancer.gov [accessed 2025-07-01]
2. Bhattad PB, Pacifico L. Empowering patients: promoting patient education and health literacy. Cureus 2022 Jul 27;14(7):e27336 [FREE Full text] [doi: 10.7759/cureus.27336] [Medline: 36043002]
3. Iannantuono GM, Bracken-Clarke D, Floudas CS, Roselli M, Gulley JL, Karzai F. Applications of large language models in cancer care: current evidence and future perspectives. Front Oncol 2023 Sep 4;13:1268915 [FREE Full text] [doi: 10.3389/fonc.2023.1268915] [Medline: 37731643]
4. Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. N Engl J Med 2023 Mar 30;388(13):1201-1208. [doi: 10.1056/NEJMra2302038] [Medline: 36988595]
5. Tozuka R, Johno H, Amakawa A, Sato J, Muto M, Seki S, et al. Application of NotebookLM, a large language model with retrieval-augmented generation, for lung cancer staging. Jpn J Radiol 2025 Apr;43(4):706-712. [doi: 10.1007/s11604-024-01705-1] [Medline: 39585559]
6. Hewitt KJ, Wiest IC, Carrero ZI, Bejan L, Millner TO, Brandner S, et al. Large language models as a diagnostic support tool in neuropathology. J Pathol Clin Res 2024 Nov;10(6):e70009 [FREE Full text] [doi: 10.1002/2056-4538.70009] [Medline: 39505569]
7. Anderson PB, Wanken ZJ, Perri JL, Columbo JA, Kang R, Spangler EL, et al. Patient information sources when facing repair of abdominal aortic aneurysm. J Vasc Surg 2020 Feb;71(2):497-504 [FREE Full text] [doi: 10.1016/j.jvs.2019.04.460] [Medline: 31353272]
8. Łaszkiewicz J, Krajewski W, Tomczak W, Chorbińska J, Nowak Ł, Chełmoński A, et al. Performance of ChatGPT in providing patient information about upper tract urothelial carcinoma. Contemp Oncol (Pozn) 2024;28(2):172-181 [FREE Full text] [doi: 10.5114/wo.2024.141567] [Medline: 39421706]

XSL•FO
RenderX

9. Gupta M, Gupta P, Ho C, Wood J, Guleria S, Virostko J. Can generative AI improve the readability of patient education materials at a radiology practice? Clin Radiol 2024 Nov;79(11):e1366-e1371. [doi: 10.1016/j.crad.2024.08.019] [Medline: 39266371]
10. Piazza D, Martorana F, Curaba A, Sambataro D, Valerio MR, Firenze A, et al. The consistency and quality of ChatGPT responses compared to clinical guidelines for ovarian cancer: a Delphi approach. Curr Oncol 2024 May 14;31(5):2796-2804 [FREE Full text] [doi: 10.3390/curroncol31050212] [Medline: 38785493]
11. Finch L, Broach V, Feinberg J, Al-Niaimi A, Abu-Rustum NR, Zhou Q, et al. ChatGPT compared to national guidelines for management of ovarian cancer: did ChatGPT get it right? - A Memorial Sloan Kettering Cancer Center team ovary study. Gynecol Oncol 2024 Oct;189:75-79. [doi: 10.1016/j.ygyno.2024.07.007] [Medline: 39042956]
12. CHART Collaborative. Reporting guideline for chatbot health advice studies: Chatbot Assessment Reporting Tool (CHART) statement. Ann Fam Med 2025 Sep 22;23(5):389-398 [FREE Full text] [doi: 10.1370/afm.250386] [Medline: 40750305]
13. ChatGPT. URL: https://chatgpt.com/ [accessed 2026-09-09]
14. Educational materials. Foundation for Women's Cancer. URL: https://foundationforwomenscancer.org/resources/educational-materials/ [accessed 2025-07-01]
15. Urman A, Makhortykh M. You are how (and where) you search? Comparative analysis of web search behavior using web tracking data. J Comput Soc Sci 2023 May 03;6(2):1-16 [FREE Full text] [doi: 10.1007/s42001-023-00208-9] [Medline: 37363807]
16. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
17. Chou HH, Chen YH, Lin CT, Chang HT, Wu AC, Tsai JL, et al. AI-driven patient support: evaluating the effectiveness of ChatGPT-4 in addressing queries about ovarian cancer compared with healthcare professionals in gynecologic oncology. Support Care Cancer 2025 Apr 01;33(4):337. [doi: 10.1007/s00520-025-09389-7] [Medline: 40167802]
18. Lambert R, Choo ZY, Gradwohl K, Schroedl L, Ruiz De Luzuriaga A. Assessing the application of large language models in generating dermatologic patient education materials according to reading level: qualitative study. JMIR Dermatol 2024 May 16;7:e55898 [FREE Full text] [doi: 10.2196/55898] [Medline: 38754096]
19. Spuur K, Currie G, Al-Mousa D, Pape R. Suitability of ChatGPT as a source of patient information for screening mammography. Health Promot Pract 2025 Jul;26(4):746-762 [FREE Full text] [doi: 10.1177/15248399241285060] [Medline: 39392690]
20. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
21. Karimov Z, Allahverdiyev I, Agayarov OY, Demir D, Almuradova E. ChatGPT vs UpToDate: comparative study of usefulness and reliability of Chatbot in common clinical presentations of otorhinolaryngology-head and neck surgery. Eur Arch Otorhinolaryngol 2024 Apr;281(4):2145-2151 [FREE Full text] [doi: 10.1007/s00405-023-08423-w] [Medline: 38217726]
22. Cavnar Helvaci B, Hepsen S, Candemir B, Boz O, Durantas H, Houssein M, et al. Assessing the accuracy and reliability of ChatGPT's medical responses about thyroid cancer. Int J Med Inform 2024 Nov;191:105593. [doi: 10.1016/j.ijmedinf.2024.105593] [Medline: 39151245]
23. Reicher L, Lutsker G, Michaan N, Grisaru D, Laskov I. Exploring the role of artificial intelligence, large language models: comparing patient-focused information and clinical decision support capabilities to the gynecologic oncology guidelines. Int J Gynaecol Obstet 2025 Feb;168(2):419-427. [doi: 10.1002/ijgo.15869] [Medline: 39161265]

Abbreviations

CHART: Chatbot Health Advice Reporting Tool
ICC: intraclass correlation coefficient
LLM: large language model

Edited by J Bender; submitted 23.Apr.2025; peer-reviewed by R Bramblet, S Matsuda, Á García-Barragán; comments to author 28.May.2025; revised version received 20.Aug.2026; accepted 25.Aug.2026; published 29.Sep.2026.

Please cite as:
Smick AH, Ayoola M, Rajpara R, Lazo I, Chase DM
Assessing the Accuracy, Completeness, and Reference Quality of GPT-4 and Google for Gynecologic Cancer Information: Comparative Quantitative Study
JMIR Cancer 2026;12:e76471
URL: https://cancer.jmir.org/2026/1/e76471
doi:10.2196/76471
PMID:

XSL•FO
RenderX

©Alexandra H Smick, Martins Ayoola, Rajeshree Rajpara, Isabel Lazo, Dana M Chase. Originally published in JMIR Cancer (https://cancer.jmir.org), 29.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Experiences of Informal Caregivers Facing Cancer While Caring for Someone With Dementia: Qualitative Interview Study

Elizabeth Travis[1], PhD; Mollie Louise Price[2], PhD; Millie Slessor[2], BSc; Jacqueline Crewe[3], MSc; Laura Ashley[2], PhD

[1]Leeds Institute of Health Sciences, University of Leeds, Worsley Building, Leeds, England, United Kingdom

[2]School of Humanities and Social Sciences, Leeds Beckett University, Leeds, England, United Kingdom

[3]University of Staffordshire, Stoke-on-Trent, England, United Kingdom

Corresponding Author:
Elizabeth Travis, PhD
Leeds Institute of Health Sciences, University of Leeds, Worsley Building, Leeds, England, United Kingdom

Abstract

Background: Informal caregivers play a fundamental role in the diagnosis, treatment, and care received by their family members. Caring for someone with dementia is particularly burdensome and can have negative effects on caregivers' health. Despite caregivers being widely recognized as vulnerable, the psychosocial needs of caregivers who are managing cancer themselves remain unknown.

Objective: This study aims to investigate the experiences of UK family caregivers who have received a cancer diagnosis while caring for someone with dementia or memory problems.

Methods: Twenty-five UK informal family caregivers who had received a cancer diagnosis were recruited through convenience sampling to take part in in-depth semistructured interviews. The interviews explored participants' experiences of managing cancer alongside dementia caregiving, including the impact on treatment and caregiving, emotional challenges, support needs, and experiences of formal and informal support. Ethical approval was obtained from the Leeds Beckett University Research Ethics Committee, and all participants provided informed consent. Reflexive thematic analysis was conducted.

Results: Three key themes were identified: (1) caring responsibilities can complicate and take priority over cancer care, (2) caregiving can intensify cancer-related distress and anxieties, and (3) managing cancer while caring increases the need for support and coping resources. For dementia caregivers, having cancer can surface/heighten feelings of predeath grief and worries about future care planning for their care recipient. Dementia caregivers viewed multifaceted social support as vital for enabling access to and coping with cancer treatment.

Conclusions: Dementia caregivers who are also living with cancer continue to shoulder the burden of illness work on behalf of their care recipients while simultaneously managing their own cancer diagnosis and treatment. Caregiving responsibilities often take priority over cancer care, intensify the practical and emotional challenges of cancer, and increase the need for and reliance on support resources. These findings highlight the need for person-centered adjustments to cancer care, dementia-friendly health care environments, and tailored psycho-oncology support that address the practical and emotional needs of dementia caregivers.

(JMIR Cancer 2026;12:e93688) doi:10.2196/93688

KEYWORDS

caregivers; older people; dementia; cancer; multimorbidity; qualitative research

Introduction

An informal caregiver is a person who provides unpaid care to a family member or friend with a health condition or illness [1]. In the United Kingdom, 5.8 million people identify as informal caregivers [2], and according to Carers UK, the number of unpaid caregivers in the United Kingdom is much higher than reported because many people do not recognize themselves as a caregiver [3]. Informal caregivers play a fundamental role in the diagnosis, treatment, and care received by their family members. Many informal caregivers provide care for individuals living with multiple long-term health conditions, which can increase caregiving demands and burden [4,5]. They provide physical, emotional, social, and financial support on a daily basis to care recipients and are often responsible for aiding communication of care recipients' medical needs with health care professionals [4,6]. Recent research has highlighted the support needs of caregivers of people with cancer and dementia, helping to inform and improve interventions and services provided to them [4]. A scoping review of 34 studies highlighted

XSL•FO
RenderX

common practical difficulties faced by caregivers of someone with multimorbidity, including having to manage multiple health care teams, appointments, and medications [7]. Notably, caregivers of people with dementia have reported practical difficulties with attending medical appointments due to "dementia-unfriendly" hospital settings [8].

Informal caregivers are 16% more likely than noncaregivers to live with 2 or more long-term health conditions [1], and there is a growing amount of evidence that unpaid care should be considered a social determinant of health [1]. Caring for someone with dementia is particularly burdensome (eg, it can involve constant supervision, emotional stress, and physical exhaustion), leading to known negative effects on caregivers' health [9-11]. Despite caregivers being widely recognized as vulnerable, those caregivers who are managing cancer themselves remain an underresearched group [12]. To our knowledge, no research has been conducted to investigate the experiences, difficulties, and support needs of informal caregivers who receive a cancer diagnosis and treatment themselves. This qualitative study therefore aimed to investigate the experiences of UK family caregivers who have received a cancer diagnosis while caring for someone with dementia or memory problems. The study aimed to specifically identify the difficulties and issues faced by patients with cancer and existing gaps in psychosocial support. Any cited recommendations for interventions, policy, and practice will also be reported and considered.

Methods

Participants

Participants were people with experience of simultaneously receiving a cancer diagnosis and caring for someone with dementia. Current or former caregivers were eligible if they had, within the last 5 years, provided unpaid informal care to a family member living with dementia or memory problems; had during this time also been diagnosed with cancer; were aged 18 years or older; were living in the United Kingdom; and were English literate. People who had a relative with dementia who had recently passed away (in the last 3 mo) were advised not to take part. None of the participants dropped out or withdrew from the study. The reporting of the current study adheres to the COREQ (Consolidated Criteria for Reporting Qualitative) guidelines [13].

Sampling and Recruitment

Participants were recruited through convenience sampling using community-based advertisements via relevant charities and support groups across the United Kingdom and the Join Dementia Research platform hosted by the National Institute for Health and Care Research (NIHR). Interested caregivers contacted the research team for more information. For those who wished to proceed, a discussion via email or telephone took place to allow the researcher to check the participant's eligibility.

The Research Team

The research team comprised four nonclinical women researchers and a women dementia support advisor/research coordinator working for a local dementia charity. The team collectively had considerable experience and expertise in applied psychosocial health research, including undertaking qualitative interview studies with informal caregivers and with people affected by cancer, dementia, and the cooccurrence of these conditions. The study was prompted by concerns raised by the dementia charity and National Health Service (NHS) oncology staff, in informal conversations with a member of the research team, regarding the challenges faced by dementia caregivers undergoing cancer treatment. Relevant personal and professional experiences within the research team included family experiences of dementia and cancer, alongside research expertise in dementia support and behavioral oncology. Consistent with reflexive thematic analysis (RTA), researchers reflected on how their experiences and perspectives may influence data interpretation through regular team discussions during coding and theme development.

Data Collection

Participants took part in one individual, audio-recorded, semistructured interview undertaken by ET via telephone or video call, depending on participant preference. Interviews explored participants' experiences of caring for someone with dementia alongside being diagnosed with cancer themselves. The interview topic guide (Multimedia Appendix 1) was informed by previous research in this area, including by our team (LA and MLP). The interview topic guide was reviewed and edited by the research team. The first interview was treated as a pilot, and no substantive changes to the topic guide were required. Interviews were conducted flexibly, allowing participants to discuss issues that were most important to them. Sensitive topics, such as future care planning, end-of-life concerns, and experiences of grief and loss, were raised by participants themselves rather than through direct questioning. All interviews were conducted by ET, an experienced qualitative researcher with expertise in psychosocial applied health research. The interviewer encouraged participants to discuss experiences that were important to them, and if participants became upset, they were reminded that they could pause, take a break, or continue the interview at a later time. The interview approach remained consistent throughout the study. Participants first provided sociodemographic information (eg, gender and age) and shared their experiences of caring for someone with dementia and how they came to find out they had cancer. Participants were recruited from 7 of the 9 regions of England and 1 region of Scotland, providing broad geographical representation within the sample. Interviews then focused in-depth on caregivers' experiences, difficulties, and needs around simultaneously caring and facing cancer and on experienced or potential sources of support or ways of helping to address caregivers' challenges and needs at this time.

Data Analysis

Data were analysed using RTA [14,15]. Analysis was an inductive, iterative, collaborative process involving both independent review and team discussion of the data, with decisions made collaboratively by consensus. Analysis was begun concurrently with data collection, allowing analytical insights to inform subsequent interviews and enabling the

XSL•FO
RenderX

research team to assess whether the data captured a sufficiently broad and rich range of experiences to address the study aims. Recruitment ended when the team agreed that the dataset provided adequate depth and diversity of perspectives relevant to the research aim. We adopted an experiential approach to RTA, informed by a hermeneutic phenomenological approach, as we aimed to understand how participants experienced and made sense of receiving a cancer diagnosis while caring for someone with dementia, including their support needs [16]. This approach complemented RTA by focusing on participants' perspectives and meanings, while recognizing that theme development involved interpretation by the research team. Transcripts were analyzed following the six stages of RTA as follows: (1) initial interview observations were recorded in MS Word by the interviewer following each interview. Interviews were then transcribed and anonymized by ET (P1-P19) and MS (P20-P25). (2) Codes were generated inductively and independently by ET, MLP, and JC to develop an analytical coding framework, which was subsequently applied to all transcripts by ET. In developing codes, MLP and JC each coded a different sample of 6 transcripts, which were also all coded by ET. (3) Similar codes were grouped and candidate themes and subthemes generated by ET. (4) ET, MLP, and MS met to review and refine candidate themes and subthemes and select indicative quotes. (5) ET and LA met to further review, restructure, and define themes and subthemes and determine the illustrating quotes. (6) Final theme and subtheme titles, descriptions, and illustrative quotes were reviewed and agreed upon by team consensus (MLP, MS, and LA). All stages of coding and data analysis were carried out in MS Excel. Participant validation (also known as member checking) was not undertaken in this study. Instead, the credibility of the analysis was supported through iterative discussion and interpretation of the data within the research team.

Ethical Considerations

The study received ethical approval from the Leeds Beckett University Psychology Local Research Ethics Committee on March 14, 2024. All participants received written information about the study and provided informed consent prior to the interview.

Results

Participants and Data Collection

Twenty-five informal family caregivers participated in the study. Sample characteristics are summarized in Table 1. No participants withdrew from the study. Interviews were conducted by telephone (n=15) or online video call (n=10) between March 25, 2024, and January 15, 2025, and ranged from 24 to 78 (mean 47.72, SD 12.92) minutes in duration. Variation in interview duration is likely a reflection of individual differences in participants' experiences, the amount of detail they wished to share, and their characteristic verbal style; however, factors contributing to interview length were not assessed.

XSL•FO
RenderX

Table . Sample characteristics.

Variables	Count (N=25)
Age (y), mean (SD; range)	68.32 (10.95; 46-81)
Gender, n (%)	
Women	14 (56)
Men	11 (44)
Ethnicity, n (%)	
British–White	23 (92)
British–South Asian	1 (4)
Norwegian–White	1 (4)
Relationship to care recipient, n (%)	
Spouse/partner	12 (48)
Adult child	12 (48)
Child-in-law	1 (4)
Socioeconomic status: index of multiple deprivation	
Quintiles	
1	2 (8)
2	1 (4)
3	4 (16)
4	6 (24)
5	7 (28)
Unknown (post code not specified)	5 (20)
Residing with care recipient, n (%)	
Yes	12 (48)
No	13 (52)
Current employment, n (%)	
Employed	5 (20)
Unemployed	2 (8)
Retired	18 (72)
Care recipient dementia diagnosis, n (%)	
Vascular dementia	2 (8)
Alzheimer disease	14 (56)
Frontotemporal dementia	3 (12)
Mixed dementia	3 (12)
Parkinson disease dementia	1 (4)
Memory problems (awaiting formal diagnosis)	1 (4)
Not specified	1 (4)
Caregiver cancer diagnosis, n (%)	
Bowel	2 (8)
Bladder	1 (4)
Breast	11 (44)
Skin	1 (4)
Kidney	1 (4)
Prostate	5 (20)

XSL•FO
RenderX

Variables	Count (N=25)
Thrombocythemia	1 (4)
Leukemia	2 (8)
Multiple myeloma	1 (4)

Key Findings

Three themes, each with interrelated subthemes, were developed from the data to capture the experiences of and support received by informal caregivers while having cancer. The findings highlight how caregiver identity and role shaped and intensified both the practical and emotional challenges of cancer, giving rise to multifaceted support needs. Caring responsibilities were central to how participants experienced cancer, often taking precedence over their own health needs and influencing engagement with their treatment. Competing caregiving demands produced both practical disruptions to engaging in cancer care and intensified emotional consequences of facing cancer, including grief and loss, and anxieties about future care for the person living with dementia. In response, participants described a need to draw significantly on practical, financial, and psychological resources to effectively manage the demands of caregiving and cancer.

Theme 1: Caring Responsibilities Can Complicate and Take Priority Over Cancer Care

Overview

Caregivers spoke about their cancer as a competing priority to their caring responsibilities. Caring responsibilities frequently acted as the priority around which cancer was managed, shaping engagement with treatment and health care services. Interviews highlighted how caregiver burden and related responsibilities can complicate and negatively impact care accessibility and experiences across the cancer continuum.

Competing Caring Demands Make Cancer Care More Difficult to Manage

Caregivers cited various difficulties experienced in managing the demands of their dual role as a caregiver and people with cancer. These complications made undergoing cancer treatment difficult and stressful for some caregivers. Logistical complications included the need for caregivers to coordinate attendance at multiple appointments and treatments, the reality of having little to no choice but to prioritize the person with dementia over their own treatment, and the challenges arising when care recipients walked away from the waiting area during their own appointments or treatment.

For many caregivers, managing their cancer treatment required considerable planning and coordination alongside upholding existing caregiving responsibilities.

> *I've moved appointments and blood tests, or we've done it all on the same day. And you know, the other week we visited three hospitals in the region of England, two for my mum and one for me.* [P7, age:56, daughtercaregiver]

Challenges were not limited to scheduling appointments and could also create concerns about the safety and supervision of the person living with dementia.

> *She sat in the waiting room. Unfortunately, it was an hour long before I could come out. When I came out, she wasn't there. So, I said to the receptionist "my wife?," "Oh yes," she said. "She went to the toilet some time ago and she hasn't come back," so I was up and down the hospital searching for her, so I never took her again.* [P4, age:79, husband-caregiver]

Prioritizing Caring Responsibilities Can Affect Engagement With Cancer Care

From cancer diagnosis to outcome, the role of being a caregiver continued to be at the forefront of caregivers' minds, viewed as the greatest need before their own cancer. Examples cited demonstrate how being a caregiver with cancer can have a detrimental impact upon the patient's cancer trajectory in terms of help-seeking, decision-making, ability to give their full attention when attending medical appointments, and the need to reset and recover from their own treatment. As expected, the caregiver's burden and challenges were thought to be greater when the care recipient's dementia was at an advanced stage rather than when it was at an earlier stage.

For some participants, caregiving responsibilities took priority over their own health needs and delayed help-seeking for cancer symptoms.

> *I would say that I put off going about the lump when it came back last year…my dad was in hospital at the time, and so we were up and down and his circumstances of him going into hospital.* [P10, age:61, daughter-caregiver]

> *They were coming to appointments with me, and I found that I was more concerned about them and their reactions than I was about myself. I couldn't listen. I was looking at their faces and you know, they had all sorts of questions.* [P2, age:73, wife-caregiver]

Caregiving responsibilities also influenced how some participants approached their treatment decisions, with a preference for options that would minimize disruption to their caregiving role.

> *Radiotherapy offered a chance that it might not knock me out so much and all the rest of it, compared to the recovery time from an operation, so I could carry on caring.* [P6, age:80, husband-caregiver]

XSL•FO
RenderX

Theme 2: Caregiving Can Intensify Cancer-Related Distress and Anxieties

Overview

Beyond its practical impact on cancer care, being a dementia caregiver also shaped and intensified the emotional sequelae of cancer for caregivers. Caring for a person with dementia while facing cancer often intensified caregivers' feelings of loss, uncertainty, and vulnerability, leading to our second theme: caregiving can intensify cancer-related distress and anxieties. Being a caregiver for someone with dementia often meant that caregivers experienced additional or heightened cancer-related distress and anxieties tied to the losses and responsibilities of their caregiving role. Caregivers' realization of this emotional distress often occurred and was only dealt with at a later stage in their cancer journey. Caregivers with a terminal diagnosis worried over practical arrangements and the emotional burden of who would assume the challenging dementia care responsibilities after their death.

Cancer Triggers Feelings of Loss and Grief Toward the Person With Dementia

Being diagnosed with cancer and undergoing treatment surfaced feelings of loss in caregivers. This was attributed to their predeath grief toward the care recipient, who, due to their dementia, could no longer provide them with the emotional support they needed. Caregivers explained that they were unable to turn to their care recipient in the way they once could, before the care recipient's dementia, for emotional support. For some caregivers, this unmet need for emotional support to help them deal with their cancer and feelings of predeath grief was only realized retrospectively and acted upon at a later stage or following cancer treatment.

Adult-child caregivers in particular described a sense of loss arising from no longer being able to rely on their parent, the care recipient, for emotional support in the way they had before dementia.

> *I would have probably liked to talk about it to my mum and dad. You always go to your mum don't you. I missed that.* [P14, age:66, daughter-caregiver]

> *I said "Dad, I've just got a diagnosis of cancer", and his reaction to that, imagine this...your son tells you, "I've got cancer, it's not a little cancer, it's a serious cancer". And there was a pause. And he said, "oh, I'm sorry to hear that. Is there anything we can do?" that was literally what he said. It was so unemotional and uncaring.* [P9, age:59, son-caregiver]

For some, the emotional impact of having cancer was only recognized later, once the immediate demands of caregiving and their cancer treatment had passed.

> *I don't think I actually acknowledged...making the decision to not tell my mum was one thing, which was upsetting, but then I had to sort of put that in the box because I've got this to deal with now [having cancer]. It's only really in the last year or so that's sort of gone, do you know that was really actually quite bad. And yeah, you know what I've missed out on and how I actually really needed my mum. But at the time I had to put that into a box because I can't deal with that now because I'd got this [having cancer] to deal with. What was happening with mum was one thing and what was happening with me, was another thing. And I didn't connect them at the time, because it was too upsetting to connect them probably. So I never spoke to anybody about it because it's not something I was thinking about.* [P18, age:49, daughter-caregiver]

Caring Heightens Cancer-Related Anxieties and Fears About Dying and Future Planning

For those caregivers with an incurable cancer diagnosis, uncertainty and a lack of control regarding timings of their own and their care recipient's future situation were a cause of great distress. Caregivers were worried about what would happen to the person they cared for when they died. Their concerns encompassed both practical matters, such as funeral arrangements, and emotional anxieties, particularly around who would take on the responsibility for caregiving, a role that is often exhausting and burdensome when supporting someone with dementia.

> *I fear for the future. I fear for what's going to happen. In terms of caring for me or caring for <Person living with dementia name> because he's not going to want to go into a home. He's got quite an awkward streak in him, always has had. I think that's probably what I fear, and for my own health I'd, you know, [Crying] I don't know what the future holds. I've felt very ill before now and a year ago and you know, if I was like that, I wouldn't be able to look after him because it is very tiring.* [P1, age:76, wife-caregiver]

> *It's terrible when you're wishing your mum to die before you do. Because I don't want to cause her the trauma of, because I've got a metastatic diagnosis, you know, gosh I'm lucky to be here now. And that is a really terrible thing, that is one of the worries.* [P7, age:56, daughter-caregiver]

> *My mom and now my sister, who's going to look after them, if I'm not there, that's more the fear now, because I know that my kids will be okay, and they'll take care of their kids. But Alzheimer's is dreadful, and we've had dreadful experiences with that.* [P16, age:69, daughter-caregiver]

Managing the demanding roles of dementia caregiver and patient with cancer in tandem, along with the resultant practical and emotional challenges, was often cited as overwhelming. Consequently, participants reported a need for greater practical, emotional, and financial support, which forms our third theme.

Theme 3: Managing Cancer While Caring Increases the Need for Support and Coping Resources

Overview

Many caregivers reported a lack of awareness of available dementia support services, limited availability of those services, and inconsistent health and social care for the person with

XSL•FO
RenderX

dementia. To overcome systemic issues, caregivers often cited instrumental, financial, and emotional avenues of support they had depended upon to enable them to continue to care for their care recipient while also attending medical appointments and treatment of their cancer.

Reliance on Practical and Emotional Support From Family and Friends

Reliance on support networks was frequently cited as a response to the competing demands described in theme 1 and the emotional burden described in theme 2. Family members and friends often provided practical care that freed up participants to attend their cancer appointments, while also offering reassurance and emotional support for their cancer-related worries. All caregivers cited their reliance and sole dependence on friends and family to be able to attend medical appointments and treatment related to their cancer, particularly when caregivers had received last-minute appointments and were unable to arrange professional care at short notice. This instrumental support provided by family and friends also included respite care, cited by caregivers who at times needed a break to recharge physically and mentally. Caregivers also talked about the emotional support they had received from friends and family.

> *if somebody else had looked after my dad, instead of us all the time. And given us a break, because then none of it would have been so difficult.* [P10, age:61, daughter-caregiver]

> *It didn't really occur to me that there might be any facilities for it. So I didn't, you know, if you're not aware of something being available, possibly you don't even think to look for it. If you're not told, that would help or not.* [P11, age:79, wife-caregiver]

These accounts suggest that limited awareness of available support services may increase caregivers' reliance on family and friends.

> *I had a bit of a meltdown. I really couldn't see how I was going to cope. And I said to my daughter, could you have him? And so, she had him to stay for two nights, one day and two nights, and during that time I managed, I got my brain sorted really.* [P1, age:76, wife-caregiver]

Support from family members often extended beyond practical help and included emotional support during cancer treatment and recovery.

> *It's my wife who has taken the real hit in terms of emotionally supporting me, yet also taking on the burden, as it were, of supporting my mother* [P25, age:67, son-caregiver]

> *They [local council] didn't say they weren't interested, but they just kept saying, oh, we're fully booked. I don't know that we can do that on the day you want, and because I wasn't getting a lot of notice to go into hospital, so I was sort of saying to them, you know, in about three or four days time. So in the end, the family just said, stop messing about with it, although it sounds a good procedure.* [P4, age:79, husband-caregiver]

Financial Security Helps Caregivers Manage Cancer and Caring Responsibilities

Financial resources appeared to help mitigate the impact of challenges described in themes 1 and 2. Participants with access to savings, pensions, or privately funded care were often better placed to negotiate competing caregiving and cancer demands, reducing some of the practical tensions between caring and their cancer treatment. Given the current issues caregivers cited with a lack of health and social care provided, many caregivers felt privileged and fortunate that they had the option to fund private care themselves through their own or family personal finances. For some, personal finances had enabled them to reduce their working hours or take early retirement. For caregivers without personal finances to draw upon, this was viewed by caregivers interviewed as an unimaginable problem to cope with and manage.

> *Just to make sure they're safe and you don't have to worry about it. Do you know what I mean? Everything that you're going through without having to worry about everybody else. That would be the ideal world. But there's no budgets anywhere, is there for that.* [P13, age:53, daughter-caregiver]

> *Especially for people who haven't got the money, yeah, it must be so difficult. I just feel so completely privileged.* [P5, age:77, wife-caregiver.]

Financial security enabled some caregivers to reduce the demands of paid work, preserving more of their personal resources for managing the demands of caregiving and their cancer treatment.

> *Since I had cancer last year, I have only gone back to work, I halved my shifts, so that is a considerable difference and I am really aware of that but, we haven't, you know, we've managed. No, no, we're lucky.* [P10, age:61, daughter-caregiver]

> *We've both got private pensions, and we got somebody round and realized that we could take a bit of that every month and not have to, to carry on working if we didn't need to, for the short term at least.* [P19, age:59, daughter-caregiver]

Personal Resilience Supports Coping With Cancer and Caregiving Demands

Caregivers' accounts highlighted emotional resilience as key to how they actively coped with the practical and emotional challenges described in themes 1 and 2. Many caregivers spoke in ways that suggested they had a positive mindset about how they perceived their situation and challenges they faced. To help with coping, caregivers appeared to engage in benefit finding, somewhat indicative of their underlying resilience, which not only helped them deal with challenges but also appeared to strengthen their resilience further. For example, some caregivers viewed an early cancer diagnosis as a blessing, given that it aligned with their perception of the illness as being less severe. Being a caregiver was viewed as a positive distraction, a role

XSL•FO
RenderX

that gave caregivers a sense of purpose and provided a good distraction from worrying too much about their own cancer. One caregiver explained how she tries to focus on today, practicing mindfulness and resilience as a way to overcome and cope with any worries about what tomorrow will bring.

> *I think having my husband to look after was a distraction from focusing too much on the cancer and kind of getting panicky, so in a way, it just gave me something else to focus on as well. I think if it had just been me and my cancer, I would have focused more on that and then that would have been maybe harder to keep any sort of anxiety out of the way.* [P5, age:77, wife-caregiver]

Caregivers often described actively focusing on the positive aspects of their situation as a way of coping with the challenges they faced.

> *The cancer was caught at a very early stage. So, fortunately the cancer never really figured it was, just. Oh, look, I've got to do X, it was no more a problem than going to the dentist.* [P3, age:76, husband-caregiver]

> *Having cancer is not really an issue in the whole thing of caring for her because it's a non-event almost.* [P22, age 79, husband-caregiver]

Other caregivers described focusing on the present rather than worrying about the future as an important coping strategy.

> *It is a progressive disease, I do know what the future holds, but, I don't think about the future, we live for today. you know, I try to live as normal life as possible. Think about today, let tomorrow take care of itself, really. That's the way I've coped really, do I have wobbles, yes I do.* [P15, age:80, husband-caregiver]

Discussion

Key Findings

To our knowledge, this is the first qualitative study to explore, in the United Kingdom and internationally, the experiences of informal caregivers who have received a cancer diagnosis while caring for someone with dementia or memory problems. The study aimed to specifically identify the difficulties and issues faced by patients with cancer and existing gaps in psychosocial support. The findings corroborate and extend the work by Price et al [4,17] on the psychosocial support needs of caregivers of people with comorbid dementia and cancer, while highlighting the distinct experiences of individuals who are themselves living with cancer while also caring for a person with dementia. Study findings were developed into three key themes: (1) caring responsibilities can complicate and take priority over cancer care, (2) caregiving can intensify cancer-related distress and anxieties, and (3) managing cancer while caring increases the need for support and coping resources. Furthermore, this study is the first to present evidence that caregivers living with cancer themselves continue to shoulder the burden of illness work on behalf of their care recipients. Illness work, as described by Corbin and Strauss [18], refers to tasks such as taking medications, monitoring symptoms, attending medical appointments, and coordinating with health care providers. Price et al [4] emphasized the relentless caring regime involved in being a caregiver, and our study findings demonstrate that the caregiver role prevails, even to the detriment of the patient's cancer trajectory. Caregivers cited examples of how their caregiving role had influenced their decisions to opt for less invasive treatments with shorter recovery times, delayed their help-seeking for cancer symptoms, required them to reschedule appointments, and left them with limited time to recover from their own treatment. These findings extend previous research on illness work by demonstrating that caregiving responsibilities often continue when caregivers themselves are managing a serious health condition such as cancer. Consistent with wider caregiving literature, caregivers commonly continue providing care despite chronic illness, pain, and other health challenges of their own, often prioritizing the needs of the care recipient over their own health and well-being [19-22]. Participants in this study similarly placed the needs of the person with dementia ahead of their own, suggesting that the caregiver role remained of greatest importance despite cancer diagnosis and treatment.

This study underscores the unique emotional challenges faced by caregivers for someone with dementia who simultaneously need support for their own cancer diagnosis and treatment. Among the adult-child caregivers in our sample, participants described their parent as having previously been their significant source of emotional support in times of need. However, due to the progression of the care recipient's dementia, they could no longer rely on them for emotional support. There is growing interest in assessing and addressing predeath or anticipatory grief experiences among dementia caregivers, with recognition that this is an underresearched issue in which people can feel unsupported [23,24]. The loss of emotional support associated with dementia progression may be particularly challenging for caregivers who are simultaneously coping with their own cancer diagnosis and treatment. These experiences suggest that anticipatory grief may be compounded by the emotional and practical demands of cancer, highlighting a need for psychosocial support that addresses both cancer-related distress and losses associated with dementia progression. Further work is required to better understand these intersecting experiences and support needs. Findings highlight an important need to provide targeted psychosocial support to caregivers with cancer to help them emotionally cope with and process additional levels of loss experienced due to predeath grief.

Many caregivers relied on the support from family and friends, their own financial security, and their own emotional resilience to cope. Ji et al [25] found significant correlations between psychological resilience, caregiving burden, social support, and quality of life among primary caregivers of patients with cancer. Our findings suggest that access to practical, emotional, and social support may be especially important for people balancing the dual roles of being an informal caregiver and a patient with cancer.

Our findings have implications for digital health. Caregivers reported limited awareness of available services, difficulties navigating support, and challenges balancing their cancer care with caregiving responsibilities. These needs may be addressed

XSL•FO
RenderX

in part through digital solutions such as integrated patient and caregiver portals, digital signposting resources, telehealth follow-up, and online psycho-oncology support. Such approaches could help improve access to information, care coordination, and emotional support while reducing the practical burdens associated with hospital visits [26,27]. Online peer support may also be valuable, particularly for caregivers with limited support from family and friends. For example, the cancer-specific section of the Alzheimer's Society online forum, codeveloped by Ashley et al [8,28], provides opportunities for shared understanding, emotional support, and the exchange of practical advice among people navigating both cancer and dementia caregiving. Future research could explore the acceptability and effectiveness of digital interventions for informal caregivers managing both cancer themselves and dementia-related caregiving demands.

Benefit finding is a psychological concept whereby someone searches for or identifies positive outcomes or personal growth when faced with stressful life experiences [29]. To help with coping, caregivers in this study indicated involvement in benefit finding, which not only helps them deal with challenges faced but also strengthens their resilience further. A recent review of 21 articles by Chiba et al [30] highlighted indicators of benefit finding among family caregivers of individuals with mental illness, including positive personal traits, good relationships with patients, and social support. Benefit finding has also been identified among caregivers of people with advanced cancer, where caregivers reported personal growth, stronger relationships, perceived social support, and a greater sense of meaning in life [31]. While benefit finding has been reported in other caregiver populations, our findings suggest that caregiving may also provide a sense of purpose and distraction from cancer-related concerns for some individuals. Caregivers in our study described engaging in positive distraction, focusing on their role as a caregiver, as a coping strategy, shifting their attention and concern away from their own cancer. Dobignes et al [32] investigated the use of an online positive psychology intervention to support caregivers of people with Alzheimer disease. Their findings reported that caregivers of people with earlier stage Alzheimer disease benefited the most from the intervention [28]. Further work in this area is warranted to explore and test interventions that support coping strategies for caregivers more widely. Several implications arising from our findings align with recommendations previously proposed by Ashley et al [8,28] to improve cancer care for people affected by dementia. These include cancer services identifying patients who also care for someone with dementia and considering reasonable adjustments that help them balance treatment and caring responsibilities. Flexible appointments, reduced hospital visits, dementia-friendly environments, and practical support with travel or attendance could also help minimize barriers to cancer care. It is also important to recognize the emotional impact of managing cancer while providing care. Regular discussions about well-being, access to caregivers' support services, and tailored psychological support can help address anxiety, future care concerns, and feelings of loss or grief, many of which may be reduced through appropriate adjustments and support [8,28].

Our findings highlight the challenges of being both an informal dementia caregiver and a patient with cancer at the same time. Participants described prioritizing the needs of the person with dementia over their own health and cancer treatment, coping with the emotional impact of dementia progression intensified in the context of coping with cancer, and needing and relying on different sources of support to manage both roles. These findings suggest that the experiences of informal caregivers with cancer are not fully reflected in the existing cancer or dementia caregiving literature. Our study helps to address this gap by providing new insight into the challenges and support needs of people who are simultaneously managing a cancer diagnosis while continuing to care for a person living with dementia.

Key Strengths and Limitations

Study strengths include recruitment of a sizeable sample with diversity across several study-relevant personal characteristics (eg, gender, age, cancer type, caregiver relationship to person with dementia, and socioeconomic status), as well as in-depth, rich data collection and collaborative, iterative data analysis involving multiple members of the research team. Limitations include the recruited caregivers who were predominantly retired and White British. Although this partly reflects the higher incidence of cancer in older adults [33] and the larger proportion of White British people affected by dementia in the United Kingdom, the Alzheimer's Society reports that just 3% of people living with dementia are from Black, Asian, and minority ethnic communities [34].

Conclusions

Our study demonstrates that for dementia caregivers facing cancer, caregiving responsibilities can compete with, complicate, and negatively impact engagement in cancer treatment and care. Our findings underscore the unique emotional challenges faced by caregivers for someone with dementia, who simultaneously need support for their own cancer diagnosis and treatment. For dementia caregivers, having cancer can surface/heighten feelings of predeath grief and worries about future care planning for their care recipient. Dementia caregivers viewed multifaceted social support as vital for enabling access to and coping with cancer treatment. This investigation provides initial evidence that caregivers living with cancer themselves continue to shoulder the burden of illness work on behalf of their care recipients. For dementia caregivers, accessibility and experiences of cancer care may be optimized by person-centered adjustments to care planning and scheduling and dementia-friendly hospital environments. Dementia caregivers may benefit from tailored psycho-oncology support encompassing caregiver-role concerns and needs (eg, predeath grief for their care recipient). Digital health interventions may also help address needs around information provision, service navigation, care coordination, and psychosocial support for dementia caregivers with cancer. The findings suggest that the experiences of dementia caregivers with cancer are not fully reflected in the existing cancer or dementia caregiving literature. By focusing on the intersection of these roles, our study provides new insight into the challenges and support needs of this overlooked group. Future research should evaluate interventions

XSL•FO
RenderX

designed to support individuals managing both dementia caregiving responsibilities and a personal cancer diagnosis and treatment.

Acknowledgments

Generative AI tools were not used at any stage in the preparation of this manuscript.

Funding

This work was supported by internal funding from the School of Humanities and Social Sciences, Leeds Beckett University, UK.

Data Availability

The data for this study are available from the corresponding author upon reasonable request. The data are not publicly available because they contain information that could compromise the privacy of research participants.

Conflicts of Interest

JC works for a dementia charity, which supports people living with dementia and their caregivers. All other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Interview topic guide for exploring the experiences of people facing cancer while caring for someone with dementia or memory problems.
[DOCX File, 22 KB - cancer_v12i1e93688_app1.docx]

References

1. Caring as a social determinant of health: findings from a rapid review of reviews and analysis of the GP Patient Survey. : Public Health England; 2021 URL: https://assets.publishing.service.gov.uk/media/60547266d3bf7f2f14694965/Caring_as_a_social_determinant_report.pdf [accessed 2025-07-03]
2. Unpaid care by age, sex and deprivation, England and Wales: census 2021. Office for National Statistics. 2023 Jan 19. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/socialcare/articles/unpaidcarebyagesexanddeprivationenglandandwales/census2021 [accessed 2025-07-03]
3. Key facts and figures about caring. Carers UK. URL: https://www.carersuk.org/policy-and-research/key-facts-and-figures/ [accessed 2025-07-03]
4. Price ML, Surr CA, Gough B, Ashley L. Understanding the experiences and psychosocial support needs of caregivers of people with comorbid dementia and cancer. Psychol Health 2024 Oct;39(10):1428-1450. [doi: 10.1080/08870446.2022.2157030] [Medline: 36536600]
5. Surr C, Griffiths AW, Kelley R, et al. Navigating cancer treatment and care when living with comorbid dementia: an ethnographic study. Support Care Cancer 2021 May;29(5):2571-2579. [doi: 10.1007/s00520-020-05735-z] [Medline: 32955656]
6. Rowland C, Hanratty B, Pilling M, van den Berg B, Grande G. The contributions of family care-givers at end of life: a national post-bereavement census survey of cancer carers' hours of care and expenditures. Palliat Med 2017 Apr;31(4):346-355. [doi: 10.1177/0269216317690479] [Medline: 28281406]
7. Price ML, Surr CA, Gough B, Ashley L. Experiences and support needs of informal caregivers of people with multimorbidity: a scoping literature review. Psychol Health 2020 Jan;35(1):36-69. [doi: 10.1080/08870446.2019.1626125] [Medline: 31321995]
8. Ashley L, Kelley R, Griffiths A, et al. Understanding and identifying ways to improve hospital-based cancer care and treatment for people with dementia: an ethnographic study. Age Ageing 2021 Jan 8;50(1):233-241. [doi: 10.1093/ageing/afaa210] [Medline: 33156901]
9. Allen AP, Curran EA, Duggan Á, et al. A systematic review of the psychobiological burden of informal caregiving for patients with dementia: focus on cognitive and biological markers of chronic stress. Neurosci Biobehav Rev 2017 Feb;73:123-164. [doi: 10.1016/j.neubiorev.2016.12.006] [Medline: 27986469]
10. Allen AP, Buckley MM, Cryan JF, et al. Informal caregiving for dementia patients: the contribution of patient characteristics and behaviours to caregiver burden. Age Ageing 2019 Dec 1;49(1):52-56. [doi: 10.1093/ageing/afz128] [Medline: 31755532]
11. Davies LE, Spiers GF, Sinclair DR, Kingston A, Hanratty B. Characteristics of older unpaid carers in England: a study of social patterning from the English Longitudinal Study of Ageing. Age Ageing 2024 Mar 1;53(3):afae049. [doi: 10.1093/ageing/afae049] [Medline: 38497238]

XSL•FO
RenderX

12. Furmenti MF, Bertarelli G, Ferrè F. Person-centred care in oncological home services: a scoping review of patients' and caregivers' experience and needs. BMC Health Serv Res 2025 Feb 11;25(1):232. [doi: 10.1186/s12913-024-12058-w] [Medline: 39934798]
13. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
14. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
15. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
16. Laverty SM. Hermeneutic phenomenology and phenomenology: a comparison of historical and methodological considerations. Int J Qual Methods 2003 Sep;2(3):21-35. [doi: 10.1177/160940690300200303]
17. Price ML, Surr C, Gough B, Howe D, Ashley L. The provision of social support in an online support forum for caregivers of people with comorbid dementia and cancer: content analysis study. JMIR Cancer 2025 Sep 29;11:e72217. [doi: 10.2196/72217] [Medline: 41022373]
18. Corbin J, Strauss A. Managing chronic illness at home: three lines of work. Qual Sociol 1985;8(3):224-247. [doi: 10.1007/BF00989485]
19. Schmaderer M, Struwe L, Pozehl B, Loecker C, Zimmerman L. Health status and burden in caregivers of patients with multimorbidity. Gerontol Geriatr Med 2020;6:2333721420959228. [doi: 10.1177/2333721420959228] [Medline: 35047651]
20. Duplantier SC, Williamson FA. Barriers and facilitators of health and well-being in informal caregivers of dementia patients: a qualitative study. Int J Environ Res Public Health 2023 Feb 28;20(5):4328. [doi: 10.3390/ijerph20054328] [Medline: 36901336]
21. Sabatini S, Martyr A, Hunt A, et al. Health conditions in spousal caregivers of people with dementia and their relationships with stress, caregiving experiences, and social networks: longitudinal findings from the IDEAL programme. BMC Geriatr 2024 Feb 19;24(1):171. [doi: 10.1186/s12877-024-04707-w] [Medline: 38373905]
22. Turner SG, Witzel DD, Garza S, Pillemer K, Reid MC. Health patterns of dementia caregivers with chronic pain: latent profile analysis of the PROMIS-29 measure. J Appl Gerontol 2026 Jun;45(6):1149-1159. [doi: 10.1177/07334648251360819] [Medline: 40738141]
23. Burton JK, Soiza RL, Quinn TJ, SIGN 168 guideline development group. Guideline summary: assessment, diagnosis, care and support for people with dementia and their carers [Scottish Intercollegiate Guidelines Network SIGN Guideline 168]. Age Ageing 2024 Jul 2;53(7):afae147. [doi: 10.1093/ageing/afae147] [Medline: 38970550]
24. Dehpour T, Koffman J. Assessment of anticipatory grief in informal caregivers of dependants with dementia: a systematic review. Aging Ment Health 2023 Jan;27(1):110-123. [doi: 10.1080/13607863.2022.2032599] [Medline: 35168426]
25. Ji Q, Zhang L, Ji P, et al. The relationship between psychological resilience and quality of life among primary caregivers of cancer patients: the mediating role of care burden and the moderating role of social support. Support Care Cancer 2025 Apr 2;33(4):343. [doi: 10.1007/s00520-025-09412-x] [Medline: 40172687]
26. Zhang Y, Flannery M, Zhang Z, et al. Digital health psychosocial intervention in adult patients with cancer and their families: systematic review and meta-analysis. JMIR Cancer 2024 Feb 5;10:e46116. [doi: 10.2196/46116] [Medline: 38315546]
27. Shaffer KM, Turner KL, Siwik C, et al. Digital health and telehealth in cancer care: a scoping review of reviews. Lancet Digit Health 2023 May;5(5):e316-e327. [doi: 10.1016/S2589-7500(23)00049-3] [Medline: 37100545]
28. Ashley L, Surr C, Kelley R, et al. Cancer care for people with dementia: literature overview and recommendations for practice and research. CA Cancer J Clin 2023;73(3):320-338. [doi: 10.3322/caac.21767] [Medline: 36512303]
29. Lechner SC. 63 benefit-finding. In: Snyder CR, Lopez SJ, Edwards LM, Marques SC, editors. The Oxford Handbook of Positive Psychology, 3rd edition: Oxford University Press; 2018:907-918. [doi: 10.1093/oxfordhb/9780199396511.013.53]
30. Chiba R, Miyamoto Y, Funakoshi A, et al. Benefit finding and growth among family caregivers of individuals with mental illness: a scoping review. J Fam Nurs 2025 Feb;31(1):16-30. [doi: 10.1177/10748407241309138] [Medline: 39840639]
31. Song Y, Wang M, Zhu M, et al. Benefit finding among family caregivers of patients with advanced cancer in a palliative treatment: a qualitative study. BMC Nurs 2024 Jun 11;23(1):397. [doi: 10.1186/s12912-024-02055-z] [Medline: 38862930]
32. Dobignies M, Larochette C, Andreotti E, et al. Experience with an online positive psychology intervention for caregivers of people with Alzheimer's disease: an interpretative phenomenological analysis. Int J Qual Stud Health Well-being 2025 Dec;20(1):2494349. [doi: 10.1080/17482631.2025.2494349] [Medline: 40292535]
33. All cancers combined incidence statistics. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/all-cancers-combined/incidence [accessed 2026-09-19]
34. Black, Asian and minority ethnic communities and dementia research. Alzheimer's Society. 2022 Jun 7. URL: https://www.alzheimers.org.uk/for-researchers/black-asian-and-minority-ethnic-communities-and-dementia-research [accessed 2026-09-19]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative

XSL•FO
RenderX

NHS: National Health Service
NIHR: National Institute for Health and Care Research
RTA: reflexive thematic analysis

Edited by M Balcarras; submitted 17.Feb.2026; peer-reviewed by C Richardson, S Choemprayong; revised version received 19.Aug.2026; accepted 08.Sep.2026; published 05.Oct.2026.

Please cite as:
Travis E, Price ML, Slessor M, Crewe J, Ashley L
Experiences of Informal Caregivers Facing Cancer While Caring for Someone With Dementia: Qualitative Interview Study
JMIR Cancer 2026;12:e93688
URL: https://cancer.jmir.org/2026/1/e93688
doi:10.2196/93688

© Elizabeth Travis, Mollie Louise Price, Millie Slessor, Jacqueline Crewe, Laura Ashley. Originally published in JMIR Cancer (https://cancer.jmir.org), 5.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identification of Early Signs of Mental Health Disorders in Older Survivors of Cancer Using Patient-Generated Health Data: Observational Study

Georgios Petridis[1*], MSc; Antonios Billis[1*], PhD; Georgios Meditskos[2*], Prof Dr; Athina Tsanousa[3*], PhD; Paraskevas Lagakis[1*], MSc; Juan Carlos Naranjo[4*], MSc; José Zenóglio de Oliveira[5*], BSc; Vania Tavares[5*], PhD; Panagiotis D Bamidis[1*], Prof Dr

[1]Laboratory of Medical Physics & Digital Innovation, School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki, Building D, Entrance 8, 3rd Floor, Aristotle University of Thessaloniki Campus, Thessaloniki, Greece

[2]School of Informatics, Faculty of Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece

[3]Centre for Research and Technology Hellas, Thessaloniki, Greece

[4]MySphera, Valencia, Spain

[5]Capgemini (Portugal), Lisbon, Portugal

[*]all authors contributed equally

Corresponding Author:
Georgios Petridis, MSc
Laboratory of Medical Physics & Digital Innovation, School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki, Building D, Entrance 8, 3rd Floor, Aristotle University of Thessaloniki Campus, Thessaloniki, Greece

Abstract

Background: Older survivors of cancer face heightened risk of depression and anxiety related to cancer experiences, fear of recurrence, and aging-related difficulties. Conventional mental health monitoring approaches, such as clinical assessments and even electronic patient-reported outcomes, are limited by recall bias, patient burden, and infrequent data collection. Emerging patient-generated health data from wearables and smart home devices offer passive, low-burden, continuous monitoring, but their ability to capture mental health risks in older survivors of cancer remains unclear.

Objective: This study aims to explore whether patient-generated health data collected in the wild, either passively or actively, can classify older survivors of cancer as having or not having signs of anxiety and depression based on Patient Health Questionnaire-4 (PHQ-4) scores and to assess the potential added value of passive monitoring modalities, such as smart plugs.

Methods: This study recruited 41 older survivors of cancer (mean age 72.3, SD 6.81 years) from the LifeChamps project. Over a 12-week period, participants were monitored using an activity tracker to measure physical activity, sleep, and physiological metrics; a smart scale to capture weight and body composition; and a smart plug to track television use as a proxy for sedentary television viewing. Mental health status was self-reported via the PHQ-4 questionnaire in a mobile app. Machine learning models were trained to classify mental health risk based on features derived from each sensor modality, both independently and in combination.

Results: Tree-based gradient boosting models showed good performance in classifying PHQ-4–defined mental health risk. The best-performing configuration, combining smart plug and smart scale features, achieved a mean F_1-score of 0.77 (SD 0.15) and a mean area under the receiver operating characteristic curve (AUC) of 0.85 (SD 0.10) across 3 repeated train-test splits. Standalone smart plug models, based solely on passive television use patterns, achieved a mean F_1-score of 0.66 (SD 0.04) and a mean AUC of 0.71 (SD 0.06), outperforming models that relied only on activity tracker data (mean F_1 0.59, SD 0.2). Multimodal combinations tended to improve average performance but did not consistently yield large gains over the strongest single-modality configurations, likely reflecting adherence-related data loss for wearables and scales. Crucially, passive monitoring of television use patterns emerged as a promising behavioral proxy measure of mental health states.

Conclusions: This study pioneers the use of passively collected data (eg, smart plugs) for mental health monitoring in older survivors of cancer, demonstrating their potential. Smart plugs capture behavioral patterns without user burden, with reasonable standalone performance (mean AUC 0.71, SD 0.06), positioning them as a promising low-burden modality. Future work should validate findings in larger independent cohorts and in prospective clinical workflows. Such technologies could transform monitoring for vulnerable populations, enabling scalable, inclusive care while reducing health care burdens.

International Registered Report Identifier (IRRID): RR2-https://doi.org/10.1016/j.soncn.2023.151437

XSL•FO
RenderX

(JMIR Cancer 2026;12:e75050) doi:10.2196/75050

KEYWORDS

behavioral data analysis; survivors of cancer; digital biomarkers; mental health monitoring; nonintrusive monitoring; patient-generated health data

Introduction

Overview

Older survivors of cancer face a unique set of challenges, including increased susceptibility to mental health issues such as depression and anxiety, 2 of the most common mental health disorders [1]. The intersection of aging, cancer survival, and mental health presents a complex clinical manifestation. These individuals often struggle with psychological distress stemming from their cancer diagnosis, treatment experiences, and the fear of cancer recurrence, all of which can significantly impact their quality of life (QoL) [2-4]. Furthermore, the combination of cancer recovery with aging processes can exacerbate mental health issues, making early detection and continuous monitoring critical for this demographic [5]. Emerging research underscores the importance of monitoring mental health to improve QoL among this demographic group, emphasizing the need for innovative approaches that are both effective and nonintrusive [6], so as to not exert additional burden on this vulnerable population.

Background and Current Approaches

To address these challenges, the efficacy of using technological advancements for mental health monitoring has been widely explored in recent years.

A 2021 study [7] explored the potential of consumer-grade wearables, specifically Fitbit devices, to assess depression severity through digital biomarkers such as step count, heart rate, and sleep patterns. Their study, which focused on a cross-sectional working population in Singapore, demonstrated that wearable metrics could be significantly associated with depressive symptoms, with a sensitivity of 73% and specificity of 79%. This work underscores the feasibility of using common activity trackers to screen for mental health issues in general adult populations and presents a scalable model for remote monitoring of behavioral health indicators.

In an older adult population of 352 individuals, Lee et al [8] used low-cost activity trackers from Fitbit to monitor circadian rhythms and sleep patterns, finding correlations between these metrics and symptoms of depression and anxiety. This study highlighted that specific circadian rhythm markers, such as intraday stability and variability, can serve as predictors for geriatric depression and anxiety. The study's findings supported the utility of passive, low-cost monitoring solutions for older adults, a critical demographic for interventions targeting mental health, by showcasing 69.8% accuracy for depression classifiers when using data only from the activity trackers.

Similarly, a study by Pedrelli et al [9] combined wearable (Empatica wristbands) and smartphone data to predict depressive symptom severity among 31 adults diagnosed with major depressive disorder. This research used multimodal data, including electrodermal activity, heart rate, and physical activity, as well as social engagement metrics from smartphones. The combination demonstrated that passive monitoring could reliably predict symptom changes, achieving high correlations between features and symptoms as well as a moderate mean absolute error between 3.8 and 4.7, emphasizing the potential of machine learning models in complex mental health assessments.

Furthermore, a systematic review [10] highlighted various studies that apply wearable sensors for passive data collection to detect symptoms of anxiety and depression. One of the significant findings was the identification of key behavioral markers detectable by wearable devices, such as heart rate variability, sleep disturbances, and physical activity levels, which consistently correlate with mental health conditions. However, the review highlighted a scarcity of research focused on older populations, where mental health burdens may intersect with age-related barriers to technology use. Additionally, Kim et al [11] recruited 20 participants aged 69 - 90 years for a 90-day period, monitoring daily activities using passive infrared motion sensors. Their neural network-based classifier achieved more than 94% area under the receiver operating characteristic curve (AUC), suggesting that passive monitoring sensors can be effectively used to detect depression levels.

Problem Statement

Despite the promising technological advancements outlined earlier, current standard-of-care methods for monitoring mental health still rely heavily on self-reported questionnaires and clinical assessments. These approaches present several challenges; they often rely on patients' ability to recall, assess, and document their experiences, which can be hindered by memory recall biases, especially among older adults [12]. Additionally, the burden of frequent clinical visits and the adherence required for consistent self-reporting can be cumbersome, leading to reduced compliance and potentially delayed identification of mental health issues [13]. Even electronic patient-reported outcomes do not fully alleviate these burdens, as they still depend on active patient participation and do require an adequate level of digital literacy [14].

The sporadic nature of these assessment methods may fail to capture timely variations in mental health status, resulting in missed opportunities for early intervention. While the integration of wearable devices and smart home technologies offers a promising avenue to overcome these limitations [10], the efficacy of these technologies in accurately identifying mental health problems within vulnerable populations, such as older survivors of cancer, remains underexplored [15]. Specifically, the use of purely passive devices, such as smart plugs that monitor energy consumption and household appliance use, has not yet been thoroughly studied and remains an emerging topic [16].

XSL•FO
RenderX

Objectives

To bridge this gap, this study aims to explore the added value that patient-generated health data (PGHD) collected in the wild, especially the ones that are passively collected, may contribute in terms of older survivors' health-related QoL monitoring. Rather than performing a full clinical validation, we treat this work as an exploratory readiness assessment of whether machine learning models built on PGHD can meaningfully classify Patient Health Questionnaire-4 (PHQ-4)–defined anxiety and depression symptoms at each questionnaire occasion based on the preceding 30 days of sensor data, and what additional signal is contributed by different sensor modalities. In addition, this work attempts to compare the different modalities and their contribution toward the most accurate identification of mental health risk.

Specifically, this study contributes to the existing body of knowledge by:

- Evaluating the performance of machine learning (ML) models using data from smart plugs, wearables, smart scales, and their combination, in classifying the mental health status of older survivors of cancer.
- Comparing the discriminative contribution and potential added value of these unobtrusive monitoring modalities—individually and in combination—with particular emphasis on fully passive smart plug data as a low-burden option for longitudinal monitoring.

By demonstrating the discriminative value of ambient home sensors compared and in addition to wearables, this work offers evidence-based insights for integrating unobtrusive, cost-effective monitoring techniques into standard postcancer treatment surveillance and care [17].

In summary, our aim was to evaluate whether PGHD from smart plugs, activity trackers, and smart scales—alone and in combination—can classify PHQ-4–defined anxiety and depression symptoms in older survivors of cancer at the time of each questionnaire occasion based on the preceding 30 days of sensor data and to quantify the incremental discriminative signal contributed by fully passive smart plug data.

Methods

Study Design and Setting

This work is an observational study embedded in the Greek pilot site of the LifeChamps project [18], a multidisciplinary initiative designed to enhance the QoL for older survivors of cancer. The pilot involved home-based remote monitoring over approximately 12 weeks per participant using wearable devices, smart-home sensors, and a mobile app. Recruitment for this study took place between October 2022 and December 2023. All study procedures were conducted in collaboration with oncology and urology clinics in public and private hospitals or clinics in Greece, where patients were approached by their treating clinicians.

Participants

Participants were recruited from the LifeChamps Greek pilot. Treating medical professionals (oncologists and urologists) identified potentially eligible survivors of breast and prostate cancer from their regular caseloads and invited them to participate during clinic visits.

The recruitment flow was as follows:

- Contacted: A total of 102 patients were approached (breast cancer: n=71 and prostate cancer: n=31).
- Consented: Of these, 50 (49%) patients (breast cancer: n=30 and prostate cancer: n=20) agreed to participate, provided written informed consent, and had the monitoring equipment installed at home.
- Excluded or refused: The remaining 52 patients either refused participation (n=35) or were deemed ineligible or unsuitable (n=17).
- Completed: Among the 50 participants who consented, 45 (90%) completed the full 12-week monitoring protocol.

For this analysis, we included the 41 participants who both completed follow-up and had at least 1 eligible postbaseline PHQ-4 assessment with corresponding sensor (smart plug) data. These 41 older survivors of cancer (female participants with breast cancer: n=24 and male participants with prostate cancer: n=17) had all completed their primary treatment and were either on adjuvant therapy or under surveillance.

Inclusion and Exclusion Criteria

Eligible participants were older survivors of breast or prostate cancer under the regular care of the participating clinicians and living in the community. The specific criteria are presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. Eligibility criteria.

Inclusion criteria

- Demographics: Age ≥65 years
- Cancer diagnosis or survivorship status: Histologically confirmed breast or prostate cancer, living beyond initial cancer treatment (surgery, chemotherapy, and/or radiotherapy), under adjuvant therapy or surveillance (as applicable), absence of diagnosed secondary malignancy
- Functional and cognitive status: Deemed physically and psychologically fit to participate (multidisciplinary team judgment); able to read, write, and understand Greek; Mini-Cog score ≥ 3
- Technological requirements: Own Android smartphone (version 10+), 24/7 home internet access (Wi-Fi and/or 4G; study provided mobile data where needed)
- Ethics or administrative: Able and willing to provide written informed consent

Exclusion criteria

- Clinical status: Terminal stage of cancer, prognosis <18 months from recruitment, currently receiving chemotherapy, active and unstable metastatic disease (clinician judgment)
- Mental or cognitive conditions impacting participation: Current diagnosis of a major mental disorder or cognitive disorder affecting the ability to participate
- Safety or device compatibility: Presence of an internally fitted medical device (eg, pacemaker), known allergy to metal or plastic that would preclude device use
- Ethics or administrative: Unwillingness or inability to provide written informed consent

Ethical Considerations

The study protocol was reviewed and approved by the Research Ethics and Deontology Committee of the Aristotle University of Thessaloniki (protocol 267203; approval date: October 13, 2022). All participants provided written informed consent before installation of any sensors. The consent process described the types of data collected (including physical activity, sleep, physiological metrics, weight and body composition, and household appliance power consumption) and the duration of monitoring and clarified that participation was voluntary and could be withdrawn at any time without consequences for clinical care. To protect participant privacy, all sensor data were pseudonymized at the source and stored using unique participant codes; the linkage key was stored separately on a secure local server accessible only to authorized clinical staff. No identifiable images were collected or used in this paper or Multimedia Appendix 1. Participation was voluntary, and no financial compensation was provided; however, participants were allowed to keep the activity tracker or smart scale upon study completion as a token of appreciation.

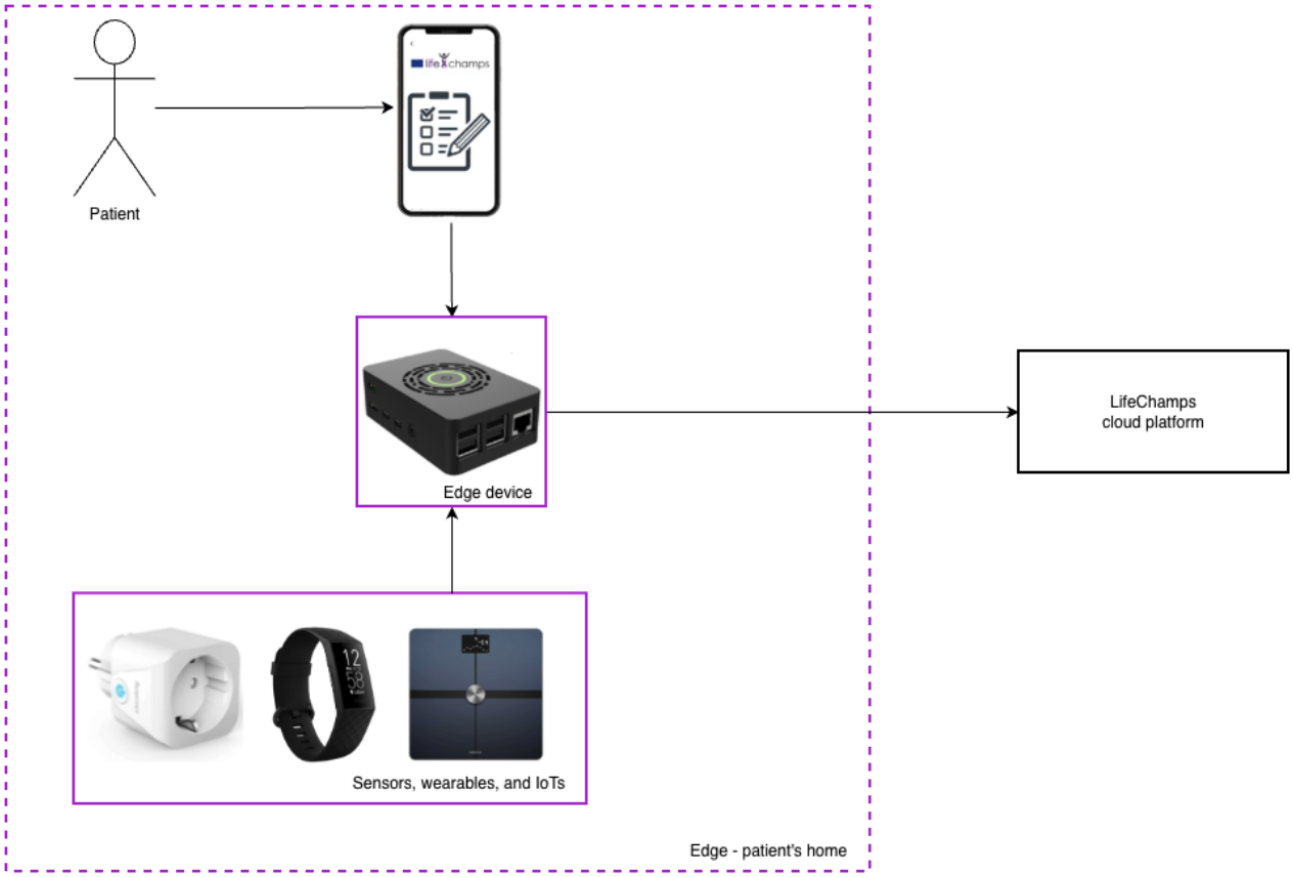
Data Collection

Overview

The study spanned a 12-week period, during which participants' homes were equipped with the LifeChamps Edge platform [19], a high-level overview of which can be seen in Figure 1. This platform included a suite of sensors and Internet of Things devices that tracked a breadth of activity and biometric data, constructing a comprehensive overview of each participant's health and daily routine.

XSL•FO
RenderX

Figure 1. Overview of the LifeChamps home monitoring platform used in this observational study, showing the integration of Fitbit, Withings smart scale, and smart plug data with the central Edge device and data warehouse. IoT: Internet of Things.



Activity Tracker

This wristband (namely the Fitbit Charge 4), worn consistently except during charging times, offered a stream of data covering physical activity (including duration, intensity, type, number of steps or floors, and calories burned), sleep metrics (including duration, efficiency, time in various sleep stages, time to fall asleep, sleep onset, wake time, and number of awakenings), and biometric parameters such as heart rate at different states (eg, during activities, resting heart rate, and sleep heart rate), and oxygen saturation levels. Data were synchronized with the vendor's cloud service and then integrated into the LifeChamps platform.

Smart Scale

Participants engaged with the smart scale (namely the Withings Body+) on a weekly basis, which measured not only weight but also provided detailed body composition analytics, including muscle mass, fat mass, bone mass, and BMI. Unlike the continuous data stream from the Fitbit, the Withings scale provided discrete, yet vital, snapshots of the participants' body composition over time.

Smart Plug

Integrated into the Edge platform and connected to the patient's television (the main television used by them in case of multiple ones available), the smart plug (specifically the Teckin SP22) was used to monitor the television's power status, recording on or off transitions every 5 minutes. This served as a proxy for one of the participants' activities of daily living, offering insights into sedentary and leisure habits. Television watching is a sedentary behavior that has been associated with mental health status in prior work using validated psychometric tools [20].

In addition to sensor data, baseline demographic and clinical characteristics (eg, age, sex [male or female], marital and employment status, living situation, cancer type and stage, and Mini-Cog score [21]) were collected at study entry and later used as candidate features in the models. Participants also received a mobile app that served as the interface for self-report questionnaires, including the PHQ-4 mental health assessment described below.

Mental Health Assessment

Participants in LifeChamps also received a mobile app, allowing them to self-report patient-reported outcome measures (PROMs) monthly. The PHQ-4 self-reported questionnaire [22] was used as the main self-report measure for mental health, assessing anxiety and depression with scores from 0 to 12, interpreted as normal mental health status (0 - 2) or mild (3-5), moderate (6-8), and severe (9-12) symptoms of anxiety and depression. It includes 2 subscales for depression and anxiety, each total subscore ranging from 0 to 6, with subscores of 3 or above indicating symptom presence. PHQ-4 was selected as a compromise between more elaborate tools like Patient Health Questionnaire-9 and ultrabrief self-reports like Patient Health Questionnaire-2, based also on previous studies that included survivors of cancer and used the same questionnaire [23]. Although questionnaires like Patient Health Questionnaire-9 are not time-consuming on their own, in the larger context of

XSL•FO
RenderX

the LifeChamps study where several PROMs were administered each month to a vulnerable group of people such as survivors of cancer, the decision was taken to opt in for brief versions of these PROMs whenever possible. On the other hand, the ultrabrief Patient Health Questionnaire-2 was not selected since it focuses only on depressive symptoms [24].

Participants were asked to report the PHQ-4 at 4 time points: at baseline and then monthly for 3 months. The first questionnaire was completed with researcher assistance, while subsequent ones were self-administered via app notifications at each specified interval. For this analysis, the baseline self-reports were not considered since they had no sensor data prior to their completion.

Each postbaseline PHQ-4 assessment, together with the corresponding 30 days of preceding sensor data, was treated as a single observation (questionnaire occasion) in the modeling analyses. Thus, the unit of analysis in this study is the PHQ-4 questionnaire occasion rather than the participant, and individual participants may contribute multiple observations over time. For the binary classification models, PHQ-4 total scores were subsequently dichotomized into "presence" versus "absence" of mental health signs, as described in the Data Preprocessing section.

Data Preprocessing

We conducted a row-wise, participant-agnostic classification analysis to classify dichotomized PHQ-4 scores from demographics, questionnaire data, and sensor-derived features, treating each questionnaire plus its preceding 30 days of sensor data as a concurrent observation. As described earlier, each postbaseline PHQ-4 assessment, together with the corresponding 30 days of preceding sensor data, was treated as a single observation (questionnaire occasion). Because participants could contribute more than 1 questionnaire over time, multiple observations could originate from the same individual.

PHQ-4 scores were transformed to suit the needs of binary classification for the presence or absence of symptoms. This transformation was guided by clinical relevance and the nature of mental health progression in older survivors of cancer. Thus, all PHQ-4 PROMs with score 3 and above were categorized as "presence of mental health signs" to avoid any ambiguity, according to the suggested cutoff point for the Greek general population [25] and the one for patients with breast cancer [26], while the rest were assigned to class "no mental health signs."

For each questionnaire occasion, we reconstructed the original feature blocks: (1) demographic and clinical characteristics (eg, age, sex, marital and employment status, living situation, Mini-Cog score, cancer type and stage, and treatment intent); (2) additional patient-reported outcomes (Brief Illness Perception Questionnaire, Medication Adherence Report Scale, Vulnerable Elders Survey-13, Edmonton Symptom Assessment System, and Linear Analog Self-Assessment); and (3) sensor-derived aggregates from the activity tracker (Fitbit), smart scale (Withings), and smart plug (television use). From the smart plug, we derived features such as daily total television on-time, daily off-time, and summary statistics (eg, mean and SD) across the 30-day window. We also experimented with features stratified by time of day (morning, afternoon, and night) and by weekday versus weekend; however, in the final models, we focused on simpler daily statistics, which yielded more stable performance in our experiments. To align with the monthly PHQ-4 assessments, we summarized daily and weekly sensor data into monthly aggregates. This process did not only involve simple aggregations, such as totals and averages, but also the computation of statistical measures that capture the variability and distributional characteristics of the data within each month. For each feature, we calculated the coefficient of variation to assess relative variability, the IQR around the median value, and the maximum and minimum values to encapsulate the range. Kurtosis was computed to understand the distribution, which can be indicative of behavioral patterns.

From the set of completed PHQ-4 questionnaires, we first excluded baseline assessments (which lacked prior sensor data and were used only for descriptive purposes) and any follow-up questionnaires without any smart plug data in the preceding 30-day window. To ensure that smart plug–derived features reflected meaningful television use patterns, we further excluded questionnaire occasions with fewer than 15 days of valid television data within the 30-day window or with 0 recorded television on-time during that period. As a valid day for the smart plug, we consider the days for which we have at least 10 hours of data. Fifteen days were selected as a threshold since they attribute 50% of the expected data within a month, a common threshold for considering missing values as manageable with imputation methods [27]. We selected smart plug data as a prerequisite, since the main aim of our analysis is to determine the efficiency of this completely passive monitoring modality as a standalone sensor and also compare it to the rest of the modalities.

Preventing data leakage, that is, avoiding any inadvertent use of information from the held-out test set during model training, was a priority, and as such, all algorithmic preprocessing steps were carefully applied to maintain the test set's integrity. Automated feature selection techniques [28], such as importance thresholding, were used alongside clinical expertise and literature to inform the selection of features related to health-related QoL. To reduce the risk of overfitting, we removed columns that were direct identifiers (eg, patient ID), and we dropped ultrasparse features with more than 70% missingness within the filtered dataset before any imputation. For the remaining features, numerical variables were imputed using an iterative multivariate imputation procedure (IterativeImputer), while categorical variables (eg, sex, treatment intent, employment status, and living situation) were imputed using the most frequent category, and then, one-hot encoding was applied. All imputation steps were fitted on the training data only within each resampling split. To address class imbalance in the binary PHQ-4 outcome, we applied the synthetic minority oversampling technique (SMOTE) strictly within the cross-validation folds to prevent data leakage and overoptimistic performance estimates. The validation sets remained composed entirely of real, unaugmented patient data, ensuring that all reported metrics reflect performance on actual participants.

XSL•FO
RenderX

The resulting feature set was scaled using standardization, adjusting the data to have a mean of 0 (SD 1), so that features were on comparable scales for algorithms that are sensitive to feature scaling, without changing their relative information content.

The final dataset, a temporal alignment of Fitbit, Withings scale, and smart plug data with monthly PHQ-4 scores, was crucial for the effective identification of patterns by our ML algorithms. The different categories of features extracted from each sensor are presented in Table S1 in Multimedia Appendix 1.

Modeling and Evaluation

We framed the task as a supervised binary classification problem, aiming to classify the presence or absence of mental health signs at each questionnaire occasion based on the preprocessed feature sets described earlier. The dichotomized PHQ-4 outcome (total score >2 vs ≤2) served as the target variable, and each PHQ-4 questionnaire plus its preceding 30 days of sensor data was treated as a separate sample. Thus, the models estimate how well PGHD can discriminate between questionnaire occasions with and without PHQ-4–defined signs within the same 30-day window, rather than predicting mental health status at a future time point.

Several model families commonly used for tabular clinical data were evaluated, including elastic net–regularized logistic regression, random forests, and HistGradientBoosting classifiers. All models were implemented in scikit-learn and combined with the same preprocessing pipeline (imputation, encoding, scaling, oversampling, and feature selection) applied in a consistent way across sensor configurations. Hyperparameters were tuned using small, predefined grids chosen to balance flexibility with the limited sample size.

To obtain more robust performance estimates and reduce dependence on a single train-test split, we used a repeated stratified train-test evaluation with an inner cross-validation loop for model selection. For each of 3 repetitions (3 different random seeds), we performed a stratified 80/20 split of the 52 samples into a training set (80%) and an independent test set (20%), preserving the class distribution. On the 80% training portion, we then carried out 10-fold stratified cross-validation to tune hyperparameters, with AUC used as the primary optimization criterion.

Within this inner cross-validation, the full preprocessing pipeline (including imputation, encoding, scaling, oversampling with SMOTE, and random forest–based feature selection) was fitted on the training folds only and applied to the corresponding validation folds, ensuring that no information from the held-out test set leaked into model training. After tuning, the selected model was refit on the full 80% training set and evaluated on the held-out 20% test set. This procedure was repeated independently for each seed and for each sensor configuration (single modalities and combinations), yielding 3 test-set performance estimates per configuration.

We assessed model performance using standard classification metrics: AUC, F_1-score, precision, and recall. For each configuration, we report the mean performance on the 3 held-out test sets together with the corresponding SD across seeds (interrun variability). Hyperparameters were tuned using inner cross-validation, with AUC as the optimization criterion. Given the modest sample size (52 questionnaire occasions from 27 participants, with 21 “presence” and 31 “no mental health signs”), these analyses are intended as exploratory and hypothesis-generating, and model complexity as well as hyperparameter ranges were deliberately constrained to mitigate overfitting. No formal a priori sample size calculation was performed; the sample size was constrained by the LifeChamps pilot and was considered appropriate for exploratory, hypothesis-generating analyses. To mitigate overfitting with 52 questionnaire occasions (21 “presence” and 31 “absence” of signs), we restricted model complexity and used repeated resampling and nested cross-validation rather than more flexible model families.

Missing Data

Missing data in this study arose from 2 main sources: participant nonadherence to the protocol and technical issues affecting data capture. For the PHQ-4 outcome, questionnaires that were not completed at a given time point were simply absent and were not imputed; only completed postbaseline PHQ-4 assessments with corresponding sensor data were included in the modeling dataset. As described earlier, we additionally excluded questionnaire occasions with no smart plug data or with fewer than 15 valid television days in the preceding 30-day window, because smart plug–derived features were central to our research question and the smart plug application programming interface did not allow backfilling of data after Edge-device downtime.

In the final analytic dataset (52 questionnaire occasions from 27 participants; 174 features after removing ultrasparse features >70% missingness), overall cell-level missingness was 32.5% (2943/9048). After applying the television coverage criteria, smart plug features were complete (0% missing), while missingness remained for demographics (17.3%), activity tracker features (33.4%), and smart scale features (53.6%).

For sensor-derived features, partial missingness was common, particularly for the activity tracker and smart scale, due to factors such as low digital literacy, forgetting to wear or charge the device, or not using the scale regularly. To ensure data quality, strict “validity criteria” were applied. A “valid day” for Fitbit data required >10 hours of wear time. Data from invalid days were not taken into account. In the preprocessing pipeline, ultrasparse features with more than 70% missingness were removed before modeling, and the remaining numeric and categorical features were imputed using the iterative and most-frequent strategies described earlier, fitted only on the training data within each resampling split. Because PHQ-4 served as the supervised learning target (label), we did not impute missing PHQ-4 outcomes; only completed postbaseline PHQ-4 assessments were included in the modeling dataset, as outcome imputation would require strong assumptions and could introduce bias if missingness was informative.

To explore the missingness mechanism in the final analytic dataset, we first note that Little’s global test for data missing completely at random (MCAR) can be numerically unstable when the number of variables greatly exceeds the number of observations, and its P values should be interpreted cautiously.

XSL•FO
RenderX

We therefore supplemented this with pairwise MCAR 2-tailed *t* test diagnostics that compare the distribution of each observed variable between rows where another variable is missing versus observed (as implemented in commonly used missingness-exploration toolkits). In total, 2269 of 19,082 (11.9%) tests were significant at $P<.05$, exceeding the approximately 954 of 19,082 (~5%) expected under MCAR, which is consistent with structured (non-MCAR) missingness. This was expected because missingness occurred largely at the device or modality level—when a participant did not use the Fitbit or Withings scale during a given period, all features derived from that device were simultaneously absent.

Importantly, we did not find clear evidence that missingness was associated with the study outcome (PHQ-4 score>2). The prevalence of psychological distress did not differ significantly between questionnaire occasions with comparatively complete versus incomplete sensor data, nor did participant-level mean missingness show a significant association with outcome status (Spearman $\rho=-0.24$; $P=.23$). Sex (male or female) distribution was comparable between completeness groups ($P=.87$), whereas age differed (68.4 vs 71.9 years; $P<.001$), indicating that younger participants tended to have more complete device data. Together, these findings suggest that missingness depended on observable covariates and device-use behavior—patterns consistent with (though not proving) a missing at random mechanism.

Results

Participant Characteristics and Final Sample

A total of 41 participants met the analytic inclusion criteria (completion of follow-up and at least 1 eligible postbaseline PHQ-4 assessment with corresponding smart plug data). These older survivors of cancer included 24 female participants with breast cancer and 17 male participants with prostate cancer. Their baseline sociodemographic and clinical characteristics are summarized in Table 1.

Table . Baseline sociodemographic and clinical characteristics of the 41 older survivors of breast and prostate cancer enrolled in the LifeChamps Greek pilot (Greece, 2022 - 2023).

Variable	Values
Sex assigned at birth, n (%)	
Male	17 (41.5)
Female	24 (58.5)
Age (years), mean (SD)	72.3 (6.81)
Employment status, n (%)	
Retired	20 (48.8)
Employed	2 (4.9)
Unemployed	1 (2.4)
Unknown	18 (43.9)
Cancer type, n (%)	
Curative	37 (90.2)
Metastatic	4 (9.8)
Living situation, n (%)	
Alone	4 (9.8)
Cohabitation	19 (46.3)
Unknown	18 (43.9)
PHQ-4[a] baseline assessment, n (%)	
Healthy	33 (80.5)
Mild symptoms	8 (19.5)

[a]PHQ-4: Patient Health Questionnaire-4.

Following the process described in the previous section, out of 123 expected postbaseline monthly follow-up questionnaires (41 participants×3 monthly follow-ups), 93 were collected from 41 participants due to varying adherence to the protocol. The distribution of PHQ-4 classes based on the participants' answers is shown in Table 2.

XSL•FO
RenderX

Table . Distribution of PHQ-4[a] categories (normal, mild, moderate, and severe) across all completed postbaseline monthly follow-up questionnaires (n=93) among the 41 participants.

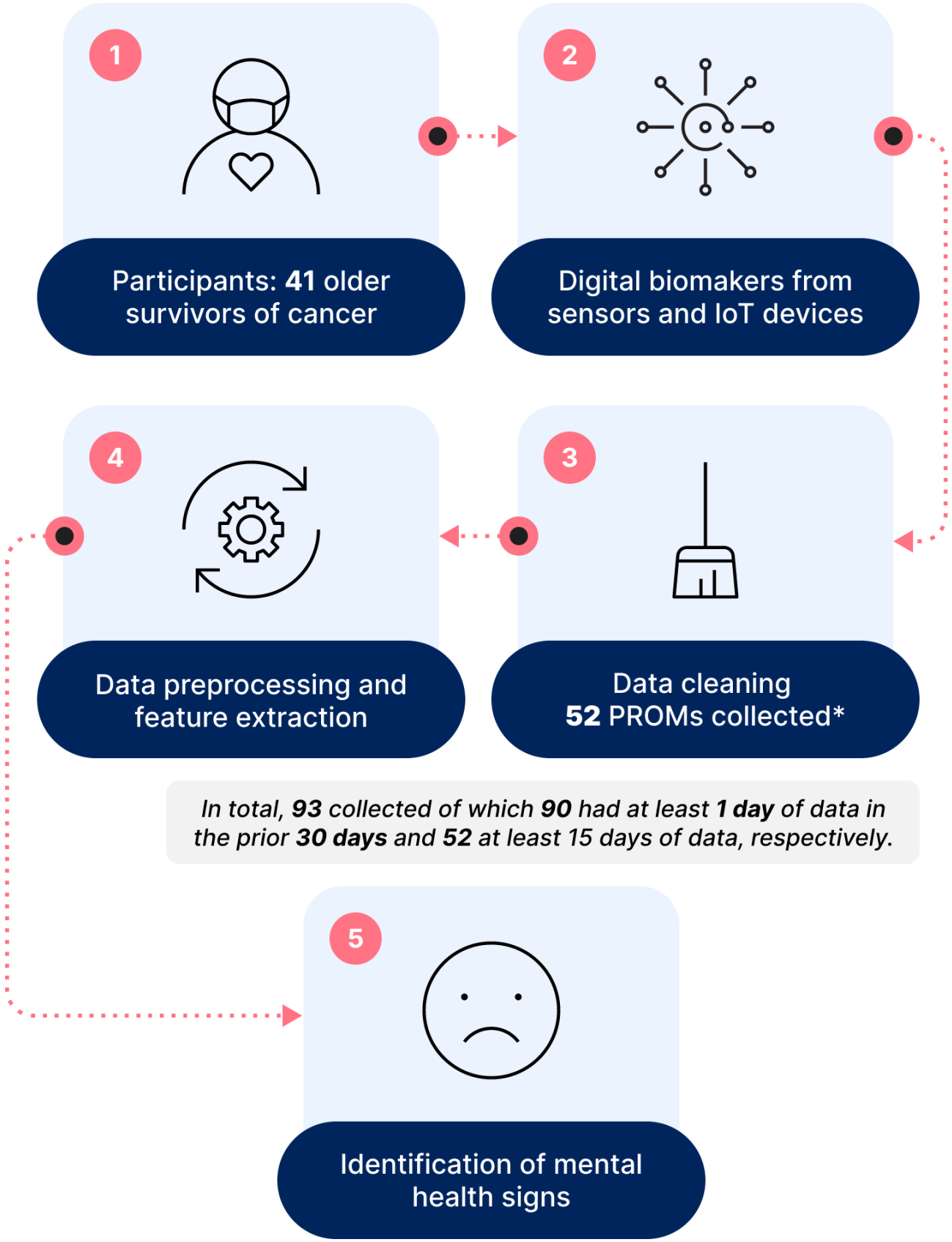
PHQ-4 class	Questionnaires in class, n (%)
0 (normal)	30 (32.3)
1 (mild)	49 (52.7)
2 (moderate)	10 (10.8)
3 (severe)	4 (4.3)

[a]PHQ-4: Patient Health Questionnaire-4.

After applying the preprocessing steps detailed previously, which are also presented in Figure 2, we derived a dataset comprising 52 questionnaire occasions contributed by 27 participants. Of them, 31 are categorized as “no mental health signs,” and 21 are identified as “presence of mental health signs.”

XSL•FO
RenderX

Figure 2. Analysis pipeline from raw sensor data and monthly PHQ-4 questionnaires to the final analytic dataset of 52 questionnaire occasions and training of machine learning models. IoT: Internet of Things; PHQ-4: Patient Health Questionnaire-4; PROM: patient-reported outcome measure.



Model Performance

Using the modeling approach described in the Methods section (3 repeated stratified 80/20 train-test splits with inner 10-fold cross-validation), we evaluated HistGradientBoosting models for each sensor configuration. Table 3 summarizes the mean test-set performance across the 3 held-out test sets, together

XSL•FO
RenderX

with the corresponding SDs (interrun variability), for each configuration.

Table . Binary classification model results.

Sensors or features used and metric	Test set, mean (SD across 3 seeds)
Smart plug	
AUC[a]	0.71 (0.06)
F_1	0.66 (0.04)
Precision	0.76 (0.08)
Recall	0.66 (0.03)
Activity tracker	
AUC	0.64 (0.14)
F_1	0.59 (0.20)
Precision	0.60 (0.20)
Recall	0.59 (0.20)
Smart scale	
AUC	0.67 (0.24)
F_1	0.67 (0.17)
Precision	0.67 (0.17)
Recall	0.67 (0.17)
Smart plug+activity tracker	
AUC	0.78 (0.14)
F_1	0.74 (0.06)
Precision	0.74 (0.05)
Recall	0.74 (0.06)
Smart plug+smart scale	
AUC	0.85 (0.10)
F_1	0.77 (0.15)
Precision	0.78 (0.17)
Recall	0.76 (0.14)
Activity tracker+smart scale	
AUC	0.73 (0.09)
F_1	0.64 (0.04)
Precision	0.74 (0.10)
Recall	0.65 (0.03)
All sensors	
AUC	0.85 (0.15)
F_1	0.76 (0.12)
Precision	0.78 (0.13)
Recall	0.74 (0.11)

[a]AUC: area under the receiver operating characteristic curve.

Across all models, the best-performing configuration was the combination of smart plug and smart scale features, which achieved a mean test AUC of approximately 0.85 (SD 0.10) and a mean F_1-score of 0.77 (SD 0.15). The model using all 3 sensor modalities (“all sensors”) showed similarly high performance (AUC=0.85; F_1=0.76), indicating that combining multiple data sources can provide reasonably accurate

XSL•FO
RenderX

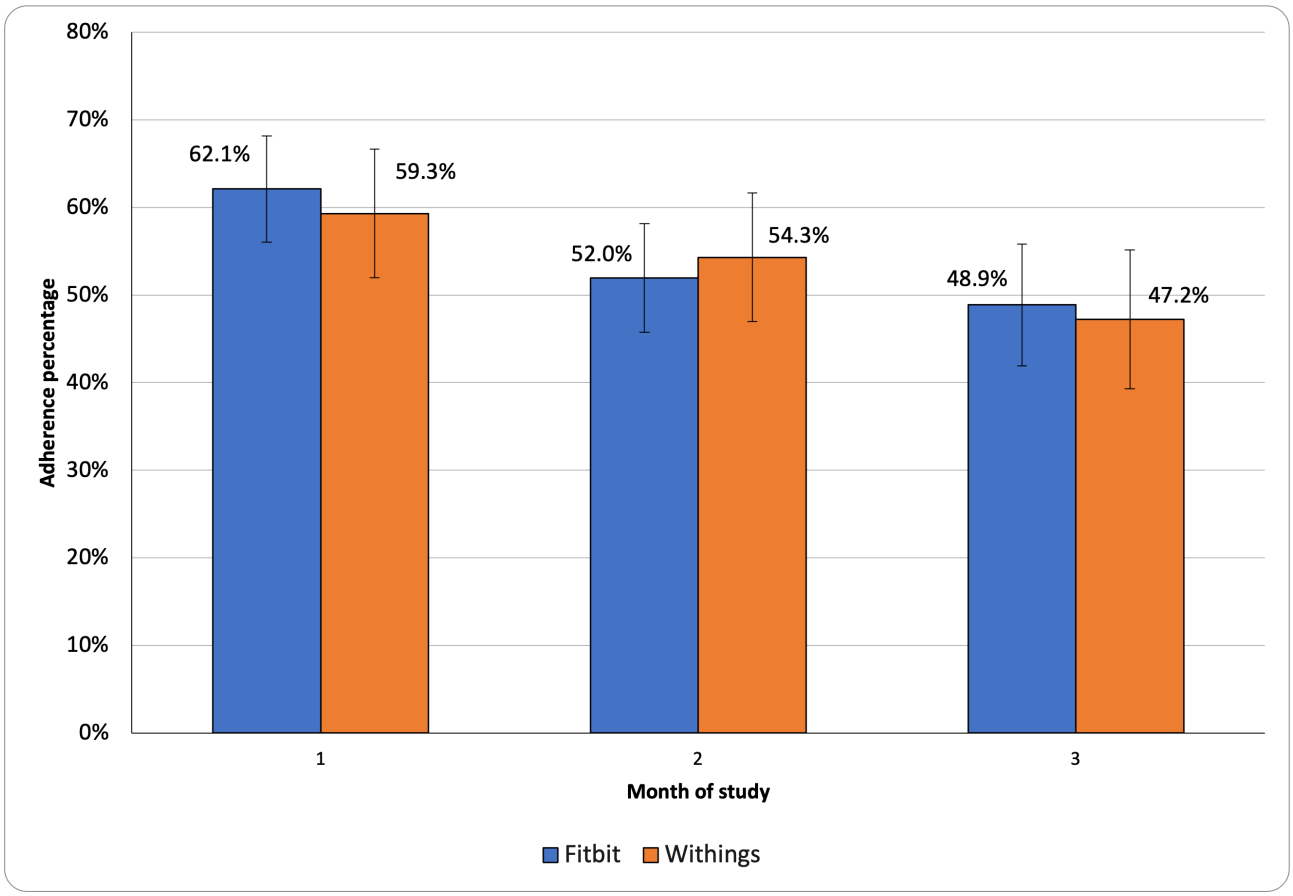
classification of PHQ-4–defined mental health signs in this small cohort.

Among single-modality models, smart plug–only and smart scale–only configurations performed comparably, with mean F_1-scores of about 0.66 (SD 0.04) and 0.67 (SD 0.17) and mean AUC values of approximately 0.71 (SD 0.06) and 0.67 (SD 0.24), respectively. The activity tracker–only model showed somewhat lower performance (F_1=0.59; AUC=0.64). Models combining smart plugs with the activity tracker (AUC=0.78; F_1=0.74) or with the smart scale (AUC=0.85; F_1=0.77) tended to outperform single-sensor models, suggesting that complementary information from different devices can improve classification in this dataset.

The comparative efficacy of the smart plug features is particularly intriguing when contrasted with the activity tracker–only model. The latter is widely regarded in the literature as capturing signals associated with mental health indicators [29]. Nevertheless, the standalone smart plug model's performance suggests that behavioral data related to sedentary habits may provide a useful behavioral signal associated with PHQ-4–defined mental health risk in our cohort.

In the case of Fitbit's and Withings' data, the models exhibit moderate effectiveness with an AUC score of 0.64 and 0.67, respectively. These results indicate that, although these devices capture clinically relevant health metrics, their standalone discriminative value for classifying mental health risk appeared somewhat lower than that of models including smart plug features in this small sample. This difference may partly reflect declining adherence to wearing the activity tracker and using the scale regularly over the course of the study (Figure 3), which reduced the completeness and temporal density of the corresponding data streams.

Figure 3. Mean adherence percentage of Fitbit (blue) and Withings (orange) through the 3 months of the study, in the original sample. Error bars represent the SE of the mean.



Given the sample size and the variability observed across random splits, the reported performance estimates should be interpreted within the context of this experimental setup. They reflect how each modality behaves under the current data conditions and provide a consistent basis for comparing their relative performance.

Discussion

Principal Findings

This study explored whether PGHD collected in the wild can be used to classify PHQ-4–defined mental health signs in older survivors of cancer and to what extent completely passive smart plug data contribute to this task. Using a small sample of 52 questionnaire occasions from 27 participants, we found that models based on smart plug data, particularly when combined with smart scale measurements, achieved promising

XSL•FO
RenderX

classification performance (test AUC around 0.85; F_1-scores around 0.76 - 0.77). Smart plug–only and smart scale–only models performed comparably to each other and better than activity tracker–only models, suggesting that television use patterns and weight or body composition metrics may contain informative behavioral signals associated with mental health risk in this cohort. These findings should be interpreted as preliminary and hypothesis-generating, but they support the idea that low-burden passive home sensing can contribute meaningfully to mental health monitoring in older survivors of cancer. Because both PGHD summaries and PHQ-4 scores refer to the same 30-day window, these analyses should be interpreted as cross-sectional discriminative associations between behavioral patterns and PHQ-4–defined mental health signs, not as prognostic models of future mental health trajectories.

Our results complement prior work on PGHD and digital mental health monitoring. Existing studies have shown that consumer wearables can capture behavioral and physiological markers associated with depression and anxiety, such as heart rate variability, sleep, and activity patterns [7-10]. Studies in older adults, including those using low-cost activity trackers or ambient motion sensors [8,11], have demonstrated associations between sensor-derived measures, circadian rhythms, and depressive symptoms, and some have reported high discrimination in small samples. However, most of these studies are either descriptive (focusing on correlations with symptom scores) or feasibility-oriented and rarely target older survivors of cancer specifically. By evaluating classification models that use smart plug–derived television use as a primary modality, alongside more conventional wearables and smart scales, our work extends this literature to a cancer survivorship context and highlights a novel, fully passive home-sensing signal.

The comparative performance of the different modalities provides some insight into their potential roles in real-world survivorship care. In our data, smart plug–only and smart scale–only models achieved moderate-to-good F_1-scores (approximately 0.66 - 0.67) and AUC values in the 0.67 - 0.71 range, while the activity tracker–only model showed lower performance. When smart plug data were combined with either the scale or the activity tracker, model performance improved, especially for the smart plug+scale configuration. This pattern suggests that behavioral information about sedentary and leisure-time routines (television use) and anthropometric trends may complement each other for mental health risk classification, whereas activity tracker data in this sample may have been more affected by missingness and adherence constraints. We deliberately refrain from interpreting these differences as evidence that one modality is intrinsically superior; rather, they likely reflect a combination of signal content and data completeness in this specific cohort and setting.

From a clinical and implementation perspective, smart plugs and other passive smart-home devices have practical advantages for older survivors of cancer. Unlike wrist-worn activity trackers and connected scales, which require regular wearing, charging, and interaction, smart plugs operate entirely in the background once installed and configured. This low-burden data collection is particularly relevant for older adults living with multimorbidity, cancer-related fatigue, or low digital literacy, where even small self-management demands can reduce adherence. Our findings suggest that such devices can capture behavioral routines that are informative for mental health risk classification, and that combining them with minimally demanding periodic measurements (eg, weekly weight) may offer a feasible compromise between burden and information content. If validated in larger and more diverse samples, similar passive-sensing setups could be incorporated into survivorship follow-up to flag individuals who may benefit from more detailed assessment rather than to replace clinical evaluation.

Methodologically, we adopted a row-wise, participant-agnostic classification framework in which each PHQ-4 questionnaire and its preceding 30 days of sensor data constituted a separate observation. This choice allowed us to maximize the use of available data in a small sample, but it also means that multiple observations from the same individual are not strictly independent. In addition, we used SMOTE-based oversampling within the training folds to mitigate class imbalance and a repeated 80/20 train-test evaluation with inner cross-validation for model selection. Although these are commonly used techniques in the applied ML literature, there is active debate about the optimal handling of imbalance and the best validation schemes for small samples. Alternative approaches, such as participant-level cross-validation (eg, leave-one-participant-out and leave-one-group-out) and formal reclassification metrics (eg, net reclassification improvement and integrated discrimination improvement), were beyond the scope of this exploratory work but would be appropriate for future, larger-scale studies focused on model development and clinical validation.

Limitations

We acknowledge limitations within our study. First, the sample size is modest (41 participants and 52 questionnaire occasions) and considering that participation required smartphone use and willingness to accept in-home monitoring, which may preferentially include more motivated and digitally literate individuals, leading to self-selection bias, performance estimates are therefore imprecise and sensitive to data partitioning. Despite using repeated stratified splits and reporting interrun variability, the reported metrics may overestimate or underestimate the true performance in other populations, and external validation in independent cohorts is needed. Second, the analytic dataset required sufficient smart plug coverage (≥15 valid television days in the preceding month), which means that we effectively conditioned the analysis on the availability of high-quality smart plug data. This design decision aligns with our primary objective—assessing the potential value of a fully passive modality—but may introduce selection bias and reduce direct comparability with models that could, in principle, be trained without smart plug data.

Third, missing data were frequent, especially for the activity tracker and smart scale, due to nonadherence (eg, forgetting to wear or charge the device) and technical issues. Although we used a principled preprocessing pipeline that removed ultrasparse features, applied imputation only within training folds, and avoided outcome imputation, residual missingness

XSL•FO
RenderX

and imputation-related uncertainty remain. In addition, the smart plug application programming interface did not allow for backfilling data during Edge-device downtime, leading to irreversible gaps even when participants adhered to normal television use. In addition, PHQ-4 noncompletion at specific time points was treated as a simple absence of an outcome. We did not attempt to model "informative" missingness in questionnaire responses, and it is possible that individuals who felt particularly unwell or distressed were less likely to complete PROMs. Fourth, PHQ-4 was used as a brief self-report proxy for mental health symptoms rather than a clinician-administered diagnostic assessment, and it is therefore subject to recall and reporting biases. In addition, PHQ-4 has a 14-day recall window, whereas in this study, it was administered approximately monthly; this mismatch may blur transient symptom fluctuations and introduce measurement noise. We used a cutoff of 3 based on prior work in the Greek general population and breast cancer samples [25,26], but this threshold has not been validated specifically for older Greek survivors of cancer, and alternative cutoffs may yield different classification boundaries. Finally, television on-time is only a proxy for sedentary viewing; although a preliminary check using overlapping Fitbit step data suggested that participants were inactive during most television-on periods, we cannot exclude scenarios where other household members used the television while the participant was elsewhere in the home.

These limitations highlight the need for follow-up work along several directions. Future studies should aim to recruit larger and more diverse samples of older survivors of cancer, include independent external validation cohorts, and adopt participant-level validation schemes that more directly reflect prospective deployment. In addition to smart plugs and smart scales, other fully passive sensors (eg, ambient motion sensors and door or bed sensors) could be integrated to capture complementary aspects of daily routines, while maintaining low participant burden. On the outcome side, combining brief screeners such as PHQ-4 with more comprehensive, clinician-administered assessments in a subset of participants would enable more rigorous evaluation of classification performance and more robust calibration of clinically meaningful thresholds. Ultimately, the goal is not to replace clinical judgment but to evaluate whether unobtrusive, home-based PGHD streams can help identify older survivors of cancer whose mental health may be changing and who might benefit from more timely support.

Conclusions

This study provides preliminary evidence that completely passive data collection, such as smart plug–based monitoring, may help identify mental health signs among older survivors of cancer, either as a standalone data source or in combination with other sensing modalities. The smart plug's ability to capture granular, behavioral data without requiring patient interaction makes it a promising tool for passive health monitoring. The standalone efficacy of smart plug data highlights the importance of selecting specific, meaningful features over more comprehensive but potentially redundant data. Future research should aim to validate these findings in larger populations and explore the integration of additional nonintrusive sensors to further enhance the classification accuracy and robustness of mental health monitoring models in vulnerable population groups, such as older survivors of cancer.

By leveraging digital biomarkers and PGHD, this study paves the way for integrating smart home technology into regular health assessments. Such integration has the potential to transform care strategies for older survivors of cancer and reduce the burden on health care systems worldwide, making mental health monitoring more accessible and inclusive.

Acknowledgments

The authors would like to thank the LifeChamps consortium. The authors especially thank the health care professionals who assisted in the recruitment of patients, namely, P Sountoulides, S Levva, D Dionysopoulos, E Karamitrousis, A Sachoulidou, I Xanthakis, A Karabeazis, and V Sakalis. Finally, the authors would like to thank Helena Benito Escriba for her help with the preprocessing of smart plug data. The authors further acknowledge Luciano Colares for his earlier contribution to the work. Generative artificial intelligence tools (eg, ChatGPT and OpenAI) were used to assist with language editing and improving the clarity of some sections of the manuscript. All scientific content, analysis, and interpretation were conceived, verified, and approved exclusively by the authors.

Funding

This work was supported in part by the European Commission under the LifeChamps Project (grant 875329) and the COMFORT project (grant 101079894). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ongoing analyses but are available from the corresponding author on reasonable request.

Authors' Contributions

GP led the manuscript preparation and coordinated the submission and revision process. GP, AB, GM, AT, PL, JCN, JZO, VT, and PB contributed substantially to the conception and design of the study, interpretation of the findings, and critical revision of

XSL•FO
RenderX

the manuscript for important intellectual content. All listed authors reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of candidate features per modality.
[DOCX File, 17 KB - cancer_v12i1e75050_app1.docx]

References

1. National Collaborating Centre for Mental Health (UK). Common mental health disorders. In: Common Mental Health Disorders: Identification and Pathways to Care: British Psychological Society; 2011.
2. Deimling GT, Bowman KF, Sterns S, Wagner LJ, Kahana B. Cancer-related health worries and psychological distress among older adult, long-term cancer survivors. Psychooncology 2006 Apr;15(4):306-320. [doi: 10.1002/pon.955] [Medline: 16041841]
3. Schroyen S, Marquet M, Jerusalem G, et al. The link between self-perceptions of aging, cancer view and physical and mental health of older people with cancer: a cross-sectional study. J Geriatr Oncol 2017 Jan;8(1):64-68. [doi: 10.1016/j.jgo.2016.09.001] [Medline: 27659547]
4. Marshall-McKenna R, Kotronoulas G, Kokoroskos E, et al. A multinational investigation of healthcare needs, preferences, and expectations in supportive cancer care: co-creating the LifeChamps digital platform. J Cancer Surviv 2023 Aug;17(4):1094-1110. [doi: 10.1007/s11764-022-01289-7]
5. Aziz NM, Rowland JH. Trends and advances in cancer survivorship research: challenge and opportunity. Semin Radiat Oncol 2003 Jul;13(3):248-266. [doi: 10.1016/S1053-4296(03)00024-9] [Medline: 12903014]
6. Elkefi S, Trapani D, Ryan S. The role of digital health in supporting cancer patients' mental health and psychological well-being for a better quality of life: a systematic literature review. Int J Med Inform 2023 Aug;176:105065. [doi: 10.1016/j.ijmedinf.2023.105065] [Medline: 37224644]
7. Rykov Y, Thach TQ, Bojic I, Christopoulos G, Car J. Digital biomarkers for depression screening with wearable devices: cross-sectional study with machine learning modeling. JMIR Mhealth Uhealth 2021 Oct 25;9(10):e24872. [doi: 10.2196/24872] [Medline: 34694233]
8. Lee TR, Kim GH, Choi MT. Identification of geriatric depression and anxiety using activity tracking data and minimal geriatric assessment scales. Appl Sci (Basel) 2022 Jan;12(5):2488. [doi: 10.3390/app12052488]
9. Pedrelli P, Fedor S, Ghandeharioun A, et al. Monitoring changes in depression severity using wearable and mobile sensors. Front Psychiatry 2020;11:584711. [doi: 10.3389/fpsyt.2020.584711] [Medline: 33391050]
10. Hickey BA, Chalmers T, Newton P, et al. Smart devices and wearable technologies to detect and monitor mental health conditions and stress: a systematic review. Sensors (Basel) 2021 May 16;21(10):3461. [doi: 10.3390/s21103461] [Medline: 34065620]
11. Kim JY, Liu N, Tan HX, Chu CH. Unobtrusive monitoring to detect depression for elderly with chronic illnesses. IEEE Sensors J 2017 Jul 20;17(17):5694-5704. [doi: 10.1109/JSEN.2017.2729594]
12. Stone AA, Shiffman S. Capturing momentary, self-report data: a proposal for reporting guidelines. Ann Behav Med 2002;24(3):236-243. [doi: 10.1207/S15324796ABM2403_09] [Medline: 12173681]
13. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11. [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
14. Meirte J, Hellemans N, Anthonissen M, et al. Benefits and disadvantages of electronic patient-reported outcome measures: systematic review. JMIR Perioper Med 2020 Apr 3;3(1):e15588. [doi: 10.2196/15588] [Medline: 33393920]
15. Shaffer KM, Turner KL, Siwik C, et al. Digital health and telehealth in cancer care: a scoping review of reviews. Lancet Digit Health 2023 May;5(5):e316-e327. [doi: 10.1016/S2589-7500(23)00049-3] [Medline: 37100545]
16. Korenhof SA, Fang Y, Luo J, van der Cammen TJM, Raat H, van Grieken A. Monitoring the well-being of older people by energy usage patterns: systematic review of the literature and evidence synthesis. JMIR Aging 2023 Mar 31;6(1):e41187. [doi: 10.2196/41187] [Medline: 37000477]
17. Muhsen IN, Rasheed OW, Habib EA, et al. Current status and future perspectives on the internet of things in oncology. Hematol Oncol Stem Cell Ther 2023 Jan 17;16(2):102-109. [doi: 10.1016/j.hemonc.2021.09.003] [Medline: 34687614]
18. Papachristou N, Kartsidis P, Anagnostopoulou A, et al. A smart digital health platform to enable monitoring of quality of life and frailty in older patients with cancer: a mixed-methods, feasibility study protocol. Semin Oncol Nurs 2023 Jun;39(3):151437. [doi: 10.1016/j.soncn.2023.151437] [Medline: 37149438]
19. Billis A, Lagakis P, Petridis G, et al. Enable care of older cancer survivors with digital health technologies: the LifeChamps project. Presented at: 2023 IEEE 19th International Conference on Body Sensor Networks (BSN); Oct 9, 2023; Boston, MA, United States p. 1-4. [doi: 10.1109/BSN58485.2023.10331429]

XSL•FO
RenderX

20. Billis A, Kartsidis P, Garyfallos DKG, Tsatali MS, Karagianni M, Bamidis P. Ecologically valid trials of elderly unobtrusive monitoring: analysis and first results. Presented at: AI-AM/NetMed 2015: 4th International Workshop on Artificial Intelligence and Assistive Medicine; Jun 20, 2015 URL: https://ceur-ws.org/Vol-1389/paper4.pdf [accessed 2026-05-29]
21. Limpawattana P, Manjavong M. The mini-cog, clock drawing test, and three-item recall test: rapid cognitive screening tools with comparable performance in detecting mild NCD in older patients. Geriatrics (Basel) 2021 Sep 16;6(3):91. [doi: 10.3390/geriatrics6030091] [Medline: 34562992]
22. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
23. Eichler M, Hechtner M, Wehler B, et al. Psychological distress in lung cancer survivors at least 1 year after diagnosis-results of a German multicenter cross-sectional study. Psychooncology 2018 Aug;27(8):2002-2008. [doi: 10.1002/pon.4760] [Medline: 29761591]
24. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
25. Christodoulaki A, Baralou V, Konstantakopoulos G, Touloumi G. Validation of the Patient Health Questionnaire-4 (PHQ-4) to screen for depression and anxiety in the Greek general population. J Psychosom Res 2022 Sep;160:110970. [doi: 10.1016/j.jpsychores.2022.110970] [Medline: 35728338]
26. Weihs KL, Wiley JF, Crespi CM, Krull JL, Stanton AL. Predicting future major depression and persistent depressive symptoms: development of a prognostic screener and PHQ-4 cutoffs in breast cancer patients. Psychooncology 2018 Feb;27(2):420-426. [doi: 10.1002/pon.4472] [Medline: 28600855]
27. Mishra S, Khare D. On comparative performance of multiple imputation methods for moderate to large proportions of missing data in clinical trials: a simulation study. J Med Stat Inform 2014 Jan 1;2(1):9. [doi: 10.7243/2053-7662-2-9]
28. Chandrashekar G, Sahin F. A survey on feature selection methods. Comput Electr Eng 2014 Jan;40(1):16-28. [doi: 10.1016/j.compeleceng.2013.11.024]
29. Correia ATL, Lipinska G, Rauch HGL, Forshaw PE, Roden LC, Rae DE. Associations between sleep-related heart rate variability and both sleep and symptoms of depression and anxiety: a systematic review. Sleep Med 2023 Jan;101:106-117. [doi: 10.1016/j.sleep.2022.10.018] [Medline: 36370515]

Abbreviations

AUC: area under the receiver operating characteristic curve
MCAR: missing completely at random
ML: machine learning
PHQ-4: Patient Health Questionnaire-4
PROM: patient-reported outcome measure
QoL: quality of life
SMOTE: synthetic minority oversampling technique

Edited by M Balcarras; submitted 15.Apr.2025; peer-reviewed by A Anastasiou, CJ Han, P Gallos, S Pilleron; revised version received 04.Mar.2026; accepted 04.Mar.2026; published 12.Jun.2026.

Please cite as:
Petridis G, Billis A, Meditskos G, Tsanousa A, Lagakis P, Naranjo JC, Zenóglio de Oliveira J, Tavares V, Bamidis PD
Identification of Early Signs of Mental Health Disorders in Older Survivors of Cancer Using Patient-Generated Health Data: Observational Study
JMIR Cancer 2026;12:e75050
URL: https://cancer.jmir.org/2026/1/e75050
doi:10.2196/75050

© Georgios Petridis, Antonios Billis, Georgios Meditskos, Athina Tsanousa, Paraskevas Lagakis, Juan Carlos Naranjo, José Zenóglio de Oliveira, Vania Tavares, Panagiotis D Bamidis. Originally published in JMIR Cancer (https://cancer.jmir.org), 12.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Fear of Cancer Recurrence Before Treatment in Patients Newly Diagnosed With Breast Cancer: Cross-Sectional Study

Yun Ding*, MS; Huiting Zhang*, MS; Qiaoling Zhong, BS; Lijuan Zhang, BS; Huiying Qin, MS

State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, 651 Dongfeng East Road, Guangzhou, China

*these authors contributed equally

Corresponding Author:
Huiying Qin, MS
State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, 651 Dongfeng East Road, Guangzhou, China

Abstract

Background: Fear of cancer recurrence (FCR) is a common psychological disorder among patients with cancer. In patients newly diagnosed with breast cancer, fear and concern about recurrence often influence treatment decision-making. However, the prevalence of FCR before treatment and its key risk factors—such as age, sleep quality, economic status, self-disclosure, and social support—remain unclear.

Objective: This study aims to describe the prevalence and severity of FCR before treatment in patients newly diagnosed with breast cancer and analyze its predictive factors, thereby providing a scientific basis for interventions to reduce such fear.

Methods: A cross-sectional study was conducted among patients newly diagnosed with breast cancer prior to treatment at a specialized oncology hospital in Guangdong Province, China, between February 2023 and July 2023. Sociodemographic data were collected, including age, gender, ethnicity, religious belief, education level, and marital status. Clinical data included pathological stage, tumor type, and family history of cancer. The Distress Disclosure Index, the Chinese version of the Medical Social Support Scale, and the short form of the Fear of Progression Questionnaire were used. Univariate analyses and bivariate correlations were conducted to explore relationships between variables. Multiple regression analysis was performed to identify predictors of FCR.

Results: A total of 350 patients newly diagnosed with breast cancer were enrolled. Among the participants, 48.3% (n=169) experienced FCR before initiating treatment. The mean total score for FCR was 33.44 (SD 8.00). Age, sleep quality, financial burden, monthly household income per capita, family relationships, self-disclosure, and social support accounted for 33.5% (R^2) of the variance in FCR.

Conclusions: FCR prior to treatment is common and relatively severe among patients newly diagnosed with breast cancer. Health care professionals should implement targeted interventions to promote self-disclosure and strengthen social support networks to alleviate FCR, facilitate informed treatment decision-making, and improve patients' psychological readiness for treatment. The findings of this study underscore the urgent need for early identification and management of FCR in patients newly diagnosed with breast cancer. High levels of pretreatment FCR may impair patients' psychological well-being and influence treatment decision-making, potentially leading to suboptimal treatment choices. Health care professionals should implement targeted interventions that promote self-disclosure and strengthen social support networks to alleviate FCR. These measures can enhance patients' psychological readiness for treatment, facilitate informed decision-making, and ultimately improve clinical outcomes.

(JMIR Cancer 2026;12:e89346) doi:10.2196/89346

KEYWORDS

fear of cancer recurrence; FCR; breast cancer; self-disclosure; social support

Introduction

According to the latest data released by the International Agency for Research on Cancer, breast cancer ranks as the second most common cancer globally, following lung cancer. In 2022, approximately 2.3 million new cases of breast cancer were reported worldwide, accounting for 11.7% of all newly diagnosed malignancies [1]. In recent years, significant advances have been made in the diagnosis and treatment of malignant tumors. However, the challenges of recurrence and metastasis remain unresolved. These continue to be formidable barriers that modern medicine has yet to overcome. Even patients with ductal carcinoma in situ of the breast face a recurrence risk ranging from 4% to 20% [2]. Fear of cancer recurrence (FCR)

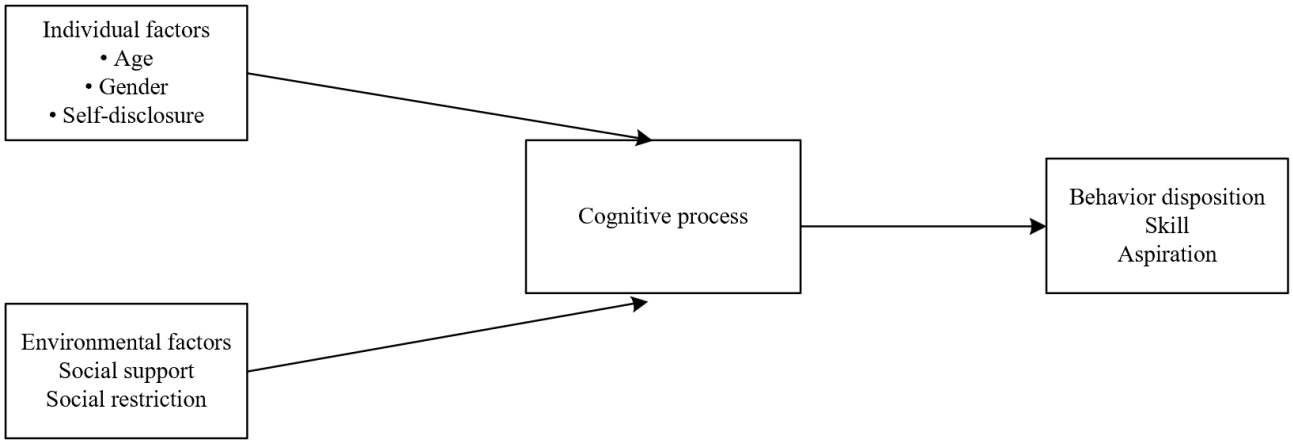
XSL•FO
RenderX

refers to the fear, worry, and anxiety experienced by patients regarding the potential return or progression of cancer [3]. This psychological state often begins at the time of cancer diagnosis and may persist for many years thereafter, affecting patients' emotional well-being and behavioral responses [4]. A study by Lidia et al [5] revealed that approximately 64.5% of patients with breast cancer experience FCR prior to treatment. Pretreatment FCR is likely to influence patients' treatment decisions. For example, a systematic review indicated that some Asian women opt for mastectomy due to concerns about recurrence and a desire to avoid further treatment. As a result, they may miss the opportunity for breast-conserving surgery [6]. A qualitative study also indicated that survivors of breast cancer carrying the breast cancer gene (BRCA1/2) mutations may undergo prophylactic mastectomy due to fears of recurrence [7]. Thus, recognizing and addressing pretreatment FCR in patients with breast cancer is essential to support optimal treatment choices.

However, current literature primarily focuses on the occurrence of FCR among patients with breast cancer during or after treatment. Findings indicate that approximately half of these patients experience moderate to high levels of FCR, which negatively impacts their quality of life, psychological well-being, and treatment adherence [5,7]. In contrast, limited research has been conducted on patients newly diagnosed with breast cancer prior to treatment. Only a few studies have reported the presence of FCR during the period between diagnosis and the initiation of treatment [5,7]. However, the severity of this fear and its associated factors remains insufficiently explored. Therefore, conducting in-depth research on FCR before treatment in patients newly diagnosed with breast cancer holds significant clinical importance. Such research not only addresses a critical gap in the existing literature but also provides a foundation for early intervention strategies.

According to social cognitive theory [8], human behavior and psychological outcomes are shaped by the dynamic interaction between personal factors, environmental influences, and cognitive processes. In the context of cancer diagnosis, personal factors such as self-disclosure and environmental factors such as social support may influence patients' psychological adaptation through cognitive appraisal mechanisms (Figure 1). Self-disclosure, a core concept in psychology, refers to the process by which individuals share personal and private thoughts and emotions with others [9]. As a positive and effective form of communication, self-disclosure has been shown to negatively predict FCR [10]. Several studies have demonstrated that interventions promoting self-disclosure can enhance communication between patients with cancer and others, correct misconceptions regarding cancer recurrence, foster accurate illness beliefs, and ultimately reduce FCR. Specifically, self-disclosure allows patients to verbalize their fears and reorganize illness-related cognitions, which may reduce maladaptive interpretations of cancer recurrence risk [11-13]. Furthermore, social support has been identified as an independent predictor of FCR and plays a crucial role in alleviating it [14]. Low levels of social support are associated with increased FCR among patients with breast cancer, whereas high levels of support can encourage positive perceptions of health status and buffer the psychological distress caused by stressful events, thereby mitigating FCR [15,16]. At the same time, social support provides informational and emotional resources that help patients reinterpret threatening stimuli, correct misconceptions about recurrence, and strengthen perceived coping efficacy. Through these cognitive processes—such as illness perception, risk appraisal, and coping self-efficacy—patients may experience lower levels of FCR. Based on this theoretical framework, we hypothesize that higher levels of self-disclosure and stronger perceived social support are associated with lower levels of pretreatment FCR in patients newly diagnosed with breast cancer.

Figure 1. Framework of social cognitive theory.



Patients newly diagnosed with breast cancer represent a unique population in which psychological factors play a pivotal role in treatment decision-making. Despite growing interest in the psychological dimensions of cancer care, no existing studies have empirically established the relationship between pretreatment FCR and self-disclosure or social support in this specific group. Given the potential influence of psychological readiness on treatment choices, it is crucial to explore whether targeted interventions that enhance self-disclosure and strengthen social support networks can effectively reduce FCR. Such interventions may ultimately empower patients to make more informed and appropriate treatment decisions. However, further research is needed to examine these associations and

XSL•FO
RenderX

assess the clinical value of early psychosocial interventions in patients newly diagnosed with breast cancer.

Therefore, this study aims to describe the current status of FCR prior to treatment in patients newly diagnosed with breast cancer and to identify its risk factors based on social cognitive theory. The objectives of this study are as follows: (1) to explore the incidence and severity of pretreatment FCR in patients newly diagnosed with breast cancer; and (2) to identify the key risk factors contributing to pretreatment FCR in this population, thereby providing empirical evidence to inform the development of effective intervention strategies aimed at reducing FCR levels.

Methods

Study Design

This study used a cross-sectional design and was conducted from February 2023 to July 2023.

Ethical Considerations

Ethical approval was obtained from the ethics committee of Sun Yat-sen University Cancer Center (B2023-700-01). The procedures used in this study adhere to the tenets of the Declaration of Helsinki, which emphasize respect for the dignity, autonomy, and privacy of research participants; the minimization of potential risks and the maximization of possible benefits; and the requirement for voluntary and informed consent prior to participation. Participants' personal data were treated with strict confidentiality, and their rights and well-being were protected throughout the research process. Participants did not receive any compensation for their participation in this study.

Participants

A convenience sampling method was used to recruit patients with breast cancer who sought medical care at a specialized oncology hospital in Guangdong Province, China, between February 2023 and July 2023. Inclusion criteria were as follows: (1) being aged 18 years or older; (2) first-time pathological diagnosis of stage 0-3 breast cancer; (3) no prior breast cancer–related treatment; and (4) an educational level of primary school or higher. Exclusion criteria included: (1) metastatic or recurrent breast cancer; (2) a history of other major organ diseases or malignancies; and (3) a diagnosis of psychiatric disorders or cognitive impairment. Written informed consent was obtained from all participants.

According to the empirical sample size calculation method based on regression analysis [17], the sample size should be 5 to 10 times the number of variables. The predictive factors included self-disclosure, social support, tumor stage, type, and a total of 23 influencing factors. Considering a 20% rate of invalid responses, the minimum sample size was calculated as 23×10/(1–0.20)=287.5 cases. Therefore, the minimum required sample size for this study was determined to be 288 cases.

Measures

Sociodemographic and Clinical Characteristics

Sociodemographic data collected in this study included age, sex, ethnicity, religious beliefs, education level, marital status, number of children, average age of the youngest child, average age of all children, educational attainment, monthly household income per capita, financial burden, place of residence, living arrangement, family relationships, occupation prior to diagnosis, and method of medical expense payment. Disease-related variables included clinical stage, tumor type, comorbidities, family history of cancer, and sleep quality. Sleep quality was assessed using a single self-reported item in which participants rated their overall sleep status as "good," "fair," or "poor" during the period prior to treatment.

The Short Form of the Fear of Progression Questionnaire

The short form of the Fear of Progression Questionnaire was developed by German scholar Mehnert et al [18] in 2006 by simplifying the original Fear of Progression Questionnaire and is designed to assess the level of fear of disease progression in patients with cancer. In 2015, Wu Qiyun et al [19] translated and culturally adapted the scale into Chinese. The short form of the Fear of Progression Questionnaire consists of 12 items, categorized into 2 dimensions: physical health and social or family concerns. Each item is rated on a 5-point Likert scale, with all items scored positively. A total score of 34 or higher is commonly used to indicate a high level of fear, based on validation studies of the Chinese version of the scale [19]. The Cronbach α coefficient of the scale in this study was 0.883, indicating good internal consistency.

Revised Distress Disclosure Index

The Distress Disclosure Index was originally developed by Kahn and Hessling [20] to assess an individual's tendency to disclose distressing experiences. In 2008, Chinese scholar Li Xinmin translated and revised the scale for use in the Chinese population to evaluate levels of self-disclosure [21]. The revised Chinese version consists of 12 items and uses a 5-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). Items 2, 4, 5, 8, 9, and 10 are reverse scored. The total score ranges from 12 to 60, with higher scores indicating a greater level of self-disclosure, while lower scores reflect a higher tendency toward concealment. The Cronbach α coefficient of the scale is 0.86, indicating good internal consistency.

Simplified Chinese Medical Outcomes Study Social Support Survey

The Simplified Chinese Medical Outcomes Study Social Support Survey (MOS-SSS-CM) was originally developed by Sherbourne et al [22] in 1991 to assess perceived social support among patients with chronic illnesses. In 2012, Li Huan [23] adapted the scale into a simplified Chinese version (MOS-SSS-CM) for use in the Chinese population. The total scale demonstrated good reliability, with a Cronbach α coefficient of 0.889 and a test-retest reliability of 0.77, indicating high internal consistency and stability. The scale contains 20 items across 4 dimensions: emotional or informational support, tangible support, affectionate support, and positive social interaction. Each item is rated using a 5-point Likert scale ranging from 1 to 5, based on the individual's perceived frequency of support. The total score is converted to a 0-to-100

XSL•FO
RenderX

scale, with higher scores reflecting a higher level of perceived social support.

Data Collection

For patients who met the inclusion and exclusion criteria, data collection was conducted after hospital admission and pathological diagnosis, but prior to the initiation of treatment. Upon obtaining written informed consent, a professionally trained research staff member distributed the questionnaires and guided participants in their completion. After completion, the researcher checked each questionnaire for completeness. Prior to the start of the study, this staff member received comprehensive training on research ethics, questionnaire content, completion guidelines, and communication skills with patients, ensuring effective interaction and the provision of appropriate support throughout the process.

To ensure data quality and minimize potential bias, standardized procedures were implemented during questionnaire administration. The research staff members received structured training on research ethics, questionnaire content, standardized instructions, and communication skills prior to data collection. All questionnaires were completed independently by participants, and responses were checked for completeness without influencing participants' answers. Specifically, the researcher responsible for questionnaire administration was not provided with information about the patients' treatment plans or clinical conditions during data collection, thus minimizing potential bias in the data collection process.

Data Analysis

Data analysis was conducted by an independent researcher with extensive experience in clinical research and professional training in oncology and statistics. The dataset was deidentified prior to analysis to ensure confidentiality and objectivity.

All statistical analyses were performed using SPSS version 25.0 (IBM Corp), with a significance level set at α=.05; *P* values less than .05 were considered statistically significant. Descriptive statistics were used to characterize participants' sociodemographic and clinical characteristics. Univariate analyses or bivariate correlations were initially conducted to identify potential factors associated with levels of FCR. ANOVA was used to examine differences in FCR scores across various sociodemographic and clinical characteristics (eg, education level, marital status, employment status, and household income). Pearson correlation analysis was performed to assess the relationships among self-disclosure, social support, and FCR in patients with breast cancer.

Variables demonstrating statistical significance in univariate analyses were subsequently entered into a multivariate regression model. A multiple linear regression analysis was conducted to identify independent predictors of FCR. The squared correlation coefficient (R^2) was used to indicate the proportion of variance in the dependent variable that could be explained by the independent variables. A *P* value of 0.05 was considered the threshold for statistical significance.

Results

Analysis of Demographic Characteristics, Self-Disclosure, Social Support, and FCR Among Patients Newly Diagnosed With Breast Cancer Prior to Treatment

A total of 362 patients met the inclusion criteria and provided informed consent. As 12 patients refused to participate in the study, 350 patients finally completed the questionnaire, resulting in a loss rate of 3.31%. The descriptive data on sociodemographic and clinical characteristics, as well as Distress Disclosure Index and MOS-SSS-CM scores, are summarized in Table 1. A total of 350 patients with breast cancer, aged between 25 and 77 years (mean age: 45.39, SD 9.68), were included in the study, with 99.1% (n=347) being female. The majority reported a per capita monthly household income above 5000 RMB (US $689.22). Financial concerns were prevalent, with 63.7% of patients experiencing some degree of economic burden and 15.7% (n=55) reporting severe financial stress. Additionally, 55.7% (n=195) rated their sleep quality as average. The mean (SD) scores for self-disclosure and social support were 38.86 (9.31) and 70.57 (14.04), respectively. The mean FCR score among patients with breast cancer was 33.43 (SD 8.00), with the physical health dimension (18.11, SD 4.12) scoring higher than the social or family dimension (15.32, SD 4.66). Notably, 48.3% (n=169) of patients reported experiencing fear of recurrence prior to the initiation of treatment.

XSL•FO
RenderX

Table . Correlation of the short form of the Fear of Progression Questionnaire (FoP-Q-SF) with sociodemographic characteristics, clinical characteristics, the Distress Disclosure Index (DDI), and the Simplified Chinese Medical Outcomes Study Social Support Survey (MOS-SSS-CM; N=350).

Variables	Value	FoP[a] score, mean (SD)	Statistics	*P* value
Age (y), mean (SD)	45.39 (9.68)	—[b]	10.99 (2, 347)[c]	<.001
Young adults (18-40), n (%)	121 (34.6)	35.34 (7.41)		
Middle-aged adults (41-60), n (%)	202 (57.7)	33.05 (8.06)		
Older adults (61-77), n (%)	27 (7.7)	27.78 (7.21)		
Gender, n (%)			–1.914 (348)[d]	.06
Men	3 (0.9)	24.67 (6.02)		
Women	347 (99.1)	33.51 (7.98)		
Ethnicity, n (%)			–0.192 (348)[d]	.85
Han ethnicity	338 (96.6)	33.0 (10.46)		
Minority	12 (3.4)	33.45 (7.91)		
Faith, n (%)			0.596 (348)[d]	.55
No	330 (94.3)	33.50 (7.97)		
Yes	20 (5.7)	32.40 (8.61)		
Marital status, n (%)			0.178 (2, 347)[c]	.84
Unmarried	17 (4.9)	33.65 (8.90)		
Divorced or widowed	19 (5.4)	34.47 (8.43)		
Married	314 (89.7)	33.36 (7.94)		
Number of children, n (%)			0.491 (3, 346)[c]	.69
Nulliparous	29 (8.3)	32.66 (8.24)		
1	140 (40)	33.05 (7.65)		
2	145 (41.1)	33.68 (8.23)		
≥3	37 (10.6)	34.58 (8.31)		
Average age of the youngest child (y), n (%)			2.487 (348)[d]	.01
18	222 (60.1)	34.24 (7.94)		
≥18	128 (36.6)	32.05 (7.94)		
Average age of all children (y), n (%)			2.433 (348)[d]	.02
<18	206 (55.1)	34.30 (7.89)		
≥18	144 (41.1)	32.20 (8.01)		
Educational level, n (%)			1.866 (2, 347)[c]	.16
Junior high school or lower	96 (27.4)	34.77 (7.73)		
High school or technical secondary school	84 (24)	33.07 (9.39)		
College degree or higher	170 (48.6)	32.86 (7.33)		
Monthly per capita household income (RMB[e]), n (%)			3.695 (2, 347)[c]	.03
≤3000	48 (13.7)	33.56 (8.30)		
3001-5000	111 (31.7)	35.05 (8.55)		
>5000	191 (54.6)	32.47 (7.46)		

XSL•FO
RenderX

Variables	Value	FoP[a] score, mean (SD)	Statistics	*P* value
Financial burden, n (%)			20.624 (2, 347)[c]	<.001
No financial burden	72 (20.6)	29.64 (7.26)		
Moderate financial burden	223 (63.7)	33.45 (7.56)		
Severe financial burden	55 (15.7)	38.36 (8.08)		
Place of residence, n (%)			3.442 (2, 347)[c]	.03
Rural area	81 (23.1)	35.26 (8.29)		
Town area	88 (25.1)	33.69 (7.37)		
Urban area	181 (51.8)	32.50 (8.05)		
Living arrangement, n (%)			0.863 (2, 347)[c]	.42
Living alone	20 (5.7)	35.60 (8.83)		
Living with friends	5 (1.4)	34.80 (11.16)		
Living with family	325 (92.9)	33.28 (7.90)		
Family relationship, n (%)			4.666)（348）[d]	<.001
Nonharmonious	55 (15.7)	37.9 (7.55)		
Harmonious	295 (84.3)	32.60 (7.81)		
Occupation before diagnosis, n (%)			4.820 (2, 347)[c]	.01
Retired	52 (14.9)	30.46 (8.27)		
Unemployed	62 (17.7)	34.85 (7.81)		
Employed	236 (67.4)	33.72 (7.86)		
Type of health care payment, n (%)			0.488 (2, 347)[c]	.61
Self-paid	7 (2)	36.0 (6.21)		
Medical insurance	335 (95.7)	33.72 (7.97)		
Government-funded health care	8 (2.3)	32.0 (10.66)		
Pathological staging, n (%)			1.428 (3, 346)[c]	.23
0	31 (8.9)	34.29 (10.95)		
I	93 (26.6)	32.34 (8.02)		
II	157 (44.9)	34.24 (6.54)		
III	69 (19.7)	32.70 (9.28)		
Tumor type, n (%)			0.582 (348)[d]	.56
Noninvasive carcinoma	37 (10.6)	34.16 (10.21)		
Invasive carcinoma	313 (89.4)	33.35 (7.71)		
Family history of cancer, n (%)			–0.381 (348)[d]	.70
No	251 (71.7)	33.33 (8.00)		
Yes	99 (28.3)	33.70 (8.03)		
Comorbidity, n (%)			0.599 (348)[d]	.55
No	283 (80.9)	33.56 (7.87)		
Yes	67 (19.1)	32.91 (8.57)		
Sleep status, n (%)			9.399 (2, 347)[c]	<.001
Poor	67 (34.6)	35.04 (8.83)		
Fair	195 (34.6)	34.28 (7.34)		

XSL•FO
RenderX

Variables	Value	FoP[a] score, mean (SD)	Statistics	*P* value
Good	88 (34.6)	30.35 (8.01)		
Self-disclosure score, mean (SD)	38.86 (9.31)	—	–0.377[f]	<.001
Social support score, mean (SD)	70.57 (14.04)	—	–0.393[f]	<.001

[a]FoP: Fear of Progression.

[b]Not applicable.

[c]Means comparing the FoP of different characteristics with an ANOVA (*F* test).

[d]Means comparing the FoP of different characteristics with the *t* test. All *t* tests were 2-tailed.

[e]Currency conversion was based on the exchange rate on July 1, 2023 (1 RMB = US $0.137844).

[f]Means Pearson correlation with FoP-Q-SF.

Analysis of Demographic Characteristics, Self-Disclosure, Social Support, and Their Correlations With FCR in Patients Newly Diagnosed With Breast Cancer Before Treatment

The results of the ANOVA indicated that there were statistically significant differences in FCR related to age, the average age of the youngest child, the average age of all children, monthly per capita household income, financial burden, place of residence, family relationships, occupation before diagnosis, and sleep status (P<.05). The Pearson correlation analysis revealed that both self-disclosure (r=−0.394; P<.01) and social support (r=−0.393; P<.01) were significantly and negatively associated with fear of cancer recurrence (Table 1).

Multiple Regression Analysis of FCR in Patients Newly Diagnosed With Breast Cancer

A standard enter method linear regression analysis was performed, with FCR set as the dependent variable. Independent variables were those found to be statistically significant (P<.05) in univariate and correlation analyses, including patient age, average age of the youngest child, average age of all children, monthly per capita household income, financial burden, place of residence, family relationship, occupation before diagnosis, sleep quality, self-disclosure, and social support. A significant regression equation ($F_{11,337}$=15.42; P<.001) was found, with an R^2 of 0.335 (Table 2). Younger age, poorer sleep quality, greater financial burden, lower self-disclosure, and lower social support were independently associated with higher FCR levels. Lower household income showed a trend toward association with higher FCR (P=.05).

XSL•FO
RenderX

Table . Multiple linear regression analysis of fear of cancer recurrence in patients with breast cancer(N=350).

Variables	B[a]	SE *B*	β[b]	*t* test (*df*)[c]	*P* value	R^2	95% CI
Constant	43.05	2.93		14.70 (337)	<.001	0.335	37.30 to 48.82
Age (y)							
Older adults (61-77)	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]
Young adults (18-40)	6.00	1.50	0.358	4.01 (337)	<.001	—[d]	3.06 to 8.94
Middle-aged adults (41-60)	3.76	1.42	0.232	2.65 (337)	.008	—[d]	0.97 to 6.54
Sleep status							
Good	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]
Fair	3.09	0.88	0.192	3.52 (337)	<.001	—[d]	1.36 to 4.82
Poor	2.68	1.15	0.132	2.32 (337)	.02	—[d]	0.42 to 4.95
Monthly per capita household income (RMB[e])							
High income (>5000)	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]
Middle income (3001-5000)	0.19	0.86	0.011	0.22 (337)	.83	—[d]	−1.51 to 1.90
Low income (≤3000)	−2.33	1.21	−0.100	−1.94 (337)	.054	—[d]	−4.71 to 0.04
Financial burden							
No financial burden	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]
Severe financial burden	4.52	1.42	0.206	3.18 (337)	.002	—[d]	1.73 to 7.31
Moderate financial burden	2.35	0.96	0.141	2.44 (337)	.02	—[d]	0.45 to 4.24
Family relationship	−1.97	1.06	−0.09	−1.85 (337)	.07	—[d]	−4.06 to 0.12
Self-disclosure	−0.25	0.04	−0.286	−5.77 (337)	<.001	—[d]	−0.33 to -0.16
Social support	−0.10	0.03	−0.17	−3.15 (337)	.002	—[d]	−0.16 to -0.04

[a]B: unstandardized regression coefficient.

[b]β: standardized regression coefficient.

[c]All *t* tests were 2-tailed.

[d]Not applicable.

[e]Currency conversion was based on the exchange rate on July 1, 2023 (1 RMB = US $0.137844).

Discussion

Patients newly diagnosed with breast cancer represent a unique population, in which psychological factors play a critical role in treatment decision-making. To our knowledge, this study is the first to investigate the relationships between sociodemographic and clinical characteristics, self-disclosure, social support, and FCR prior to treatment, and to identify significant predictors of FCR in this population.

FCR Before Treatment in Patients Newly Diagnosed With Breast Cancer

This study found that FCR was a prevalent psychological issue among patients newly diagnosed with breast cancer prior to treatment, with approximately 50% experiencing clinically significant levels of fear. This proportion was lower than that reported by Han et al [24] but higher than that found by Li et al [25], suggesting heterogeneity in FCR across different patient groups. The discrepancy may be due to differences in study populations and timing of assessment. Han et al [24] focused on patients undergoing chemotherapy, during which FCR may be intensified by side effects and prolonged stress (including financial, occupational, and familial stressors). In contrast, the

XSL•FO
RenderX

participants in this study were at the pretreatment stage, not yet exposed to treatment-related stressors, which may explain their relatively lower FCR. Li et al's [25] study mainly involved postoperative patients. In our study, patients were newly diagnosed, facing difficulty accepting the diagnosis and perceiving uncertainty about future treatment, potentially contributing to high FCR. These findings highlight the importance of early psychological intervention for newly diagnosed patients.

Younger Patients, Poor Sleep Quality, Low Income, and Financial Burden Exhibit Higher Levels of FCR

This study identified age as a significant factor influencing FCR in patients with breast cancer (P<.001), consistent with the findings of Zhang et al [26]. Younger patients tend to be more concerned with body image, and treatment-related changes, such as hair loss from surgery or chemotherapy, may trigger feelings of inferiority or depression, exacerbating FCR.

Sleep quality was significantly associated with FCR in the multivariate analysis, with poorer self-reported sleep linked to higher levels of fear [27]. However, as sleep quality was assessed using a single self-reported item and the study used a cross-sectional design, the directionality of this association cannot be determined. It is possible that sleep disturbances and FCR are interrelated through shared psychological mechanisms, such as heightened cognitive arousal or stress sensitivity. These findings suggest that sleep quality may be an important factor to consider when assessing psychological vulnerability in newly diagnosed patients. Future longitudinal studies using standardized sleep measures (eg, the Pittsburgh Sleep Quality Index) are needed to further clarify the nature and direction of this relationship.

Both univariate and multivariate analyses suggested that lower household income tended to be associated with higher levels of FCR; however, this association did not reach conventional statistical significance in the multivariate model (P=.054). In contrast, perceived financial burden remained a significant independent predictor of FCR, with patients reporting moderate or severe financial burden exhibiting higher FCR levels. These findings are consistent with the results reported by Niu et al [14]. Possible explanations include the substantial cost of cancer treatment and the additional financial concerns related to potential recurrence. Patients facing greater economic strain may experience heightened anxiety and uncertainty about future treatment expenses, which could amplify fear of recurrence. Notably, patients with medical insurance reported lower levels of FCR compared with those who paid out-of-pocket, suggesting that financial protection mechanisms may play a buffering role in psychological adaptation. Given that patients with breast cancer often require costly adjuvant therapies such as chemotherapy or radiotherapy, limited reimbursement for anticancer medications may exacerbate financial stress among economically vulnerable patients. Therefore, health care providers should assess patients' financial situations when formulating treatment plans, offer cost-effective therapeutic options when feasible, and provide targeted psychological support. At the policy level, further strengthening the medical insurance system and expanding coverage for major illnesses may help alleviate financial burden and indirectly reduce FCR.

Lower Levels of Self-Disclosure Are Associated With Higher FCR

This study found a significant negative correlation between self-disclosure and FCR. Self-disclosure, typically expressed through verbal communication of one's thoughts and emotions, serves both as a personal coping mechanism and a psychological intervention strategy. It has been shown to alleviate psychological distress and improve quality of life [28]. Prior research supports these findings. Xu et al [11] reported that effective communication helps patients correct misconceptions about recurrence and fosters healthier beliefs, thereby reducing FCR. In a controlled trial by Chen et al [12], a self-disclosure intervention significantly increased disclosure levels, reduced negative emotions, and lowered FCR in postoperative patients with breast cancer undergoing chemotherapy. Another study found that enhancing spousal self-disclosure also reduced FCR [13]. This may be due to the fact that individuals with higher levels of self-disclosure receive more external feedback—such as encouragement, experiential guidance, or health-related information—which improves their understanding of the disease and relieves emotional stress. Sharing one's illness experience and feelings can also enhance emotional support. Therefore, clinical staff should encourage newly diagnosed patients to engage in active and meaningful communication, such as journaling or peer discussion, to express their thoughts and expectations. This practice can enhance self-disclosure, help patients cope with stress, and assist them in making informed treatment decisions.

Lower Levels of Social Support Are Associated With Higher Levels of FCR

Social support was significantly and negatively correlated with FCR, encompassing both the physical and social or family dimensions. This finding is consistent with the results of Niu et al [14] and Yu et al [15]. A meta-analysis also demonstrated that low levels of social support are associated with increased FCR in patients with breast cancer, while higher levels of support promote positive health perceptions and reduce stress-related recurrence fears [16]. As a protective factor, social support facilitates emotional regulation through interpersonal interaction and helps alleviate symptom distress and fear of recurrence [27]. In the context of FCR, social support acts as a "buffer," mitigating the negative psychological impact of stressful events [29]. Strong social support not only reduces symptom sensitivity and psychological distress, such as anxiety and depression, but also enhances mental health, decreases FCR, and improves quality of life. Furthermore, it enables patients to approach treatment and recovery with greater positivity and optimism.

Limitations

This study has several limitations. First, participants were recruited from a single tertiary cancer hospital in Guangzhou using convenience sampling, which may introduce selection bias and limit generalizability. Patients in a specialized urban cancer center may differ from those in community or rural

XSL•FO
RenderX

settings in terms of socioeconomic status, access to insurance, and psychosocial resources. Given Guangzhou's developed health care infrastructure, findings related to financial burden may be influenced by regional insurance policies and economic conditions. In contrast, associations between psychosocial factors (eg, self-disclosure and social support) and FCR may reflect more broadly applicable cognitive and interpersonal processes. Multicenter studies across diverse regions are needed to confirm the generalizability of these findings. Second, the study captured self-disclosure, social support, and FCR at a single time point. Consequently, it lacks longitudinal data to assess how these psychological variables change over time or impact patients across different stages of treatment and disease progression. Third, the cross-sectional design precludes causal inference. Although significant associations were identified between self-disclosure, social support, financial burden, and FCR, the temporal sequence of these variables cannot be determined. Longitudinal or interventional studies are needed to clarify these potential causal pathways. Fourth, sleep quality was assessed using a single self-reported item rather than a validated instrument, which may introduce reporting bias and limit measurement accuracy. Future studies should use standardized measures, such as the Pittsburgh Sleep Quality Index, to provide a more comprehensive assessment.

Conclusions

Prior to initial treatment, approximately 50% of patients newly diagnosed with breast cancer reported high levels of FCR. Younger age, poorer sleep quality, greater financial burden, lower self-disclosure, and lower perceived social support were associated with higher FCR levels. These findings highlight the importance of early identification of patients at risk for elevated FCR. Although causal relationships cannot be established due to the cross-sectional design, the results provide preliminary evidence to inform future longitudinal and interventional studies aimed at examining whether enhancing psychosocial resources and addressing sleep and financial concerns may help reduce FCR and improve psychological adaptation during treatment decision-making.

Acknowledgments

The authors thank all the participants for their contributions to this study. The authors confirm that no generative artificial intelligence tools were used in the writing, analysis, or preparation of this manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets used in this study can be obtained from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: YD (equal), HZ (equal), QZ (supporting), LZ (supporting), HQ (supporting)
Data curation: YD (lead), HZ (supporting)
Formal analysis: YD (lead), HZ (supporting)
Investigation: YD (lead), HZ (supporting), QZ (supporting)
Methodology: YD (supporting), HZ (supporting), QZ (supporting), LZ (lead), HQ (lead)
Project administration: HQ (lead), LZ (equal)
Resources: HQ (lead), LZ (equal)
Supervision: HQ (lead), LZ (equal)
Validation: YD (supporting), HZ (supporting), QZ (supporting), LZ (supporting), HQ (lead)
Writing – original draft: YD (lead), HZ (lead)
Writing – review & editing: YD (supporting), HZ (supporting), QZ (supporting), LZ (lead), HQ (lead)

YD and HZ are co-first authors.
HQ (qinhy@sysucc.org.cn) and LZ (zhanglij@sysucc.org.cn) are co-corresponding authors of this paper and can be reached at: State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou, China.

Conflicts of Interest

None declared.

References

1. Cancer today. Global Cancer Observatory. URL: https://gco.iarc.who.int/today/en [accessed 2026-03-19]
2. The Society of Breast Cancer China Anti-Cancer Association, Breast Oncology Group of the Oncology Branch of the Chinese Medical Association. Guidelines for breast cancer diagnosis and treatment by China Anti-Cancer Association (2024 Edition). China Oncol 2023;33(12):1092-1187. [doi: 10.19401/j.cnki.1007-3639.2023.12.004]

XSL•FO
RenderX

3. Butow P, Sharpe L, Thewes B, Turner J, Gilchrist J, Beith J. Fear of cancer recurrence: a practical guide for clinicians. Oncology (Williston Park) 2018 Jan 15;32(1):32-38. [Medline: 29447419]
4. Lebel S, Rosberger Z, Edgar L, Devins GM. Emotional distress impacts fear of the future among breast cancer survivors not the reverse. J Cancer Surviv 2009 Jun;3(2):117-127. [doi: 10.1007/s11764-009-0082-5] [Medline: 19322661]
5. Schapira L, Zheng Y, Gelber SI, et al. Trajectories of fear of cancer recurrence in young breast cancer survivors. Cancer 2022 Jan 15;128(2):335-343. [doi: 10.1002/cncr.33921] [Medline: 34614212]
6. Galeano Machuca MP, Cheng SCC, Jou THT, Cheng CT. Determinants of breast-conserving therapy in the Asian population: a systematic review. World J Surg 2021 Mar;45(3):799-807. [doi: 10.1007/s00268-020-05814-6] [Medline: 33051701]
7. Savard J, Filion C, Brearty CM, Caplette-Gingras A, Chiquette J, Dorval M. Fear of cancer recurrence in breast cancer survivors carrying a BRCA1/2 mutation: a qualitative study. Cancer Med 2024 Jan;13(1):e6889. [doi: 10.1002/cam4.6889] [Medline: 38192174]
8. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52(1):1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
9. Mugo C, Firdawsi O, Wang J, et al. “When they are all grown, I will tell them”: experience and perceptions of parental self-disclosure of HIV status to children in Nairobi, Kenya. BMC Public Health 2023 Mar 17;23(1):519. [doi: 10.1186/s12889-023-15387-3] [Medline: 36932351]
10. [FREE Full text]
11. Xu W, Wang J, Schoebi D. The role of daily couple communication in the relationship between illness representation and fear of cancer recurrence in breast cancer survivors and their spouses. Psychooncology 2019 Jun;28(6):1301-1307. [doi: 10.1002/pon.5082] [Medline: 30946501]
12. Chen J, Cheng H, Wei Q, Zhong C, Xu Y, et al. Self-disclosure intervention for patients undergoing chemotherapy following breast cancer surgery. J Nurs Sci 2022;37(14):86-89. [doi: 10.3870/j.issn.1001-4152.2022.14.086]
13. Xianxian Z, Lixia Z, Zhihui J, Yuanzun C. Application and effect evaluation of marital self-disclosure intervention on the fear of cancer recurrence in young breast cancer patients. Chin J Nurs 2019;54(11):1610-1614. [doi: 10.3761/j.issn.0254-1769.2019.11.002]
14. Niu L, Liang Y, Niu M. Factors influencing fear of cancer recurrence in patients with breast cancer: evidence from a survey in Yancheng, China. J Obstet Gynaecol Res 2019 Jul;45(7):1319-1327. [doi: 10.1111/jog.13978]
15. Yu Z, Sun D, Sun J. Social support and fear of cancer recurrence among Chinese breast cancer survivors: the mediation role of illness uncertainty. Front Psychol 2022;13:864129. [doi: 10.3389/fpsyg.2022.864129] [Medline: 35369168]
16. Jin LI, Na LU, Feng-wa ZH, et al. Influence factors of cancer recurrence fear in breast cancer patients: a meta-analysis. J Nurs 2021;28(14):1-7 [FREE Full text] [doi: 10.16460/j.issn1008-9969.2021.14.001]
17. Ni P, Chen J, Liu N. The sample size estimation in quantitative nursing research. Chin J Nurs 2010;45(4):378-380. [doi: 10.3761/j.issn.0254-1769.2010.04.037]
18. Mehnert A, Herschbach P, Berg P, Henrich G, Koch U. Fear of progression in breast cancer patients—validation of the short form of the Fear of Progression Questionnaire (FoP-Q-SF). Z Psychosom Med Psychother 2006;52(3):274-288. [doi: 10.13109/zptm.2006.52.3.274] [Medline: 17156600]
19. Wu QY, Ye ZX, Li L. [FREE Full text]
20. Kahn JH, Hessling RM. Measuring the tendency to conceal versus disclose psychological distress. J Soc Clin Psychol 2001 Mar;20(1):41-65. [doi: 10.1521/jscp.20.1.41.22254]
21. Li X. A study on the relationship between adult attachment, self-disclosure, and depression in college students [Master’s Thesis]. : Hebei Normal University; 2009.
22. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32(6):705-714. [doi: 10.1016/0277-9536(91)90150-b] [Medline: 2035047]
23. Li H. Psychometric measurement of the simplified chinese medical outcomes study social support survey (MOS-SSS-c) and its application in HIV/AIDS patients [Master’s Thesis]. : Central South University; 2012.
24. Han M, Chen H, Li J, et al. Correlation between symptom experience and fear of cancer recurrence in postoperative breast cancer patients undergoing chemotherapy in China: a cross-sectional study. PLoS One 2024;19(9):e0308907. [doi: 10.1371/journal.pone.0308907] [Medline: 39292722]
25. Li Y, Fang C, Xiong M, Hou H, Zhang Y, Zhang C. Exploring fear of cancer recurrence and related factors among breast cancer patients: a cross-sectional study. J Adv Nurs 2024 Jun;80(6):2403-2414. [doi: 10.1111/jan.16009] [Medline: 38041587]
26. Zhang HF, Lin ZH, Li J, Chen JJ, Wu WB, Cui ZT. Latent profile analysis of fear of cancer recurrence of patients with breast cancer. Chin Nurs Res 2022;22:4013-4017 [FREE Full text] [doi: 10.12102/j.issn.1009-6493.2022.22.012]
27. Mahendran R, Liu J, Kuparasundram S, et al. Fear of cancer recurrence among cancer survivors in Singapore. Singap Med J 2021 Jun;62(6):305-310. [doi: 10.11622/smedj.2020007]
28. Du X, Wang D, Du H, Zou Q, Jin Y. The correlation between intimate relationship, self-disclosure, and adaptability among colorectal cancer enterostomy patients. Medicine (Baltimore) 2021 May 14;100(19):e25904. [doi: 10.1097/MD.0000000000025904] [Medline: 34106651]

XSL•FO
RenderX

29. Hong SJ, Shin NM, Jung S. A predictive model of fear of cancer recurrence for patients undergoing chemotherapy. Support Care Cancer 2020 Sep;28(9):4173-4181. [doi: 10.1007/s00520-019-05245-7] [Medline: 31900611]

Abbreviations

FCR: fear of cancer recurrence
MOS-SSS-CM: Simplified Chinese Medical Outcomes Study Social Support Survey

Edited by M Balcarras; submitted 10.Dec.2025; peer-reviewed by P Sivaranjani, Z Ping; accepted 04.Mar.2026; published 10.Apr.2026.

Please cite as:
Ding Y, Zhang H, Zhong Q, Zhang L, Qin H
Fear of Cancer Recurrence Before Treatment in Patients Newly Diagnosed With Breast Cancer: Cross-Sectional Study
JMIR Cancer 2026;12:e89346
URL: https://cancer.jmir.org/2026/1/e89346
doi:10.2196/89346

© Yun Ding, Huiting Zhang, Qiaoling Zhong, Lijuan Zhang, Huiying Qin. Originally published in JMIR Cancer (https://cancer.jmir.org), 10.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Pathways and Roadblocks in Navigating Online Cancer Communities: Qualitative Study Among Young Adult Cancer Survivors

Qi Chen[1], MSW, PhD; Erin Donovan[2], PhD; Leyi Zhou[3], MSW; Lailea Noel[4], MSW, PhD; Barbara Jones[5], MSW, PhD

[1]Silberman School of Social Work, Hunter College, City University of New York, 2180 3rd Ave, Room 453, New York, NY, United States

[2]Department of Communication Studies, Moody College of Communication, The University of Texas at Austin, Austin, TX, United States

[3]School of Social Welfare, University of California, Berkeley, Berkeley, CA, United States

[4]Steve Hicks School of Social Work, The University of Texas at Austin, Austin, TX, United States

[5]School of Social Work, Boston University, Boston, MA, United States

Corresponding Author:
Qi Chen, MSW, PhD
Silberman School of Social Work, Hunter College, City University of New York, 2180 3rd Ave, Room 453, New York, NY, United States

Abstract

Background: Online cancer communities provide young adult (YA) cancer survivors with access to informational and emotional support that may not be available in traditional care settings. While these platforms offer vital connection opportunities, the unique pathways YA survivors take to find online communities and the challenges they encounter remain underexplored.

Objectives: This study aimed to (1) examine how YA cancer survivors locate and access online cancer communities and (2) identify barriers that impede their participation or sustained engagement.

Methods: The first author conducted semistructured interviews with 12 YA cancer survivors aged 18 to 39 years who had experience using online resources after their diagnosis. Participants were recruited through purposive and snowball sampling from YA cancer-focused nonprofit organizations and through social media. Interviews were conducted via Zoom and analyzed using thematic analysis. The analytic process followed Braun and Clarke's 6-phase framework and was supported by MAXQDA software. Codes and themes were generated inductively and refined iteratively.

Results: Participants ranged in age from 24 to 39 years (mean 32, SD 5.08 years), with an average of 3 (SD 2.98) years since diagnosis. Most identified as female (n=9, 75%) and non-Hispanic White (n=7, 58%). Over half held a graduate degree (n=7, 58%), half were employed full time (n=6, 50%), and most resided in suburban areas (n=8, 67%). Cancer diagnoses included leukemia (n=3, 25%), lymphoma (n=4, 33%), and other solid tumors such as testicular, colon, and uterine cancers. At the time of the interview, 3 (25%) participants were in active treatment and 9 (75%) had completed treatment. Participants described five primary pathways to discovering online cancer communities: (1) direct searching using hashtags or keywords, (2) community hubs on public accounts, (3) referrals from health providers or social networks, (4) algorithm-recommended content, and (5) connections formed within preexisting online interest-based groups. Despite the promise of digital tools, participants encountered five roadblocks: (1) platform fragmentation and digital literacy complicated initial discovery; (2) lack of representation made it difficult for some to find communities where they felt seen; (3) emotional overload and engagement fatigue, along with shifting group hierarchies and boundaries, further hindered sustained participation; and (5) lastly, concerns about cyberbullying discouraged open engagement, prompting some to withdraw or limit their presence in online communities.

Conclusions: YA cancer survivors navigated a fragmented and emotionally complex digital landscape in search of social support. Their ability to access and engage with online communities was shaped not only by individual agency and digital literacy but also by structural and relational factors. This study underscores the need for more discoverable, inclusive, and sustainable digital support environments. Oncology professionals and peer survivors can play a vital role in facilitating safe, informed access to online cancer communities. Multilevel psychoeducation and training with health care providers, YA cancer survivors, and online community facilitators are warranted to bridge gaps and enhance equity in digital survivorship care.

(JMIR Cancer 2026;12:e79893) doi:10.2196/79893

KEYWORDS

young adult; cancer survivors; online social networking; social support; health disparities

XSL•FO
RenderX

Introduction

Before the widespread adoption of communication technologies, patients having cancer and their families largely relied exclusively on oncology providers for support and information. The rise of social media and online platforms has transformed this landscape, enabling cancer survivors to connect, share experiences, and find vital health information from the comfort of their homes. A report by the Pew Research Center revealed that 11% of adults in the United States have been following their friends' health updates on online communities, and 5% have actively engaged by sharing their own health-related information, questions, or comments [1]. A national survey showed that the general population's online support group participation grew from 5% in 2008 to 9% in 2017 [2].

Online cancer communities refer to groups of people with expressed interest in cancer who communicate using a website, social media platform, instant messaging, or app [3]. These communities, often created by patients, nonprofits, or health communication professionals, serve as supportive spaces for patients, survivors, and caregivers to seek information and emotional connection. Over time, members contribute to an evolving body of shared knowledge, fostering a sense of belonging based on common diagnoses or life circumstances. Cancer survivors report positive attitudes toward these online communities and the social support they provided [4].

Young adult (YA) cancer survivors, a distinct group of cancer survivors defined by the National Cancer Institute [5], often face unique challenges: social and developmental disruptions, feelings of isolation, and limited access to in-person support groups tailored to their needs [6]. Because YAs are often geographically dispersed, online cancer communities provide accessible spaces to exchange health information, share experiences, and find emotional and informational support among peers who understand their situations [7-9]. These communities foster empowerment as both a process and an outcome of participation, complementing traditional health care by promoting agency, peer learning, and self-management [10-12]. Participation can help survivors reclaim a sense of connection, reduce isolation, and build hope for life beyond cancer [13,14].

However, these benefits are not experienced uniformly. Engagement in online cancer communities is shaped by both individual characteristics, such as education level, health literacy, psychological well-being, and personality traits, and community-level factors like dynamics, including reciprocity, inclusiveness, and the quality of moderation [9,15,16]. Understanding how these factors interact to shape YA survivors' participation remains critical to improving access to meaningful online support.

Finding an online community that meets one's emotional and informational needs can be challenging for YA cancer survivors. Although there are various types of online cancer communities, 2 broad classifications, open forums and secret groups, pose distinct challenges for YA cancer survivors seeking meaningful support [17]. Open forums, typically hosted by charities or nonprofits, provide accessible and reliable information but foster limited personal connection, while secret groups, often patient-initiated and invitation-only, offer greater intimacy and emotional support but raise privacy concerns and require trust-building [17]. The diversity of platforms and fragmented resources can make exploration overwhelming. Privacy concerns are also particularly pressing for YAs, who fear being identified by future employers or wider social circles, making them gravitate toward more private groups that require referral to join [18]. However, the intimacy of secret groups comes at the cost of additional risks, including loss of control over shared stories [19,20], emotional distress following other members' death [13,21], and exposure to misinformation or harmful content [22,23].

Taken together, these findings highlight the need to better understand how YA cancer survivors identify and engage with online health communities that align with their emotional and informational needs. Although prior studies have documented both the benefits and challenges of these spaces, little is known about how YAs navigate online resources, find suitable communities, and sustain meaningful participation. The purpose of this qualitative study is to understand (1) how YA cancer survivors locate and access their online cancer communities and (2) what hinders their ongoing engagement.

Methods

Design

This qualitative study used semistructured interviews to explore how YA cancer survivors find, access, and engage with online cancer communities. The study aimed to understand both the pathways through which participants discovered these communities and the barriers they encountered in participating or sustaining engagement over time.

Participants

Purposive and snowball sampling was used to recruit participants. To be eligible for the study, participants needed to (1) receive a diagnosis of cancer at the age between 18 and 39 years, (2) be between 18 and 39 years at the time of the interview, (3) have experiences using social media platforms, and (4) be able to participate in a videoconference interview. The age range was based on the National Cancer Institute's definition of adolescent and young adult (AYA) cancer survivors [5] and was selected to capture the relational and developmental contexts of young adulthood, a life stage marked by key social transitions and identity formation that shape survivorship experiences [6].

The research team developed a flyer that they asked nonprofit organizations dedicated to YA cancer to post on their social media accounts. Cactus Society and Livestrong Cancer Institute, 2 US-based organizations supporting AYA cancer survivors, shared the flyer on their social media accounts. The study flyer contained study aims, the principal investigator's contact information, and a QR code of intake survey on Qualtrics. The principal investigator (QC) followed up with each potential participant to schedule Zoom interviews and gather verbal informed consent.

XSL•FO
RenderX

Data Collection

All interviews followed the same interview guide with some variations based upon the flow of conversation between the participant and interviewer (see Multimedia Appendix 1 for the interview guide, including prompts and questions). The interview guide was developed by the research team. It contained questions related to the ways YA cancer survivors connected to peer survivors on social media, including positive and negative experiences during the process. The interviews were conducted in a one-on-one setting, free from the presence of any nonparticipants. Prior to each interview, participants were instructed to join from a private location and to ensure that no nonparticipants were present or able to overhear the conversation. Interviews were video recorded and transcribed by Zoom. Transcripts were made available for participant review upon request. The interviewer (QC) read through transcripts to ensure the quality of transcription and kept field notes that captured the main points of conversation after each interview.

Data Analysis

Reflexive thematic analysis is a method for identifying, analyzing, and reporting patterns within data, defined by Braun and Clarke [24]. All the transcripts were deidentified before importing into MAXQDA (VERBI Software, 2021), a qualitative software program for data coding and comparison.

Two members of the research team, QC and LZ, independently coded all transcripts using an inductive, consensus-driven approach. The data coding approach was inductive. Codes and theme development were directed by the content of the data, rather than being theoretically driven [24]. Both coders initially reviewed a subset of transcripts to develop a preliminary codebook, which was refined through iterative discussion. Each transcript was then coded separately by both coders, followed by comparison and discussion to reconcile differences and reach agreement on code application and interpretation. Discrepancies were resolved through dialogue until full consensus was achieved, ensuring analytic consistency and rigor.

Following the six-phase process of thematic analysis, QC and LZ began with a process of (1) *data familiarization*, in which QC noted initial analytic observations about each data item and the entire dataset by reading the transcription and field notes. Two transcriptions were coded to (2) *generate initial codes* and develop the first-version codebook with columns of code category, code name, definition, eligibility, and example based on the initial codes. QC and LZ coded the other 2 interviews and revised the codebook accordingly by adding new codes, refining code's eligibility criteria, and adding second-level codes to group the existing codes. After the codebook was finalized with all 12 interviews, the data were organized into a matrix with cases and second-level codes. QC and LZ (3) *searched for themes* by reading and identifying the similarity in data across cases in the matrix. The 2 authors (4) *reviewed the themes* and gave (5) *definitions and names to the themes*. All authors reached consensus regarding the themes' consistency. The last phase was to move from semantic description to interpret the significance of patterns and their broader meaning and implications for the (6) final report.

Rigor

Steps to ensure rigor were based on the trustworthiness criteria of data analysis posited by Lincoln and Guba [25]. The criteria involved establishing credibility, dependability, confirmability, and transferability. Throughout the research process, the first author reported the coding process and reflection of the transcription with the research team to assure the credibility of the analysis. The research team provided consultation and guidance on the naming of codes and themes to ensure the accuracy of language. The first author kept an audit trail (Koch [26]) to record each decision made with modifying the research question, finalizing the data analytic approach, and coding and analyzing the pattern to improve the dependability of the finding. Qualitative software was used, as it permits greater audibility and hence increases the reliability of the data analysis process.

Confirmability was sought by regular review and discussion of the codes with the research team by discussing reflexivity and how the interactive nature of the interview process influenced data collection. Member checking was not conducted, as this study employed a reflexive thematic analysis approach in which meaning is understood as coconstructed through researcher interpretation rather than verified by participants. Following Braun and Clarke's distinction, traditional member checking assumes the existence of a single, verifiable truth and seeks to "correct" researcher subjectivity—an assumption that sits conceptually at odds with reflexive thematic analysis, where researcher subjectivity is viewed as an analytic resource, not a bias to eliminate [27]. Consistent with the interpretivist assumptions of reflexive thematic analysis, data saturation was not used as a criterion for determining sample adequacy because Braun and Clarke [28] advocated for achieving meaning through researcher interpretation.

Ethical Considerations

Institutional review board approval was granted for this study by the University of Texas at Austin (protocol STUDY00003004). All study procedures complied with institutional and national research ethics standards for research involving human participants. Participants were informed of the study purpose, confidentiality protections, the benefits and risks of participation, and their rights to withdraw at any time. Prior to participation, all participants provided verbal informed consent to take part in the study and to be video recorded. Participation was voluntary, and participants were reminded that they could decline to answer any question or withdraw from the study at any time without penalty. All data were deidentified during transcription and analysis. Identifying information such as names or specific personal details was removed or replaced with pseudonyms. Digital recordings and transcripts were stored in a secure and access-restricted folder on UT Box, accessible only to the research team. Each participant received a US $20 Amazon electronic gift card, sent via email upon completion of the interview, as compensation for their time and contribution to the study.

Positionality

The research team represents a diverse group of scholars and practitioners whose social identities, disciplinary backgrounds,

XSL•FO
RenderX

and professional experiences shaped the design and interpretation of this study. QC (Asian woman, Assistant Professor of Social Work, and a PhD candidate at the time of the study), ED (White woman, Professor of Communication), LN (Black woman, Assistant Professor of Social Work), and BJ (White woman, Professor of Social Work) are seasoned qualitative researchers with a PhD degree and extensive experience studying psychosocial oncology, communication, and health equity among AYA cancer survivors. LZ (Asian woman, social work clinician, and doctoral student), the second coder of the transcript, brought both clinical practice experience and methodological expertise in qualitative research. The interviewer QC's interest in this topic stems from her professional experience in psychosocial oncology and her commitment to addressing disparities in access to cancer support as well as the rising use and impact of communicative technology during the COVID-19 pandemic. She acknowledges that her background as a social work researcher and her own positive experiences with online communities as an immigrant may have shaped assumptions about online support and digital disparities. As someone who is not a cancer survivor herself, she remains mindful of the potential distance between her professional perspective and participants' lived experiences. QC had limited prior relationships with participants. One participant was a personal friend of QC, while all others had no prior relationship with the interviewer. Participants were informed about QC and the study through the introductory phone calls or emails and the consent form, which described QC's role as the principal investigator and explained the study's purpose, goals, and procedures. No additional personal or professional information was shared beyond what was stated in the consent materials.

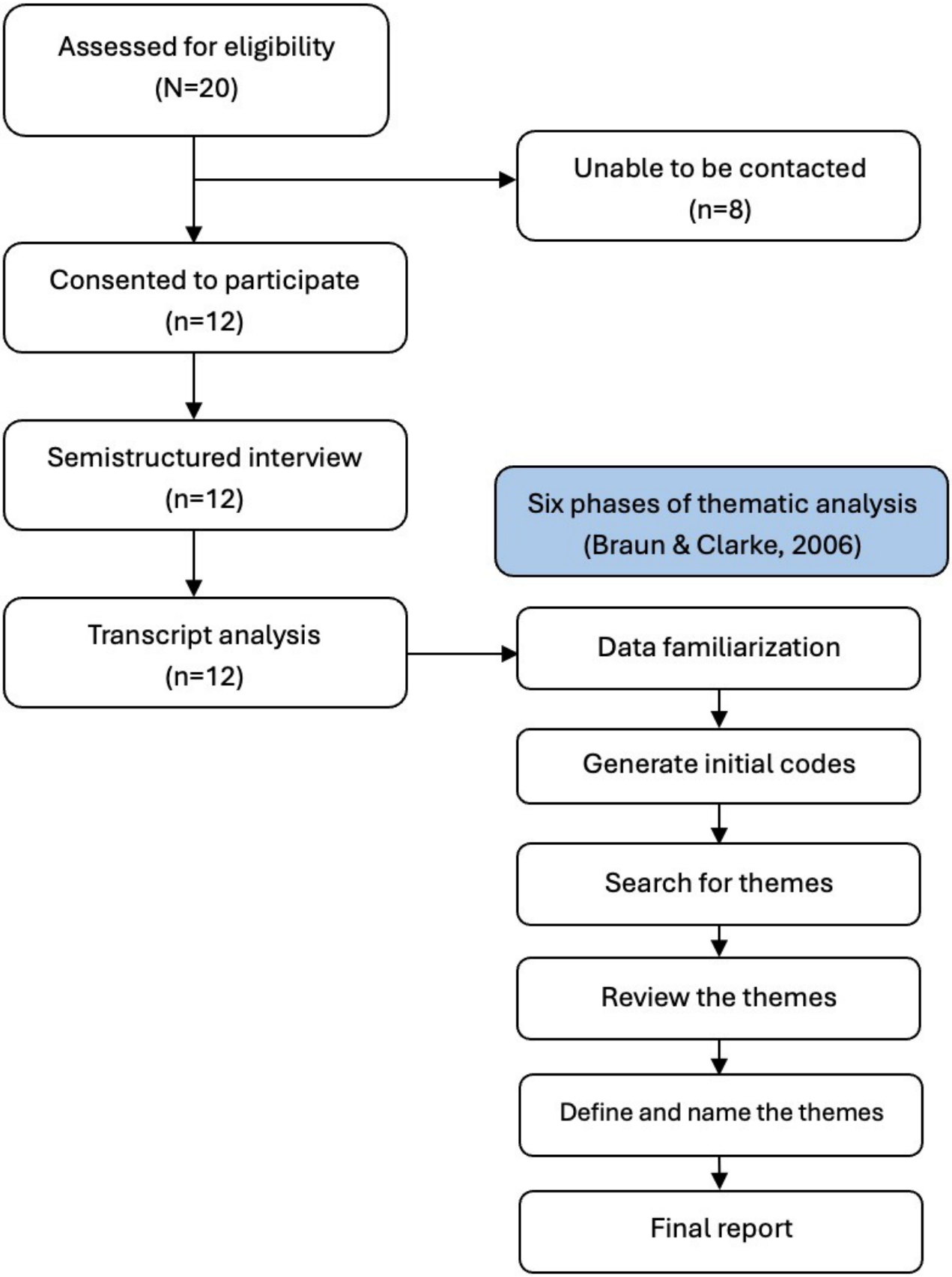
Results

Participants' Characteristics

A total of 20 potential YA participants who filled out the screening questionnaire were all eligible for the study. Of these 20 YAs, 8 did not respond to the principal investigator's emails or phone calls, and 12 consented and were recruited to participate in the interviews, yielding an enrollment rate of 60%. Figure 1 presents the recruitment and analysis process flowchart. Semistructured individual Zoom interview data were collected from September 2022 to February 2023, ranging from 39 to 110 minutes (mean 67.2 min, SD 21.3 min). No repeated interview was carried out, and no participants were requested to review or edit their transcripts. The demographic information of participants is presented in Table 1. Participants had a mean age of 32 (SD 5.08; range 24 - 39) years and an average of 3 (SD 2.98; range 0 - 11) years since diagnosis. Most participants were identified as female (n=9, 75%). The majority were non-Hispanic White (n=7, 58%), followed by Hispanic (n=2, 17%) and Asian (n=2, 17%), and 1 (8%) identified as multiracial. Over half (n=7, 58%) of participants held a graduate degree, while 4 (33%) had completed some college or a college degree, and 1 (8%) had completed high school. In terms of employment, 6 (50%) participants were employed full-time, 2 (17%) part-time, and 4 (33%) not employed. Participants resided in 8 different states. Most participants resided in suburban areas (n=8, 67%). Cancer diagnoses among participants varied, with leukemia being the most common (n=3, 25%), followed by Hodgkin's lymphoma (n=2, 17%), testicular cancer (n=2, 17%), and individual cases of non-Hodgkin's lymphoma, colon, brain, soft tissue, and uterine cancers (each n=1, 8%). At the time of the interview, 9 (75%) participants had completed treatment, while 3 (25%) were still in active treatment.

XSL•FO
RenderX

Figure 1. Participant recruitment and thematic analysis process flowchart [24].



XSL•FO
RenderX

Table . Participant demographics and background information (n=12).

Demographics	Count
Age, mean (SD; range)	32 (5.08; 24-39)
Years since diagnosis, mean (SD; range)	3 (2.98; 0 - 11)
Sex, n (%)	
Female	9 (75)
Male	3 (25)
Race, n (%)	
Non-Hispanic White	7 (58.3)
Hispanic	2 (16.7)
Asian	2 (16.7)
Multiracial	1 (8.3)
Education attainment, n (%)	
High school	1 (8.3)
Some college or college	4 (33.4)
Graduate school	7 (58.3)
Employment, n (%)	
Full time	6 (50)
Part time	2 (16.7)
Not employed	4 (33.3)
Location (State), n (%)	
Texas	3 (25)
New York	2 (16.7)
Illinois	2 (16.7)
Pennsylvania	1 (8.3)
Wisconsin	1 (8.3)
Florida	1 (8.3)
Maryland	1 (8.3)
Michigan	1 (8.3)
Urbanicity, n (%)	
Urban	4 (33.3)
Suburban	8 (66.7)
Cancer diagnosis, n (%)	
Leukemia	3 (25)
Hodgkin’s Lymphoma	2 (16.7)
Testicular cancer	2 (16.7)
Non-Hodgkin’s Lymphoma	1 (8.3)
Colon cancer	1 (8.3)
Brain cancer	1 (8.3)
Soft tissue cancer	1 (8.3)
Uterine cancer	1 (8.3)
Treatment status, n (%)	
Completed	9 (75)
In active treatment	3 (25)

XSL•FO
RenderX

Qualitative Findings

This study explored 2 central questions: (1) how YA cancer survivors locate and access online cancer communities and (2) what challenges they encounter in sustaining engagement. The findings revealed a range of entry points that survivors used to discover supportive online spaces, as well as structural, interpersonal, and emotional factors that limited their continued participation. While participants represented a range of ages within the YA category (24 - 39 y), their approaches to engaging with online cancer communities and the obstacles they experienced were broadly similar. No age-related patterns were observed in terms of search strategies, preferred platforms, or engagement levels.

The results are presented in 2 sections, corresponding to each research question. The first section, *Pathways*, included 5 themes describing how YA cancer survivors found online communities: (1) search, connect, and join/build, (2) community hubs, (3) third-party referrals, (4) algorithm recommendations, and (5) online interest-based groups. The second section, *Roadblocks*, outlined 5 themes that capture barriers to sustained engagement: (1) platform fragmentation and digital literacy, (2) group boundaries and hierarchies, (3) lack of representation, (4) emotional overload and engagement fatigue, and (5) cyberbullying. Figures 2 and 3 depicted 2 coding trees of Pathways and Roadblocks to illustrate the codes identified and themes emerged in the analysis. Table 2 presents each theme along with its definition and representative participant quotes.

Figure 2. Coding tree of "Pathways" illustrating how young adult cancer survivors discovered online cancer communities.

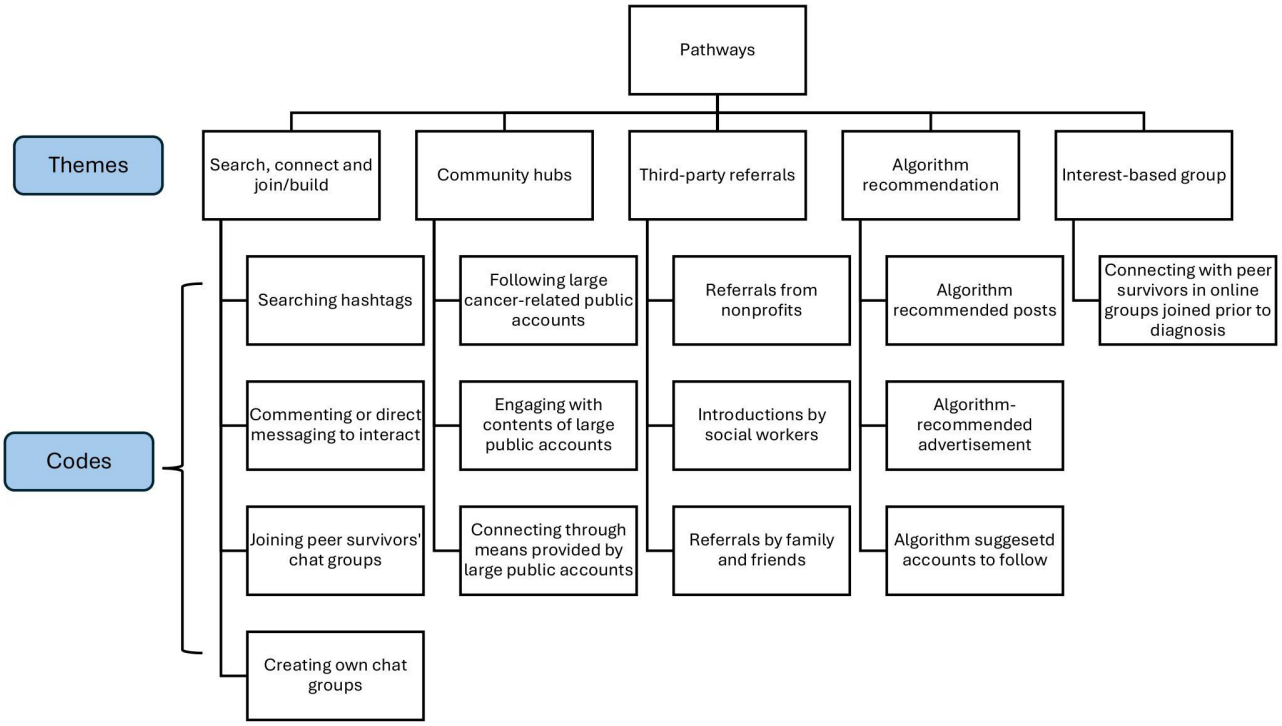
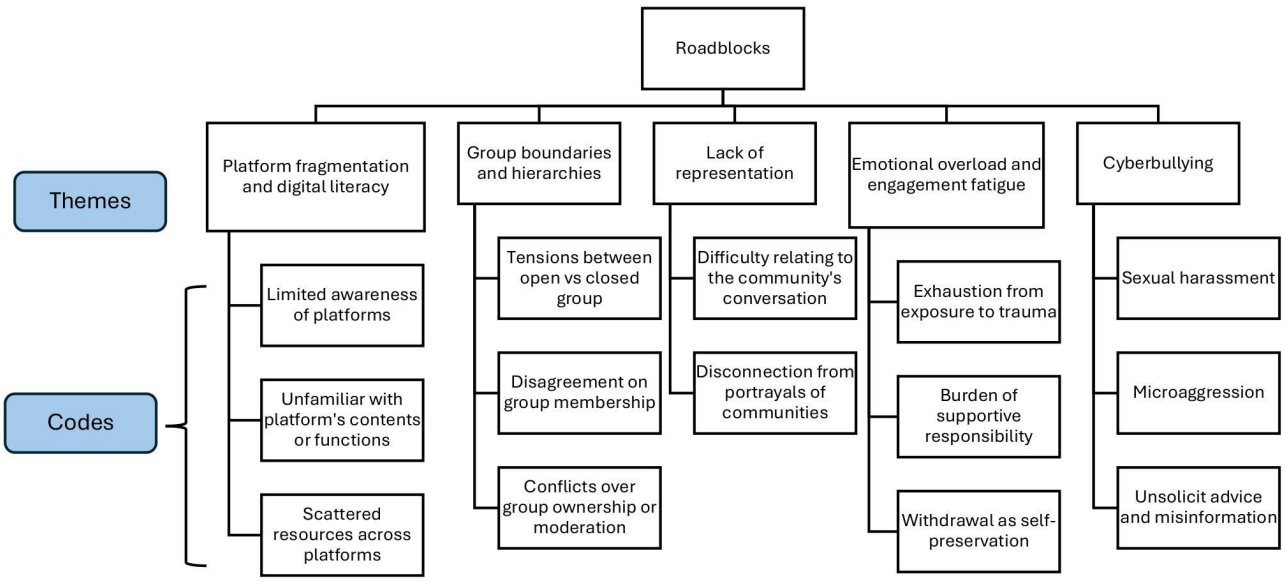


Figure 3. Coding tree of "Roadblocks" illustrating barriers that young adult cancer survivors encountered in sustaining engagement with online cancer communities.



XSL•FO
RenderX

Table . Themes, description, and representative quotes: pathways and roadblocks.

Themes	Description	Representative quote
Pathways		
Search, connect, and join/build	YA survivors actively searched relevant hashtags and initiated contact with peer survivors through comments or direct messages, leading to joining an existing community or forming their own community	• *She has a decent following, so I didn't expect her to see it or anything, but she responded back. So I messaged her, and then suddenly started messaging and then exchanged phone numbers and she's like, I've met a few other girls, who also going through cancer treatment this summer. So then she made a group chat for us. And then we said we all will support each other through cancer treatment this summer, when we should like having fun or whatever. That was like eight months ago, we still talk almost like every day.* (Participant 09) • *First of all, I was adding people that were in my state to try and build a community in my state, then it widened out more and more. And we would eventually share each other's Instagrams and that's how I built it up. Sometimes I'll still search up care packages, just to see if there's any new ones to add to my list, but also to see any people out there so I can connect with them.* (Participant 10)
Community hubs	YA cancer survivors followed large public accounts serving as information hubs for peer interaction.	• *The person who runs the account made a spreadsheet where people could put their name, their age, their diagnosis, and where they lived and the Instagram handle. And then, if anybody wanted to talk to them they could reach out. So I put my information on that spreadsheet, and I had one person reach out*. (Participant 08)
Third-party referrals	Nonprofit organizations or health care providers referred YA cancer survivors to existing online groups.	• *The other organization does a weekly happy hour every Friday, and that is the same host every time, and I started going and I found out from people who go to that also go to Stupid Cancers zoom session. And so, there's a lot of cross involvement with different survivors are part of and take part in programming for different AYA groups. Some of my closest friends that I've connected with there, I have almost all of them on Instagram and/or Facebook. I talk with them on an Instagram group Chat* (Participant 06).
Algorithm recommendations	YA cancer survivors encountered groups through social media algorithm suggestions.	• *When I search for something on Google, Facebook picks up my information and pushes related content… I can even see on Facebook what I had just searched for on Google.* (Participant 01) • *That was through algorithm. I didn't search for anything during that time. I think I just got lucky that one day, and she already knew people. That was like pure luck, like it was meant to happen. I don't think I find that many people on social media share similar experiences*. (Participant 09)
Online interest-based groups	YA cancer survivors found peer support within preexisting online interest or hobby groups.	• *It wasn't a cancer group, but people there had similar ages and life experiences, so I felt connected.* (Participant 07)

XSL•FO
RenderX

Themes	Description	Representative quote
Roadblocks		
Platform fragmentation and digital literacy	YA cancer survivors struggled to locate reliable supportive communities due to scattered platforms and lack of digital literacy.	• *I didn't really use Facebook and Instagram much so I was not aware of those accounts. I used to only use Chinese-language platforms and those that are available for Chinese people. I like hanging out with other Chinese people more*. (Participant 01)
Group boundaries and hierarchies	Conflicts over group ownership, membership criteria, or tone led to disengagement.	• *Sometimes caregivers come in, and they'll tell us 'Oh, how inspirational you are.' And we just don't know how to take that because we're like, we're just trying to live our life or don't feel inspirational*. (Participant 10) • *It fell apart because there was a disagreement over who was old enough to be in the group, and that hurt some people's feelings… and another group (that was dismissed), it turned out one person was making stuff up and people just didn't want to be a part of the group*. (Participant 06)
Lack of representation	Limited diversity (age, diagnosis, culture) left survivors feeling disconnected or unseen.	• *I couldn't really find a relevant account about infertility that felt like it described my experience. I followed some infertility-related accounts that weren't cancer specific, but just about people who had infertility in general, and those accounts are mostly run by white women who are really religious and really Christian. And so I think I've decided not to engage with those accounts anymore, just because I couldn't find anybody who had experiences that were similar to me or identities that were similar to mine.* (Participant 08)
Emotional overload and engagement fatigue	Exposure to others' trauma and constant emotional demands led to burnout and withdrawal.	• *I already felt miserable and no one cared about me... but seeing them vent still made me feel awful… On top of my own problems, I have to carry others' misery… thinking about their fate makes it unbearable.* (Participant 01) • *Sometimes it can feel a little repetitive. Like there are certain types of posts that I feel like you see again and again and again. So I think that can make it hard to stay engaged.* (Participant 03)
Cyberbullying	Harassment or microaggressions led survivors to privatize accounts, reducing visibility and engagement.	• *There's a subsection of people who are really weird that like bold girls. They will message cancer patients with really graphic descriptions. I've got that kind of message*. (Participant 10)

Pathways

YA cancer survivors accessed online communities through a range of entry points that reflected both deliberate strategies and serendipitous discovery.

Search, Connect, and Join/Build

This was the most common way YA cancer survivors mentioned regarding finding and joining their online cancer community. Usually, it would start with searching for keywords related to their cancer diagnosis or type of treatment, or the hashtag #AYAcancer, #YoungAdultCancer, #AYAcancer, or #cancerinyour20s, to identify peer YA survivors who shared similar cancer experiences. Some YA cancer survivors would initiate the conversation by commenting on the post of accounts they identified, such as sending wishes and making comments on similar cancer-related situations they share, and some would directly message to connect. After the connection was established and a larger network based on similar cancer experiences was formed, sometimes a chat group/channel would be created by one of the group members. There were also some

XSL•FO
RenderX

other YA cancer survivors taking a leading role in creating a community for people they knew and actively searching out to YA cancer survivors who might be looking for support on social media.

Community Hubs

Community hubs were centralized online spaces managed by individuals or small groups, serving as anchors or gatekeepers for connection, information exchange, and peer support. We found that there were such hubs that existed for YA cancer survivors to connect: almost all of the participants followed the accounts @thecancerpatient and @NoGriefTouristAllowed on Instagram. These accounts shared similar traits: large number of followers and mostly YA cancer survivors. They served as an information hub for YA cancer survivors. A YA cancer survivor community was naturally formed based on these 2 accounts, and these accounts were acknowledged and advertised by these cancer survivors at platforms other than Instagram. These public accounts connected peer survivors through posts that reflected shared cancer experiences. People responded in the comments or initiated conversations or through "story" (posts visible for 24 h) featuring followers' requests to connect. Some accounts would create ways to collect the contact information of those who want to connect.

Third-Party Referrals

YA cancer survivors also found nonprofit organizations or social workers at the hospital or cancer center helpful in referring to reliable online support resources. Nonprofit organizations dedicated to adolescent and YA cancer care, such as Stupid Cancer, Elephants and Tea, and Cactus Cancer Society, usually host online support groups and have their own forum or social media accounts. YA cancer survivors described being connected across different organizations, participating in multiple online cancer communities, and then forming their own communities in other social media platforms. YA cancer survivors also shared that when their family and friends found related resources, they would also recommend them and encourage them to connect. For instance, one participant mentioned that her sister has been an active user on the Reddit Cancer Caregiving subreddit and recommended her to join a cancer-related Reddit channel to find the community.

Algorithm Recommendation

As algorithm recommendation technology continued to advance, YA cancer survivors also benefited from recommendations on websites and apps they visited. Some described seeing the advertisement of Leukemia Society on Facebook after multiple rounds of internet search about Leukemia and survival rates and ended up visiting the site and found peer support programs. Some described their experiences of being recommended social media personal accounts with contents that were relevant to their life experiences and being able to connect with those people.

Online Interest-Based Groups

YA cancer survivors were comfortable with communication technologies, and they browsed different online communities at an early age. Therefore, many of them were already part of some online communities before they got cancer. One participant described her experience being in one huge interest group community, and there was a subgroup with people diagnosed with different types of cancer and stages of cancer. Even though it was not specifically for YAs, the participant described that since these interest-based groups were usually formed with people of similar age, demographic, and political outlook, or shared hobbies, it was easier to feel connected to these people.

Roadblocks

While online communities offered vital spaces of support, YA cancer survivors reported a number of challenges that shaped their participation.

Platform Fragmentation and Digital Literacy

Participants described difficulties locating relevant support communities due to the fragmented nature of social media platforms. Information and resources were scattered across platforms such as Instagram, Facebook, Reddit, and organization-specific forums, many of which lacked centralized searchability. Some participants were unfamiliar with widely used platforms in the United States, particularly immigrant YA cancer survivors who previously used regional platforms to maintain relationships with family and friends back in their home countries. This limited exposure made it challenging to access content or locate established YA cancer communities on English social media platforms.

Others noted that even when they were familiar with a platform, they often lacked the knowledge of which accounts or groups to follow. Many were also unaware of nonsocial media–based online communities. For instance, one participant shared her experience of only recently learning about COLONTOWN, a web-based platform uniquely for colorectal cancer survivors. Furthermore, many shared that YAs who were connected through one platform might move to another to create their own group, sometimes choosing platforms others had never used or did not have accounts on. These access disparities underscored how platform fragmentation and varying levels of digital literacy could hinder survivors' ability to locate and engage with appropriate sources of support.

Group Boundaries and Hierarchies

Online community boundaries could be challenging to manage. In open forums, lacking group boundaries would trigger uncomfortableness and conflicts and thus decrease their interests in staying in the group. Participants stated that for some groups that also welcomed the participation of caregivers, family members, or providers would interrupt the group dynamics and disrupt the sense of safety and belonging of the community. In secret/closed groups, group hierarchy, group ownership, and group membership could also become problematic. If agreement could not be achieved between group owner and group members, it would cause conflict, and the online group would have to be dismissed if conflicts continue to escalate.

Lack of Representation

Not all online cancer communities effectively met the informational and emotional support needs of YA cancer survivors. Several participants in the study described feeling like "outsiders" in these spaces. A major barrier to meaningful

XSL•FO
RenderX

engagement was the perceived lack of inclusiveness and representation. Participants emphasized that the characteristics of community members, such as age, diagnosis, religion, health beliefs, cultural background, and treatment experience, strongly influenced their sense of connection. When these attributes did not align, it was difficult to relate to others or feel understood. For example, wide variations in treatment plans and side effects among YA patients having cancer made it challenging to find peers with similar experiences. The more disconnected participants felt from the community, the less likely they were to engage.

Some YA cancer survivors expressed discomfort with content on accounts that were community hubs and overly romanticized the cancer experience. They found such narratives to be superficial and disconnected from their own realities, making it harder to relate or feel represented. One participant shared that she avoided public accounts that presented highly curated or idealized portrayals of survivorship, as they felt inauthentic and failed to meet her informational or emotional needs.

Emotional Overload and Engagement Fatigue

For many YA cancer survivors, ongoing exposure to peers' illness narratives in online communities could generate a deep sense of emotional fatigue, like one participant said: "Everyone's unloading their trauma and it's overwhelming." (Participant 05). This exhaustion stemmed from not only the content's intensity but also the complex dynamics of shared vulnerability and responsibilities. This form of emotional overload often coexisted with a sense of obligation to support, which heightened the burden. In addition to this emotional overload, participants described seeing similar posts over time could become draining. These revealed not a loss of interest but the emotional cost of showing up, where survivors must balance their need for support with the strain of witnessing others' ongoing pain. Self-preservation, through withdrawal or silence, became a necessary response for survivors negotiating their own healing within communal spaces of vulnerability.

Cyberbullying

For people to be able to search one's posts or follow one's account, one needs to keep their account and posts public. However, being public and exposing themselves to the internet could be risky, especially for YA survivors. Many participants reported microaggression and cyberbullying on social media, such as unfriendly comments on their appearance, unsolicited advice, or even sexual harassment. As a result, YA cancer survivors who keep their accounts public to share their cancer experiences and help peer survivors connect initially would change their account setting to private to avoid these negative impacts. The unsafe environment on social media would limit the information source that peer YA cancer survivors could access.

Discussion

Overview

This study provided timely insights into how YA cancer survivors navigated the increasingly complex landscape of online cancer communities. It identified 5 key pathways to discover online communities and 5 major barriers in finding and sustaining engagement. The findings revealed that survivors' participation was influenced not only by personal agency and constraints but also by systemic factors such as platform accessibility and usability. The findings also further illuminated understudied interpersonal dynamics, such as emotional labor and shifting hierarchies and boundaries within communities, that affected survivors' sustained involvement in online communities. By centering YA cancer survivors' lived experiences in exploring online communities, this research advanced a more nuanced understanding of how to foster inclusive, meaningful, and sustainable online support communities.

Principal Findings

This study uncovers new forms of digital space navigation and structural inequity that shape survivors' ability to benefit from online support. Our findings showed that with more online platforms available, YA survivors joined or formed communities across different platforms and often switched between them. While new features such as algorithmic recommendations, hashtags, stories, and group chat channels allowed them to find or form online cancer communities more easily, the benefits of these tools were not experienced equitably. Participants' capacity to navigate and engage effectively was strongly influenced by their familiarity with functions of different platforms, many of which were not accessible to all users. For instance, immigrant survivors originally from countries where platforms like Facebook and Instagram are banned often lacked exposure to dominant US platforms. It demanded additional effort to learn and transition to a different platform solely for the purpose of accessing online support. Furthermore, limited access to certain social media platforms among users from specific cultural backgrounds results in fewer online support communities offering inclusive and culturally relevant content in their native languages. This disproportionately affected YA of color. These gaps in representation were compounded by algorithmic patterns that reinforced misinformation and unsolicited advice, disproportionately affecting those with lower health literacy or limited English proficiency. Together, these findings reveal how structural digital inequities shaped YA survivors' capacity to benefit from online support.

This study reveals that YA survivors' engagement with online cancer communities was emotionally dynamic and relationally fluid. Their interactions were marked with uncertainty and change. Navigating online cancer communities involved ongoing assessment of the relevance of the communities, negotiation of personal or group boundaries, and weighing between benefits and the risks. Some participants moved fluidly between communities or platforms as new communities emerged and migrated, while others withdrew temporarily or permanently when the emotional toll became too high. Rather than serving as a stable or consistently supportive resource, online cancer communities could also generate emotional distress, uncertainty, and exposure to unsafe or overwhelming virtual environments.

Comparison to Prior Work

This study extends prior digital health disparity research on YA survivors' online support by shifting attention from individual

XSL•FO
RenderX

constraints in digital literacy to structural inequities in virtual spaces. Prior research has largely emphasized individual-level factors such as education, socioeconomic status, or their social capital as the main drivers of digital divide in access and utilization of online resources [29,30]. Our study reveals that digital disparities persisted even among digitally savvy YA survivors, driven by the nature of online platforms. Participants' platform preferences were shaped by the habits as well as the accessibility of apps and websites of their families and social networks [30]. One prior study has similarly identified representation barriers, which noted that YA survivors often avoided older platforms such as Facebook, where most established cancer support groups exist, due to perceived misalignment with their peer networks and identity [13]. Familiarity with specific platforms, access to culturally relevant content, and the design of digital infrastructures emerged as key determinants equally impacted YA. This shift from individual to systemic explanations underscores the importance of improving platform accessibility and usability.

This study also advances existing literature by revealing how interpersonal dynamics within online communities could introduce more emotional distress and uncertainties besides support. Consistent with the prior findings, our study found that YA cancer survivors had strong needs of connecting to peer survivors [31] and they would search online proactively to find peer survivors [32]. Our study pointed out that, however, engagement was not always self-initiated; algorithms or others' suggestions often prompted participation, leading to interpersonal strain. Prior work has well documented the benefits for YA cancer survivors involving online cancer communities, emphasizing information exchange, social support, and empowerment [31,33,34]. Risks such as exposure to distressing content or misinformation have also been acknowledged, though often discussed as isolated, individual experiences [35]. One prior study has examined the dynamics of social media engagement among YA cancer survivors, but focusing primarily on individual motivation and preferences [32]. Our findings expand these understandings by showing that group-level factors such as changing boundaries, evolving hierarchies, and the associated emotional labor directly influence engagement and withdrawal. By framing these experiences as relational and evolving, this study highlights that community organizing and group moderation is necessary for virtual spaces.

Limitations

This study has several limitations. First, our sample included only participants who had internet access and prior experience using social media. It remained unclear whether YA cancer survivors without such access or experience would be open to, or able to engage with, online cancer communities. This inclusion criterion was necessary to ensure that participants could reflect meaningfully on their experiences navigating online communities. However, it may have limited the perspectives of YA cancer survivors with limited digital access or literacy or who lack interests in online cancer communities. Future studies should consider various recruitment strategies (eg, phone interviews or community-based outreach) to capture the experiences of survivors who have not been involved with any online cancer communities.

Second, the sample lacked representation from African American participants. Prior research indicated that African American cancer survivors were underrepresented in online support groups and less likely to use the internet for health information compared to their White counterparts, due in part to digital access barriers, differing help-seeking preferences, and lower trust in online health sources [36,37]. Although we recruited through multiple nonprofit networks serving diverse populations and encouraged participant referrals, our sample may have unintentionally excluded survivors not affiliated with such organizations or those relying on offline or culturally specific support systems. Future research should employ more inclusive recruitment strategies, including outreach through hospital registries, partnerships with African American survivor networks or community centers, and targeted social media campaigns or accounts that have higher number of African American followers. Addressing these structural and digital inequities is essential to advancing equity and representation in digital cancer survivorship research.

Future Directions

Future intervention research and clinical practice should focus on improving the accessibility and sustainability of online cancer communities through both structural support and individual capacity-building.

At the system level, oncology social workers have long played a critical role in connecting patients with fragmented resources [38], and our findings further confirmed this essential function. However, many social workers, particularly those outside of large cancer centers, were not well-versed in the availability of online cancer communities and virtual support groups for minority YA survivors on various scattered platforms [31]. While social support from family, peers, and health care providers can facilitate access to online communities, the active involvement of health care professionals or trained moderators in online health communities enhances trust, reduces the spread of misinformation, and mitigates risks such as cyberbullying and communication breakdowns [39,40]. Expert-led moderation has been associated with improved emotional safety, greater engagement, and more supportive peer interactions [41]. This underscores the importance of digital literacy training with health care professionals to strengthen their competencies in navigating and moderating different online cancer communities, so they can more effectively refer, prepare, and support YA survivors.

At the individual level, private or peer-led groups typically rely on YA survivors themselves to establish norms and maintain boundaries, placing added emotional and relational burdens on those already coping with cancer. Our findings indicated that many YA survivors not only passively join a group but actively take part in building or contributing to their own communities. One promising approach is to train YA cancer survivors who are interested in leadership roles in online communities on skills to identify harmful content, establish group norms, and manage conflict. Additionally, while the prevalence of cyberbullying and emotional risk in online communities has been documented [42], few interventions exist to prepare survivors for these challenges. Education around safe storytelling, self-care, and

XSL•FO
RenderX

managing online engagement fatigue should be integrated into survivorship support programming [35]. This includes guidance on when to disengage, how to set personal boundaries, and how to respond to toxic interactions.

Conclusion

This study highlights that YA cancer survivors navigate online communities through multiple pathways but face significant barriers to sustained engagement, highlighting the need for provider training and peer moderator support. To foster safer and more inclusive online environments, health care providers should be trained to guide survivors toward reliable support resources; community moderators should be trained to manage emotional labor and maintain healthy group boundaries; and YA cancer survivors should receive survivorship support to navigate online spaces effectively. Strengthening these multilevel supports can enhance the sustainability and psychosocial benefits of online cancer communities for YAs.

Funding

QC was supported by the American Cancer Society Social Work Doctoral Training Grant (134807-DSW-20-086-01-SW).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the sensitive and potentially identifiable nature of qualitative interview data and the absence of participant consent for open sharing, but are available from the corresponding author on reasonable request.

Authors' Contributions

QC contributed to conceptualization, data collection, project administration, data analysis, and original draft writing. ED was responsible for conceptualization, methodology, review, and editing of the manuscript. LZ contributed to data analysis. LN contributed to methodology and review and editing. BJ was involved in conceptualization, supervision, and funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 22 KB - cancer_v12i1e79893_app1.docx]

Checklist 1
COREQ checklist.
[PDF File, 381 KB - cancer_v12i1e79893_app2.pdf]

References

1. Fox S. The social life of health information. : Pew Research Center; 2014 URL: https://www.pewresearch.org/short-reads/2014/01/15/the-social-life-of-health-information/ [accessed 2025-12-23]
2. Chou WYS, Gaysynsky A, Trivedi N, Vanderpool RC. Using social media for health: national data from HINTS 2019. J Health Commun 2021 Mar 4;26(3):184-193. [doi: 10.1080/10810730.2021.1903627] [Medline: 33856286]
3. Gupta T, Schapira L. Online communities as sources of peer support for people living with cancer: a Commentary. J Oncol Pract 2018 Oct 18;14(12):JOP1800261. [doi: 10.1200/JOP.18.00261] [Medline: 30335558]
4. Jansen F, van Uden-Kraan CF, van Zwieten V, Witte BI, Verdonck-de Leeuw IM. Cancer survivors' perceived need for supportive care and their attitude towards self-management and eHealth. Support Care Cancer 2015 Jun;23(6):1679-1688. [doi: 10.1007/s00520-014-2514-7] [Medline: 25424520]
5. Adolescent & Young Adult Health Outcomes & Patient Experience Study (AYA HOPE). National Cancer Institute (NCI). 2025. URL: https://healthcaredelivery.cancer.gov/aya/ [accessed 2020-09-15]
6. Zebrack BJ, Block R, Hayes-Lattin B, et al. Psychosocial service use and unmet need among recently diagnosed adolescent and young adult cancer patients. Cancer 2013 Jan 1;119(1):201-214. [doi: 10.1002/cncr.27713] [Medline: 22744865]
7. Sansom-Daly UM, Wakefield CE, Ellis SJ, et al. Online, group-based psychological support for adolescent and young adult cancer survivors: results from the recapture life randomized trial. Cancers (Basel) 2021 May 18;13(10):2460. [doi: 10.3390/cancers13102460] [Medline: 34070134]
8. Viola A, Panigrahi G, Devine KA. Digital interventions for adolescent and young adult cancer survivors. Curr Opin Support Palliat Care 2020 Mar;14(1):51-59. [doi: 10.1097/SPC.0000000000000480] [Medline: 31842020]
9. Vlooswijk C, Janssen SHM, Sleeman SHE, et al. Identifying the informational needs and sources of support of adolescent and young adult (AYA) cancer survivors to inform the development of a digital platform. J Cancer Surviv 2024 Oct 18:1-21. [doi: 10.1007/s11764-024-01679-z] [Medline: 39424710]

XSL•FO
RenderX

10. Johansson V, Islind AS, Lindroth T, Angenete E, Gellerstedt M. Online communities as a driver for patient empowerment: systematic review. J Med Internet Res 2021 Feb 9;23(2):e19910. [doi: 10.2196/19910] [Medline: 33560233]
11. Verberne S, Batenburg A, Sanders R, van Eenbergen M, Das E, Lambooij MS. Analyzing empowerment processes among cancer patients in an online community: a text mining approach. JMIR Cancer 2019 Apr 17;5(1):e9887. [doi: 10.2196/cancer.9887] [Medline: 30994468]
12. Frost J, Massagli M. PatientsLikeMe the case for a data-centered patient community and how ALS patients use the community to inform treatment decisions and manage pulmonary health. Chron Respir Dis 2009;6(4):225-229. [doi: 10.1177/1479972309348655] [Medline: 19858352]
13. Lazard AJ, Collins MKR, Hedrick A, et al. Using social media for peer-to-peer cancer support: interviews with young adults with cancer. JMIR Cancer 2021 Sep 2;7(3):e28234. [doi: 10.2196/28234] [Medline: 34473063]
14. Daniels SR, Yang CC, Toohey SJ, Willard VW. Perspectives on social media from adolescents and young adults with cancer. J Pediatr Oncol Nurs 2021;38(4):225-232. [doi: 10.1177/1043454221992319] [Medline: 33729902]
15. Badreddine BM, Blount Y. Understanding influential factors behind lurking behaviour in online cancer communities. Behav Inf Technol 2021 Apr 26;40(6):542-564. [doi: 10.1080/0144929X.2019.1709545]
16. Park KA, Eum SY, Oh H, et al. Factors affecting online health community participation behavior in patients with thyroid cancer. PLoS ONE 2020;15(6):e0235056. [doi: 10.1371/journal.pone.0235056] [Medline: 32579575]
17. Harkin LJ, Beaver K, Dey P, Choong KA. Secret groups and open forums: defining online support communities from the perspective of people affected by cancer. Digit Health 2020;6:2055207619898993. [doi: 10.1177/2055207619898993] [Medline: 32010450]
18. Clerici CA, Quarello P, Bergadano A, et al. Proper use of social media by health operators in the pediatric oncohematological setting: consensus statement from the Italian Pediatric Hematology and Oncology Association (AIEOP). Pediatr Blood Cancer 2018 May;65(5):e26958. [doi: 10.1002/pbc.26958] [Medline: 29316150]
19. Easley J. Motivations for cancer history disclosure among young adult cancer survivors. J Cancer Surviv 2019 Jun;13(3):447-458. [doi: 10.1007/s11764-019-00766-w] [Medline: 31102132]
20. Pini S, Hugh-Jones S, Shearsmith L, Gardner P. “What are you crying for? I don’t even know you” - the experiences of teenagers communicating with their peers when returning to school. Eur J Oncol Nurs 2019 Apr;39:28-34. [doi: 10.1016/j.ejon.2018.12.010] [Medline: 30850135]
21. Harkin LJ, Beaver K, Dey P, Choong K. Navigating cancer using online communities: a grounded theory of survivor and family experiences. J Cancer Surviv 2017 Dec;11(6):658-669. [doi: 10.1007/s11764-017-0616-1] [Medline: 28470506]
22. DiFonzo N, Robinson NM, Suls JM, Rini C. Rumors about cancer: content, sources, coping, transmission, and belief. J Health Commun 2012;17(9):1099-1115. [doi: 10.1080/10810730.2012.665417] [Medline: 22724591]
23. Hughes B, Joshi I, Wareham J. Health 2.0 and Medicine 2.0: tensions and controversies in the field. J Med Internet Res 2008 Aug 6;10(3):e23. [doi: 10.2196/jmir.1056] [Medline: 18682374]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. Lincoln YS, Guba EG. Naturalistic Inquiry: Sage Publications, Inc; 1985.
26. Koch T. Establishing rigour in qualitative research: the decision trail. J Adv Nurs 1994 May;19(5):976-986. [doi: 10.1111/j.1365-2648.1994.tb01177.x] [Medline: 8056928]
27. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
28. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport Exerc Health 2021 Mar 4;13(2):201-216. [doi: 10.1080/2159676X.2019.1704846]
29. Zhang Y, Xu P, Sun Q, Baral S, Xi L, Wang D. Factors influencing the e-Health literacy in cancer patients: a systematic review. J Cancer Surviv 2023 Apr;17(2):425-440. [doi: 10.1007/s11764-022-01260-6] [Medline: 36190672]
30. Mohsin FM, Ali SH, Chong SK, Parikh RS, DiClemente RJ, Hu L. Social media utilization within Asian American families and its role in healthy lifestyle behavioral influence: results from a nationwide survey. Social Media + Society 2023 Jul;9(3):20563051231196544. [doi: 10.1177/20563051231196544]
31. Kent EE, Smith AW, Keegan THM, et al. Talking about cancer and meeting peer survivors: social information needs of adolescents and young adults diagnosed with cancer. J Adolesc Young Adult Oncol 2013 Jun;2(2):44-52. [doi: 10.1089/jayao.2012.0029] [Medline: 23781400]
32. Lazard AJ, Collins MKR, Hedrick A, et al. Initiation and changes in use of social media for peer support among young adult cancer patients and survivors. Psychooncology 2021 Nov;30(11):1859-1865. [doi: 10.1002/pon.5758] [Medline: 34165848]
33. Zebrack B. Information and service needs for young adult cancer patients. Support Care Cancer 2008 Dec;16(12):1353-1360. [doi: 10.1007/s00520-008-0435-z] [Medline: 18386075]
34. Lazard AJ, Meernik C, Collins MKR, et al. Social media use for cancer support among young adults with cancer. J Adolesc Young Adult Oncol 2023 Oct;12(5):674-684. [doi: 10.1089/jayao.2023.0025] [Medline: 37257189]

XSL•FO
RenderX

35. De Clercq E, Rost M, Gumy-Pause F, Diesch T, Espelli V, Elger BS. Moving beyond the friend-foe myth: a scoping review of the use of social media in adolescent and young adult oncology. J Adolesc Young Adult Oncol 2020 Oct;9(5):561-571. [doi: 10.1089/jayao.2019.0168] [Medline: 32397793]
36. Fogel J, Ribisl KM, Morgan PD, Humphreys K, Lyons EJ. Underrepresentation of African Americans in online cancer support groups. J Natl Med Assoc 2008 Jun;100(6):705-712. [doi: 10.1016/s0027-9684(15)31346-8] [Medline: 18595573]
37. Fareed N, Swoboda CM, Jonnalagadda P, Huerta TR. Persistent digital divide in health-related internet use among cancer survivors: findings from the Health Information National Trends Survey, 2003-2018. J Cancer Surviv 2021 Feb;15(1):87-98. [doi: 10.1007/s11764-020-00913-8] [Medline: 32671557]
38. Zebrack B, Kayser K, Padgett L, et al. Institutional capacity to provide psychosocial oncology support services: a report from the Association of Oncology Social Work. Cancer 2016 Jun 15;122(12):1937-1945. [doi: 10.1002/cncr.30016] [Medline: 27070342]
39. Wadden D, August T, Li Q, Althoff T. The effect of moderation on online mental health conversations. arXiv. Preprint posted online on May 19, 2020. [doi: 10.48550/arXiv.2005.09225]
40. Huh J, Marmor R, Jiang X. Lessons learned for online health community moderator roles: a mixed-methods study of moderators resigning from WebMD communities. J Med Internet Res 2016 Sep 8;18(9):e247. [doi: 10.2196/jmir.6331] [Medline: 27608721]
41. Bizzotto N, de Bruijn GJ, Schulz PJ. Buffering against exposure to mental health misinformation in online communities on Facebook: the interplay of depression literacy and expert moderation. BMC Public Health 2023 Aug 18;23(1):1577. [doi: 10.1186/s12889-023-16404-1] [Medline: 37596592]
42. Reuman H, Kerr K, Sidani J, et al. Living in an online world: social media experiences of adolescents and young adults with cancer. Pediatr Blood Cancer 2022 Jun;69(6):e29666. [doi: 10.1002/pbc.29666] [Medline: 35293691]

Abbreviations

AYA: adolescent and young adult
YA: young adult

Edited by N Cahill; submitted 30.Jun.2025; peer-reviewed by C Walsh, D Rodrigulez; revised version received 09.Nov.2025; accepted 25.Nov.2025; published 12.Jan.2026.

Please cite as:
Chen Q, Donovan E, Zhou L, Noel L, Jones B
Pathways and Roadblocks in Navigating Online Cancer Communities: Qualitative Study Among Young Adult Cancer Survivors
JMIR Cancer 2026;12:e79893
URL: https://cancer.jmir.org/2026/1/e79893
doi:10.2196/79893

© Qi Chen, Erin Donovan, Leyi Zhou, Lailea Noel, Barbara Jones. Originally published in JMIR Cancer (https://cancer.jmir.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Benefits and Challenges of a Digital Exercise and Mind-Body Program During Active Cancer Treatment: Qualitative Study of Patients' Perceptions

Karolina L Bryl[1], MS, PhD; Sierra Silverwood[2], BA; Krupali Desai[1], MPH; Kelsey Schobert[1], MPH; Xiaotong Li[1], PhD; Susan Chimonas[3], PhD; Jun J Mao[1], MD, MSCE; Erin F Gillespie[4,2], MD, MPH

[1]Integrative Medicine & Wellness Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[2]Department of Radiation Oncology, University of Washington School of Medicine, Fred Hutch Cancer Center, Seatttle, WA, United States

[3]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[4]Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, United States

Corresponding Author:
Karolina L Bryl, MS, PhD
Integrative Medicine & Wellness Service, Department of Medicine
Memorial Sloan Kettering Cancer Center
321 East 61st Street, 4th Fl.
New York, NY, 10065
United States
Phone: 1 6466088573
Email: brylk@mskcc.org

Abstract

Background: Individuals undergoing cancer treatment often face a high symptom burden that impairs quality of life. Exercise and mind-body therapies have been shown to reduce symptoms but are underused. We developed a digital exercise and mind-body therapy program that effectively reduces symptoms while overcoming in-person delivery barriers. Understanding patient experiences can inform treatment mechanisms and guide digital health interventions in cancer care.

Objective: This study aimed to explore patient experiences with Integrative Medicine at Home (IM@Home), a 12-week live digital program delivering exercise and mind-body therapies tailored to the needs of individuals undergoing cancer treatment.

Methods: This qualitative study was embedded in a randomized clinical basket trial (NCT05053230) evaluating the effects of IM@Home versus enhanced usual care on symptoms and acute health care utilization in adults with solid tumors undergoing active treatment and experiencing moderate or greater fatigue. Using maximum variation sampling, 20 participants were selected for semistructured interviews. Interviews explored participants' experiences with the program, its impact on outcomes, unmet needs, and suggestions for improvement. Transcripts were analyzed using a combined inductive and deductive thematic analysis.

Results: Twenty participants (mean age 63, SD 9.6 years; 18/20, 90% female) were interviewed. Five major themes emerged: (1) IM@Home alleviated symptom burden and supported symptom self-management; (2) IM@Home facilitated social support and information exchange; (3) IM@Home offered a flexible, tailored program in a group setting; (4) IM@Home facilitated accessible, cost-effective support; and (5) recommendations for program enhancement. IM@Home was perceived as an accessible, flexible, and supportive program that promoted physical and emotional well-being during treatment.

Conclusions: IM@Home demonstrates a promising model for delivering integrative supportive care during cancer treatment. Findings highlight patient-valued features such as real-time guidance, tailored content, and community support. These insights can inform future implementation, integration into clinical care, and efforts to enhance digital mind-body interventions in oncology.

Trial Registration: ClinicalTrials.gov NCT05053230; https://www.clinicaltrials.gov/study/NCT05053230

International Registered Report Identifier (IRRID): RR2-10.1038/s41746-024-01387-z

(JMIR Cancer 2026;12:e80075) doi:10.2196/80075

KEYWORDS

integrative oncology; integrative medicine; digital health; supportive cancer care; mind-body therapies; exercise; qualitative research; symptom management; telehealth

XSL•FO
RenderX

Introduction

In 2024, an estimated 2,001,140 new cancer cases were expected in the United States [1], with most individuals undergoing systemic therapy or radiation [2]. While these treatments have significantly improved survival rates, people with cancer often experience a high symptom burden, including fatigue, pain, and insomnia, which can lead to reduced quality of life and, in some cases, treatment discontinuation [3,4]. A recent systematic review found that up to 90% of patients undergoing active cancer treatment reported unmet needs, with psychological (eg, anxiety and fear) and physical (eg, tiredness and lack of energy) challenges being most prevalent [5]. Exercise and mind-body therapies (eg, meditation, yoga) are endorsed by leading oncology organizations, including the American Society for Clinical Oncology, the Society for Integrative Oncology, and the National Comprehensive Cancer Network, as effective strategies for managing cancer treatment-related symptoms [6-10], therefore, improving treatment adherence and survival rates [11-13]. However, clinical implementation remains limited due to barriers, such as limited access, restricted modality options, and logistical challenges (eg, travel) [14-19]. Innovative delivery methods are urgently needed to address these barriers and integrate these therapies into oncology care.

Digital health technologies, including remote delivery of clinical or supportive care via telecommunications (telehealth), health services delivered through mobile health devices, and internet-based health services (eHealth), show promise for improving access [19-22] and are widely used in cancer care for symptom management [23-27], treatment supervision [28,29], and clinical trials [30-33]. Examples relevant to cancer care include mobile apps for symptom tracking and medication [34,35] and wearable activity monitors to promote physical activity and exercise [36,37]. However, their adoption for delivering exercise and mind-body therapies during cancer treatment remains limited, with concerns including ensuring safety when patients engage remotely, tailored patient-centered program design for online group formats, and addressing technical barriers [19]. To address these gaps, at Memorial Sloan Kettering Cancer Center (MSK), we developed a novel Integrative Medicine at Home (IM@Home), a 12-week live, digital telehealth program delivering group-based exercise and mind-body therapies [38,39], designed specifically for individuals with cancer. IM@Home combines multiple evidence-based modalities—including yoga, meditation, and fitness—with intensity and activities tailored to participants' needs and treatment-related limitations. Patients have unlimited access to classes, allowing them to choose sessions based on their preferences, providing flexible and ongoing support. A randomized clinical trial (NCT05053230) of 200 patients in active treatment demonstrated IM@Home is safe and effective in reducing fatigue, anxiety, and depression, while lowering acute health care utilization [40].

While quantitative data support IM@Home's clinical benefits, understanding the end user experience is essential to inform broader implementation and patient-centered refinement, particularly in oncology, where patient safety and care needs are heightened [41-43]. To fully understand how complex interventions such as IM@Home lead to change in outcomes, it is important to examine the broader context in which these interventions operate. This includes moving beyond standardized outcome assessments to actively include the voices of individuals directly targeted by the intervention, whose experiences and perspectives are crucial to refining the delivery and ensuring real-world relevance [44]. Yet, few studies have examined patient perspectives on synchronous digital mind-body interventions during active treatment, a period marked by significant physical, psychological, and logistical challenges [41,43]. This study addresses that gap by exploring participants' views on the program's digital delivery, usability, and impact on symptom relief and quality of life. The findings offer insights to support the safe, effective, and accessible integration of digital integrative oncology interventions into routine cancer care.

Methods

Study Design

This qualitative study was conducted as part of the Integrative Medicine for Patient-reported Outcomes, Values, and Experience (IMPROVE; NCT05053230) trial, a randomized clinical basket trial assessing the effects of the IM@Home program on symptoms and acute health care utilization in patients with moderate or greater fatigue undergoing active treatment for melanoma, head and neck, lung, gynecological, or breast cancer. From October 2021 to March 2023, a total of 200 participants were randomized to 12 weeks of either IM@Home intervention (live exercise and mind-body classes via Zoom [Zoom Video Communications]) or enhanced usual care (standard of care plus on-demand meditation resources). Study design, primary results, and detailed program description have been published elsewhere [38,40].

The IM@Home program offers 23 live classes per week, including movement-focused classes (eg, fitness and dance), mind-body classes (eg, yoga and tai chi), and mind-focused classes (eg, meditation and music therapy). Classes are delivered synchronously (in real time) via Zoom, each lasting 30-60 minutes, and led by licensed MSK Integrative Medicine Service clinicians—including certified fitness trainers, yoga and mindfulness instructors, and licensed music therapists with experience in oncology—who guided participants, offered modifications, and provided verbal cues and encouragement to support engagement and safety. Upon randomization, a member of the research team contacted participants by phone to introduce the program, provide an overview of class types, and assist with registration. Participants could attend as many classes as desired each week, select any combination of classes, and modify their selections as needed.

Participants

We used maximum variation sampling, targeting approximately 20% of IM@Home participants to capture diverse experiences across cancer types and different stages of program participation (study time points: Weeks 4-7, Weeks 8-12, or after study completion). The sampling framework, including the number of participants selected from each cancer type and time point, can be found in Multimedia Appendix 1. Selected participants were contacted by a member of the research team via phone to

XSL•FO
RenderX

invite them to participate in the qualitative interviews. All participants approached for interviews agreed to participate.

Data Collection

We conducted semistructured interviews with 20 IM@Home participants, led by 2 trained qualitative researchers with backgrounds in integrative therapies and psychology (KS and KB). Interviewers had previously briefly interacted with some participants during their participation in the IM@Home program, either during consent or later while providing assistance with class registration. At the start of each interview, participants were informed of the interviewers' professional roles. The research team developed and used an interview guide to explore key aspects of participants' experiences while allowing flexibility in responses (Multimedia Appendix 2). To ensure participant-driven insights, interviewers used open-ended probing questions and adapted them in real time to explore emerging topics.

Interviews were scheduled according to participant preference and availability, conducted via HIPAA (Health Insurance Portability and Accountability Act)-compliant Zoom, lasted up to 20 minutes, and were audio-recorded with the participant's consent using the Zoom recording function. Audio files were securely stored on an encrypted, password-protected institutional server accessible only to the study team. The audio recordings were transcribed verbatim (KS), reviewed for accuracy (KB), and de-identified by both coders (KS, KB) prior to analysis. Reflexive journaling and team debriefings were used to minimize interviewer bias.

Data Analysis

We analyzed all interview transcripts using a 2-phase approach that combined inductive and deductive thematic content analysis [45,46]. Our analytic process followed the 6-phase reflexive thematic analysis framework described by Braun and Clarke [46] familiarization with the data; generating initial codes; searching for themes; reviewing themes; defining and naming themes; and producing the final report.

In Phase 1, two researchers (KS and KB) independently read 5 transcripts to identify key narrative content and determine the unit of analysis (eg, sentences or full passages) (Phase 1: familiarization). Through an inductive, open coding process, they developed descriptive and interpretive codes (Phase 2: generating initial codes), allowing themes to emerge from the data. The researchers then met to discuss coding and emerging themes and iteratively refined and grouped codes into higher-order categories, forming thematic units (Phase 3: searching for themes; Phase 4: reviewing themes). Analytic memo writing supported the identification of emerging patterns and conceptual relationships. This iterative process ensured that findings were grounded in participant experiences rather than predetermined categories.

In Phase 2, using a deductive approach, the finalized coding scheme was applied across all transcripts by 2 researchers (KS and KB) to assess coding consistency and intercoder agreement. A third researcher (KD) reviewed coded transcripts to enhance credibility, with any discrepancies resolved through discussion. Once coding was completed, the final themes, categories, and representative quotes were reviewed and discussed with the broader research team (Phase 4: reviewing themes**;** Phase 5: defining and naming themes; and Phase 6: producing the report). To further enhance confirmability, an independent reviewer (SC) examined the coding and interpretation for potential bias or inconsistency. NVivo software (version 14; QSR International) was used to support data management and analysis. Thematic saturation was evaluated continuously throughout the analysis.

Strategies to Ensure Trustworthiness

Credibility was supported through researcher triangulation (3 coders participated in data coding, analysis, and interpretation), peer debriefing with the broader research team, and reflexive memo writing. Transferability was addressed through purposeful sampling and thick description of the context, participants, and findings. Dependability was ensured via a detailed audit trail documenting all stages of data collection, coding, and analysis, along with intercoder agreement among coders.

Confirmability was enhanced by maintaining an audit trail, practicing reflexivity, and conducting an external audit by an independent reviewer (SC) to assess for bias or inconsistency.

Ethical Considerations

The study was approved by the Institutional Review Board (IRB) at MSK (IRB# 21-369), registered at ClinicalTrials.gov (NCT05053230). The ethics approval for the qualitative interviews was part of the original IRB approval. Participant confidentiality was maintained; data were collected and stored in secure, password-protected systems and analyzed in deidentified form. Participants did not receive compensation for study participation. All participants provided written informed consent. Findings are reported per the Standards for Reporting Qualitative Research guidelines [47] and COREQ (Consolidated Criteria for Reporting Qualitative Research) extension [48]. The COREQ checklist is provided as a Multimedia Appendix 3.

Results

Participant Characteristics

All participants approached for interviews agreed to participate. Of the 20 participants interviewed (mean age 63, SD 9.6 years), a majority were female (18/20, 90%), White (17/20, 85%), and had breast cancer (9/20, 45%). Table 1 summarizes the demographic and clinical characteristics of the 20 participants included in the qualitative interviews. This qualitative study sample was broadly reflective of the overall IMPROVE trial population (N=200) and the IM@Home arm (n=99) with respect to gender, age, and cancer type, although there was a slightly higher proportion of White and Asian participants and a slightly lower proportion of Black participants compared with the IM@Home arm.

XSL•FO
RenderX

Table 1. Participant characteristics.

Participant characteristic	Values
Age (years), mean (SD)	63 (9.6)
Gender, n (%)	
Female	18 (90)
Male	2 (10)
Race, n (%)	
Asian	2 (10)
Black	1 (5)
White	17 (85)
Ethnicity, n (%)	
Non-Hispanic or Latino	18 (90)
Hispanic or Latino	0 (0)
Unknown	2 (10)
Cancer type, n (%)	
Breast	9 (45)
Gynecological	3 (15)
Head and neck	4 (20)
Lung	3 (15)
Melanoma	1 (5)
Time of the interview, n (%)	
After study completion	2 (10)
Week 4-7	7 (35)
Week 8-12	11 (55)

Qualitative Findings

Saturation was achieved by the 17th transcript; however, all 20 interviews were coded to capture potential nuances and ensure comprehensive representation. Ultimately, 5 major thematic units were identified. These thematic findings are summarized in a narrative account below, while detailed theme descriptions, including definitions, categories, subcategories, and additional illustrative quotes, are provided in Table 2.

XSL•FO
RenderX

Table 2. Qualitative findings.

Themes and their definitions, and categories and subcategories	Example quotes
Theme 1. IM@Home alleviated symptom burden and supported symptom self-management: *IM@Home helped participants manage cancer-related symptoms by improving physical strength, reducing fatigue, alleviating stress and anxiety, and providing accessible self-coping strategies to enhance overall well-being.*	
Reduced treatment-related symptom burden	• "The thing that definitely did help was also, you know, in terms of trying to sleep the guided part of getting you to relax, because I've always been a bit of an insomniac, and I think this one didn't. This whole episode of going through all this didn't help in that one. So that part definitely helped just being able to kind of visualize every part of you, relaxing and going to sleep." [P003] • "It's hard to explain, I know it's good for me and I do it. My back hurts and my neck hurts, and whatever I'm doing…I don't know if it's the Chair yoga or one of the other things, but my back feels better. The neck, I think, got a little worse. But the back feels better, so I know, whatever I'm doing it is helpful in the long run." [P002] • "I really like it. I feel like a different person. I feel so much looser. I just feel like 'wow', and it's not like you're sweating, and it's really tough, but you can really feel these small movements, and I just feel good. I feel so much better like it gives me quite a bounce to my day." [P017] • "I do feel that it was um I do feel like it did give me a greater range of motion, and I do feel that I slept better." [P010]
Enhanced symptom self-management and sense of control	• "I would describe it [IM@Home program] as a tool for coping. Because I think being able to figure out how to um move in the world emotionally and physically with the illness, or the recovery from an illness is coping and coping and adapting is critical." [P010] • "I think it's such a stressful time in your life, because you really don't know what to what to expect. What you're going through and all, you don't know how you're physically the physically going to be reacting to anything, and it's such a stressful time. I think if you can plug that [IM@Home] into your day in some way, even, maybe a couple of times a day to help you kind of get yourself grounded, and I think that that would be helpful. I think you just have to try and stay focused as much as you can, even though it's quite difficult." [P011] • "I definitely feel more like myself […] where I'm not so consumed with cancer and treatments, like I almost feel like I can. I feel mentally better. That's one thing I would definitely say, I don't feel as upset all the time. I don't feel as down, because it was definitely getting, like, as the as the radiation went through like definitely emotional and teary and sort of feeling sorry for myself. So I definitely feel that I do feel more positive and just mentally like a bit more clear and like a bit more positive and going forward." [P017] • "Yes, actually, psychologically better because I finished it. Finally, you got up off the couch from watching TV or something on a day when I was doing that. Not as good as I felt if I could get out shopping for three hours. That is the ultimate kick in the pants. Yeah that [IM@Home] gets me going." [P005]
Theme 2. IM@Home facilitated social support and information exchange: *the program fostered a sense of community, offering emotional support, reducing isolation, and enabling peer-to-peer learning through shared experiences and instructor guidance.*	
Building a sense of community	• "There's also a sense of community because usually in the beginning, at the end there's little discussion, a talk, and you hear people talking giving little advices, there is a little community there and I like it." [P004]
Feeling seen and supported	• "Beyond the exposure to something that can provide comfort and health benefits, there is a certain companion ability among people if you attend the same class. There are some of the same people in those classes, and they do talk to the instructor primarily after the class. […] And you know, in the chat function afterwards, somebody will mention [personal stories] and just a nice kind of things that people would say, or someone talks about something they're having a hard time with, or something that went well. And there are all these other people don't know, but who are saying, you know congratulations, or you know that's tough, or maybe throw out a resource they've heard of. So it's not a big bulletin board, it's not a big chat, but there's a little bit of a sense of community I think, is nice." [P012]
Knowledge and resource sharing	• "I think it's very good for that [sharing personal experiences] it seemed like a lot of the conversations and the chats and the things that they [patients] were talking to each other. You know this is the best speaker [platform] to use. You know, I think, all that neighborliness is very useful too. I think a sense of community is always good." [P010] • "I don't have my camera on but some people do and it builds a camaraderie, which some people might need. Thank God, I have a lot of support from friends, from family members, you know, but some people might not have that. And to get a support from the group of people, I think that that would be very beneficial." [P002]

XSL•FO
RenderX

Themes and their definitions, and categories and subcategories	Example quotes
Theme 3. IM@Home offered a flexible, tailored program in a group setting: *participants valued the program's safe, flexible, and personalized approach within a group format, allowing them to engage in oncology-tailored exercises suited to their needs, schedules, and symptom burdens.*	
Ensuring safety and personalization	• I think the most helpful was just the fact that the instructors were clear […] they made it very clear like 'Well, here's our warm-up,' and they really did make sure, like I've been to some fitness classes where there's really no emphasis on the warm-up or on the stretching afterwards. And that's such an important part of the program. So I thought that that was probably the most beneficial. And, in my opinion, is that they really did look at this holistically and say, like, here are people who are recovering and or bring in something so let's make sure that they stretch appropriately before we dive into the program, and then let's make sure at the end." [P020] • "Well, I've been sick, for almost two years, I'm getting better now, and it makes me feel good. I mean after yoga, gentle yoga it's very gentle but it's just my speed, my level and I can't believe how good it is for me and then like its wow that feels so easy, then I feel the stretches, I feel the muscles that I haven't used for a while. So [the classes] make me feel good." [P004] • "There was a large group of people of varying levels. And so instructors wanted to make sure that everybody was able to participate. I like the emphasis on breathing. I like the emphasis on just moving, being able to do what you know. Just move. Just do whatever you can to move." [P020] • "It's something that's very good for you, and you don't need like a lot of equipment, either, because that's the other thing that people find discouraging is oh, like I don't have weights. Well, you could use water bottles, or you could use towels, or you could use. I think what she said was, you could use your thumb and your index finger, like those types of modifications, were very helpful. So I think that's how I would sell it to her judgment-free, all activity levels the privacy of your own home with instructors that give you different options for your fitness, level." [P010]
Offering variety and flexibility	• "There are abundant resources through the program for it [symptoms], access to, particularly to mind, body experiences that are interesting to explore if you've never done them before, and if you've done them before they're a way to maintain a practice […] It's good for you. It feels good to do it, and it can help you step outside yourself a little bit." [P012] • "You can choose which ones you want, and there's a variety and different levels." [P015] • "I think that's what I liked best about this program that it was just at a level and at a pace where you can kind of ease into it and then modify as you got stronger." [P020] • "I didn't find it difficult or challenging, because I think people provided enough guidance to say, 'Hey, here, if you are a beginner, here's where you could do it.' This is, for example […] push-ups right, you could use a wall or a chair or the floor or do a full-on push-up. So, there's different types of push-ups you can do depending on your level. I didn't really think it was all that challenging. And then there was always the caveat that if you weren't comfortable doing that, then, you know, and just keep moving or do something else." [P020] • "It was a good variety. It gave you just different the different chances to do the different types of classes. Also, for the length of time that I had on my schedule so that I was able to do it. I thought the variety was good. Having them [classes] at the different time sequences, also the different types they had, where they kind of had you going you know mentally and all that. That was also very good, because it gave some variety." [P011]
Role of instructors	• "And the most helpful is because the instructors also say you can do this, keep going and they motivate you to go, and that helps me. It's just a way that they, I can't explain it, it's just the way that they do the classes that I enjoy and it's like, you know, like when I get tired, I just keep going." [P018] • "I think the other positive was that there was pretty good instruction, like sometimes when you do online classes. and you know, and the cameras aren't turned on for most of the people like you don't know if they're going to do something that might injure themselves, but they manager themselves. But I feel like people were, you know, in a pretty good... they gave pretty good instruction. But I know that's always a challenge with online classes, especially when you're when you have a group of people of varying levels. Making sure that no one gets hurt." [P020]
Theme 4. IM@Home facilitated accessible, affordable, and ongoing support: *the digital format removed barriers such as cost, travel, and infection risk, making exercise and mind-body therapies accessible to diverse patients while supporting long-term wellness beyond active treatment.*	

XSL•FO
RenderX

Themes and their definitions, and categories and subcategories	Example quotes
Improving accessibility and convenience	• "I would be concerned about going to the gym and even though that don't have to have a mask for indoors now I'll continue wearing a mask because I don't need to catch anything else, and a lot of people that are basically homebound at this point either because of the situation, because of COVID, or because of the way they feel. It's good to have something that they could do at home due to their fitness level. That's not difficult and is enjoyable. I would recommend it." [P002] • "Especially after surgery and radiation, when you're tired and you want to regain your energy, it's most people aren't self-motivated to join a gym or do anything like that, but just the fact that it's in privacy of your home that there are people that are there that are just like you. And that it's geared and tailored for patients." [P020] • "I was a very regular gym goer and I always would go in, lift weights and do some cardio, and would always take a yoga class. I probably did like 4 or 5 a week […] since the pandemic, I'd stopped going to the gym, for fear of infection. And because of the pandemic I'm still not there, I'm not going to a gym. I'm just not going to be in a big building with people breathing hard and no masks. So that that's the biggest benefit that it's just the ability to get back into some sort of fitness routine." [P015] • "It was simple. I mean I just chose a class clicked on it got a registration form put in my name, address phone number whatever and then the next day or I got a link to the class the simple." [P002]
Providing structure and motivation	• "Sometimes the treatment can have a very demotivating effect, and having something to look forward to, I feel helped so much, mentally." [P016] • "Even though I get tired I can keep going better than before I was doing the classes, so I really think that they're starting to help build me up again, […] which is a good thing, because I was almost to the point where I was just sitting in my chair and said, 'Forget this, I'm tired' So now I'm actually looking forward to doing the classes." [P018] • "And the most helpful is because the instructors also say you can do this, keep going and they motivate you to go, and that helps me. It's just a way that they, I can't explain it, it's just the way that they do the classes that I enjoy and it's like, you know, like when I get tired, I just keep going." [P018]
Theme 5. Recommendations for program enhancement: *participants provided several suggestions for optimizing the program's implementation and maximizing its benefits.*	
Need for additional guidance	• "I think the one thing that can be a little difficult, which maybe in-person would have helped, is the Tai chi. I tried doing it, and it's so hard to, because you have like a mirror image, and the instructor is telling you to twist your body a certain way. It's hard to figure out exactly how to do that. So I think that was the only class that I was very confusing in terms of doing it online." [P016] • "One thing I will say about the course description, that I read before I went to each class, was I didn't feel that they were very clear on the type of material you should bring to the class […] I don't recall now what they said but it's like some classes, we use stretchy bands and I know [instructor], would talk a lot about what you could use to substitute and maybe you don't want to discourage people from coming you know joining thinking that they don't have all the materials that are maybe used in the class." [P006]
Expanding delivery methods	• "Let me just say this, if you could record the classes and rerun them in the evening, right the same class. So, for example, you know [instructor's name] does a tai chi class in the morning, and then you ran a recording of it at night where people couldn't attend in the morning. I mean it's I know it's a whole technological issue to get that set up and to do it, but you wouldn't have the instructor having to instruct again in the evening, you could replay a class." [P006]
Timing considerations for the program introduction	• "Having the program being offered to patients in conjunction with we start treatment and have someone help us come up with a training and exercises plan from the beginning would help to be active." [P008] • "Why can't the doctors mention it to the patients when they do like the five week follow up from their surgery and say, 'Oh, we have an Integrative program. I think it would be good for you to have exercise in your healing process to recover better.' And maybe you have to like kind of coordinate it with the doctors." [P009]

XSL•FO
RenderX

Theme 1. IM@Home Alleviated Symptom Burden and Supported Symptom Self-Management

Participants reported that the IM@Home program was valuable for managing cancer-related symptoms during treatment, offering physical, mental, and social benefits.

Reduced Treatment-Related Symptom Burden

Participants highlighted significant improvements in physical and mental health, including reductions in cancer treatment-related symptoms, particularly radiation-related fatigue, overall stress, and anxiety. Fatigue reduction led to increased muscle strength, flexibility, and mobility, while decreased stress and anxiety enhanced their motivation to stay active.

> *It really helped. I feel like my muscles are stronger, my walking is a lot easier, and I'm not as tired as I used to be.* [P006]

> *What I noticed was with the radiation fatigue, I think about week three, I was like, 'Wow, I'm good.' I've got my normal energy back, like there's no more fatigue. That was probably the biggest benefit, just like flipping a switch, 'I'm good!'* [P015]

These improvements collectively increased their ability and willingness to continue classes and carry out daily activities and further contributed to the reduction of comorbid symptoms such as insomnia. Yoga and fitness classes were frequently mentioned for fatigue management and meditation classes for reducing stress.

Enhanced Symptom Self-Management and Sense of Control

Many participants described the challenges they faced during treatment, particularly after leaving the hospital. This was when treatment-related symptoms emerged, requiring them to manage these difficulties on their own, often leading to a sense of lost control. In this context, participants noted that IM@Home provided valuable, accessible, and easy coping skills, such as breathing techniques, to help them navigate symptoms or manage difficult moments at home. Additionally, completing classes provided participants with a sense of control, giving them the courage and confidence to continue their ongoing treatments.

> *With all the stress of the diagnosis and treatment that's one [breathing and relaxation] that I realized is a good thing for the stress release […] there were days when things were happening to me, and my ability to just lay there and count my breaths or breathe through things was really a great thing.* [P015]

The anticipation of classes also provided mental relief, with one participant noting, "Having something to look forward to helped so much mentally" (P016). Another participant highlighted the program's impact on self-commitment and achievement.

> *At the conclusion of each class, there is a feeling of accomplishment. I did something for me, I kept a commitment to myself. I showed up and got through the whole thing. I feel good about doing it.* [P012]

Theme 2. IM@Home Facilitated Social Support and Information Exchange

The program fostered a sense of community, helping participants feel less isolated. As one participant noted, "There is a feeling that you're doing this together" (P012). Another described class interactions as feeling like "a friend came into their home" (P005).

Participants valued the emotional support of being seen and not feeling alone.

> *You see the other people that are in the class with you, that's comforting not to be alone, because a lot of us are alone.* [P013]

Receiving support and encouragement from peers and instructors contributed to a more positive experience during the treatment process.

> *It's comforting seeing the instructor and hearing them acknowledge you makes a difference.* [P013]

Participants described how postclass discussions provided opportunities to exchange advice and learn from others' experiences, further enriching the program's social aspect.

> *I learned a lot from the instructor's comments about the effects of radiation on muscles and fascia, which came from an exchange with another class participant […] It was enlightening for me, and it was what I was experiencing.* [P012]

Theme 3. IM@Home Offered a Flexible, Tailored Program in a Group Setting

Participants described their experience with IM@Home as safe, patient-centered, and flexible. They noted differences compared to other general fitness classes they had attended, expressing that IM@Home felt safe and comfortable. The activities were tailored to fitness levels and the unique physical and mental needs of cancer patients, including those with metastatic disease and a high symptom burden.

With good tolerance and noticeable benefits, they continued engaging with IM@Home and used it as a valuable tool for maintaining their quality of life. As one participant explained,

> *Exercise and [being active] can keep things from reoccurring, maximize your lifestyle and functionality, and help you regain skills you've lost. When you're living with an illness, dealing with something chronically, you need as many tools as you can to alleviate pain and symptoms and to continue to live as fully as you can, even though you're still sick.* [P010]

Participants also noted the key contributions of instructors specialized in oncology care.

> *The instructors were really top-notch, and they were specifically tailoring the exercises and the activity levels.* [P010]

Additionally, even in a group setting, patients felt that the program was individualized to their preferences, schedules, and daily routines. IM@Home offered a variety of classes, including

XSL•FO
RenderX

yoga, meditation, and fitness, empowering participants to choose activities aligned with their preferences and physical needs. Its flexibility allowed them to continue participating in the face of such challenges as scheduling conflicts with treatment appointments, changes in symptoms or needs, or concerns about monotony.

> *I thought it was great that it would be in my own home, and with little to no equipment. All those things they thought were fantastic because when you have ongoing illness you there are things like side effects, where you might have to run to a bathroom and to physically be in a class it would make all those things difficult, whereas here you had some latitude as to when you could attend or leave the class or come back to the class, and so I thought I was primed for success at it.* [P010]

Theme 4. IM@Home Facilitated Accessible, Affordable, and Ongoing Support

Participants appreciated the program's digital format for increasing accessibility in several ways: (1) compared to in-person classes, the digital format was more cost-effective, which was important given the financial burden of cancer treatments; (2) the digital format eliminated the need for travel, reducing barriers for those with mobility limitations; and (3) it minimized infection risks for patients who could not safely attend in-person sessions, particularly those who were immunocompromised. They emphasized that the digital format allowed not only the general cancer population but also those facing significant barriers to in-person attendance to receive support and stay active.

> *I don't have to go anywhere, I don't have to travel, I turn on my computer.* [P003]

> *While you're going through this process and can't get out and do some of the things you used to, or that you thought about doing […] This is a productive and psychologically healthy way to spend some of that extra time. It feels good to do it.* [P011]

Additionally, participants emphasized the importance of having resources to maintain healthy behaviors despite treatment side effects and even after completing cancer treatments. They appreciated the online program provided ongoing access, allowing patients to engage whenever needed and supporting their commitment to long-term wellness.

> *Knowing there is an interface out there I can go to any time, if I want to do some meditation or some light exercises like chair yoga, it really helps that there is something out there I can lean on.* [P016]

> *I was not exercising at all. I was becoming very lethargic. Having this program helped me on a daily basis. I signed up, I keep signing up for more and more classes, and it keeps me starting my day with exercise and it made me feel better.* [P003]

Theme 5. Recommendations for Program Enhancement

Participants provided several suggestions for optimizing the program. Some suggested offering more detailed guidance alongside the classes to help them understand complex movements or better prepare for class participation.

> *It's so hard to figure out exactly how to do [movement]. It was very confusing in terms of doing it online.* [P016]

They also recommended incorporating additional delivery methods, besides live online classes, to benefit patients without reliable internet access or devices. Furthermore, participants proposed introducing the program earlier in the cancer care journey, such as during consultation, to maximize its benefits.

> *Could there be a consultation where, after meeting your surgeon, you also meet your oncologist and learn about programs like IM@Home to help you along the road?* [P017]

Discussion

Overview

This qualitative study explored the experiences of patients undergoing active cancer treatment who participated in a 12-week cancer-specific digital exercise and mind-body program, IM@Home. Participants described a broad range of benefits, including improved symptom management and enhanced social support, and reflected on the program's novel, patient-centered, group-based online delivery. They also identified opportunities for improvement, such as adding real-time feedback, peer mentorship, and recorded sessions to complement live classes. Taken together, the findings suggest a preliminary program model in which digitally delivered, flexible integrative medicine programming promotes symptom relief and peer connection, mechanisms that supported patients during treatment. These mechanisms appear shaped by contextual factors (eg, treatment-related fatigue, mobility limitations, and scheduling barriers) and can inform refinements to optimize delivery. Overall, these findings provide timely insights into how live, structured, digitally delivered integrative care can address supportive care needs during active treatment and reinforce prior evidence on the value of synchronous, patient-centered digital interventions in oncology.

Principal Findings

Exercise and mind-body therapies are well-documented for managing cancer-related symptoms [7,40,49-53] and for improving quality of life during treatment [40,54-60]. Our interviews' findings on symptom management (ie, reduced fatigue and mental stress) align with previous trials and are consistent with results from our randomized clinical trial [40]. This suggests these symptoms may be primary targets for online exercise and mind-body therapies in cancer care. Additionally, our findings indicate that symptom relief came not only from the interventions but also from patients' own continued efforts, as they learned self-coping strategies and received encouragement and support from other participants. Existing literature underscores the importance of self-coping in symptom management and healing [61-63], highlighting its role not only in managing symptoms during treatment but also in supporting long-term survivorship.

XSL•FO
RenderX

Our study provides valuable insights into the design and delivery of digital interventions. The increasing popularity of digital health reflects its ability to overcome traditional barriers to care such as accessibility and scheduling [64]. Remote delivery of clinical or supportive care via telecommunications (ie, telehealth) is particularly beneficial for immunocompromised patients, offering privacy and reduced infection risks [65]. However, many digital programs face implementation challenges in cancer care, such as balancing intervention effects with patients' tolerance, delivering the program in a group setting while maintaining patient-centered care, and engaging patients in routine participation without adding to their burden [19,50,66,67]. Participants indicated that IM@Home successfully addressed these challenges by integrating multiple evidence-based therapies tailored to patients with cancer, accommodating different exertion levels to ensure safety and align with patient tolerance, and allowing flexible access to various treatment modalities. This flexibility enabled participants to select activities based on preferences, symptoms, and schedules, while maintaining regular engagement. These features enhanced safety, improved adherence, and expanded access, ultimately increasing the effectiveness of the interventions.

A distinctive feature of IM@Home was its live, group-based delivery format, which fostered community, reduced feelings of isolation, and created a safe, supportive environment, crucial for maintaining quality of life and coping during cancer treatment [68,69]. Participants emphasized the unique value of real-time interaction with instructors and peers, describing the experience as supportive and empowering. This structure stands in contrast to many digital interventions that rely on asynchronous content and may lack opportunities for interpersonal connection or individualized guidance. Social support was repeatedly cited as a motivating factor, as participants found validation in shared experiences and encouragement from others. These relational aspects, often overlooked in digital health design, are critical for sustained engagement and psychological support during cancer care [70]. This finding aligns with broader calls for a "patient revolution" in the health care system, emphasizing the need to empower patients to take a more active role in their care [71]. A key component of this shift is enabling patients to seek advice and support from others who have faced similar challenges. Based on patient feedback, they received empathy, insight, and practical strategies from both peers and instructors, fostering a greater sense of control and satisfaction with care and alleviating pressure on the health care system.

Our findings underscore the need to integrate digital exercise and mind-body programs like IM@Home into standard oncology care. Early incorporation into treatment planning, perhaps during initial consultations, could maximize benefits by establishing wellness routines before treatment-related symptoms worsen. Participants also highlighted the need for additional guidance or on-demand resources to complement online classes. Potential solutions, such as enhancing real-time feedback, incorporating peer mentorship, and providing recorded sessions, should be considered in future program development. Additionally, multicenter trials are needed to assess program effectiveness across diverse populations, alongside cost-effectiveness and scalability. Furthermore, improving patient education and clinician-patient communication about the benefits of digital interventions could promote adoption and adherence. By addressing current gaps, digital exercise and mind-body programs like IM@Home can be optimized to meet the complex needs of patients with cancer and expand access to effective, comprehensive care. These insights offer actionable strategies for future program refinement and broader clinical integration.

From an implementation perspective, IM@Home demonstrates a scalable model for integrating digital exercise and mind-body therapies into routine oncology care. By reducing barriers related to cost, travel, and immunocompromise, the program has the potential to increase access and address equity gaps in supportive care. Embedding such interventions early in the cancer care trajectory—and ensuring they are adaptable to individual needs—may help improve adherence and patient-centered outcomes and promote more equitable access to evidence-based supportive care. Future adaptations might include multilingual content, device-lending strategies, or hybrid models to serve patients with limited digital access.

Limitations and Future Directions

This study has several limitations. First, the small, predominantly female and White sample limits the transferability of findings. While we used maximum variation sampling to capture a range of experiences across cancer types and study time points, the demographic homogeneity may have influenced the range of perspectives reported. Future research should aim to recruit more diverse participants to enhance the transferability of findings. Second, we did not include participants who dropped out or conduct interviews at earlier time points. Including these participants could have provided additional insights into barriers and facilitators of engagement. Third, although formal member checking—returning preliminary themes to participants to assess whether the interpretations resonate with their experiences—was not conducted, casual interactions with participants supported the credibility of the findings. Finally, future studies should also explore long-term outcomes and implementation strategies (eg, integration into routine clinical workflows and program scalability) to advance equitable integration of digital integrative care into cancer treatment pathways.

Conclusions

IM@Home represents a novel, patient-centered, and scalable model for digital integrative care. Participants viewed the program as supportive, empowering, and effective in managing treatment-related symptoms. These findings lay the foundation for refining and scaling IM@Home to benefit diverse cancer populations and inform the development of future digital health interventions.

XSL•FO
RenderX

Acknowledgments

We thank all patients who participated in the study and acknowledge our multidisciplinary team of physicians, nurses, and practice providers. We also recognize research coordinators, administrators, IM@Home staff, and the Integrative Medicine Service at MSK that have supported study execution.

No generative artificial intelligence, including tools like ChatGPT or other language models, was used in any portion of the manuscript generation. All content was created and written by the authors.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Funding

This work was supported in part by the JDJ Foundation, AKTIV Against Cancer, a National Institutes of Health, National Cancer Institute Cancer Center Support Grant (P30 CA008748) and the Translational and Integrative Medicine Research Fund at Memorial Sloan Kettering Cancer Center. A National Cancer Institute (NCI) Career Development Award No. K08 CA252640 provided additional support to EFG. All funders or sponsors have not had any role in the design and conduct of this study and will not have any role in data collection, analysis of data, interpretation of the data, and dissemination of the findings.

Authors' Contributions

Conceptualization: JJM, EFG, KB; Data curation: KB, KS, KD; Formal analysis: KB, KS, KD; Funding acquisition: JJM, EFG; Investigation: KB, KS, KD; Methodology: JJM, KB, SC; Project administration: KB; Supervision: JJM, EFG; Writing - original draft: KB; Writing - review & editing: KB, SS, KS, KD, XL, SC, JJM, EFG;. All authors read and approved the final manuscript.

Conflicts of Interest

JJM reports receiving grants from Tibet Tibet CheeZheng Tibetan Medicine Co Ltd and Zhongke Health International LLC to Memorial Sloan Kettering outside the submitted work and serving in an unpaid role as co-president of the Society for Acupuncture Research. JJM reports financial interest in Greatly Health. MSK has institutional financial interests relative to Greatly Health. The remaining authors disclosed no conflict of interest.

Multimedia Appendix 1
Maximum variation sampling framework for qualitative interviews.
[DOCX File , 17 KB - cancer_v12i1e80075_app1.docx]

Multimedia Appendix 2
Interview guide.
[PDF File (Adobe PDF File), 879 KB - cancer_v12i1e80075_app2.pdf]

Multimedia Appendix 3
COREQ (Consolidated criteria for Reporting Qualitative research) checklist.
[PDF File (Adobe PDF File), 495 KB - cancer_v12i1e80075_app3.pdf]

References

1. National CI. Cancer Stat Facts: Cancer of Any Site. Cancer Stat Facts. 2024. URL: https://seer.cancer.gov/statfacts/html/all.html [accessed 2025-12-20]
2. Muhandiramge J, Warner ET, Zalcberg JR, Haydon A, Polekhina G, van Londen GJ, ASPREE Investigator Group. Cancer treatment patterns and factors affecting receipt of treatment in older adults: results from the ASPREE Cancer Treatment Substudy (ACTS). Cancers (Basel) 2023 Feb 05;15(4):1017 [FREE Full text] [doi: 10.3390/cancers15041017] [Medline: 36831362]
3. Whisenant MS, Williams LA, Mendoza T, Cleeland C, Chen T, Fisch MJ, et al. Identification of breast cancer survivors with high symptom burden. Cancer Nurs 2022;45(4):253-261 [FREE Full text] [doi: 10.1097/NCC.0000000000001019] [Medline: 34608052]
4. Shi Q, Smith TG, Michonski JD, Stein KD, Kaw C, Cleeland CS. Symptom burden in cancer survivors 1 year after diagnosis: a report from the American Cancer Society's studies of cancer survivors. Cancer 2011;117(12):2779-2790 [FREE Full text] [doi: 10.1002/cncr.26146] [Medline: 21495026]
5. Puts MTE, Papoutsis A, Springall E, Tourangeau AE. A systematic review of unmet needs of newly diagnosed older cancer patients undergoing active cancer treatment. Support Care Cancer 2012;20(7):1377-1394. [doi: 10.1007/s00520-012-1450-7] [Medline: 22476399]

XSL•FO
RenderX

6. Deleemans JM, Mather H, Spiropoulos A, Toivonen K, Baydoun M, Carlson LE. Recent progress in mind-body therapies in cancer care. Curr Oncol Rep 2023 Apr;25(4):293-307. [doi: 10.1007/s11912-023-01373-w] [Medline: 36753025]
7. Duan L, Xu Y, Li M. Effects of mind-body exercise in cancer survivors: a systematic review and meta-analysis. Evid Based Complement Alternat Med 2020;2020:7607161 [FREE Full text] [doi: 10.1155/2020/7607161] [Medline: 32952591]
8. Mao JJ, Ismaila N, Bao T, Barton D, Ben-Arye E, Garland EL, et al. Integrative medicine for pain management in oncology: Society for Integrative Oncology-ASCO duideline. J Clin Oncol 2022;40(34):3998-4024 [FREE Full text] [doi: 10.1200/JCO.22.01357] [Medline: 36122322]
9. Carlson LE, Ismaila N, Addington EL, Asher GN, Atreya C, Balneaves LG, et al. Integrative oncology care of symptoms of anxiety and depression in adults with cancer: Society for Integrative Oncology-ASCO guideline. J Clin Oncol 2023;41(28):4562-4591. [doi: 10.1200/JCO.23.00857] [Medline: 37582238]
10. Bower JE, Lacchetti C, Alici Y, Barton DL, Bruner D, Canin BE, et al. Management of fatigue in adult survivors of cancer: ASCO-Society for Integrative Oncology guideline update. J Clin Oncol 2024;42(20):2456-2487. [doi: 10.1200/JCO.24.00541] [Medline: 38754041]
11. Misiąg W, Piszczyk A, Szymańska-Chabowska A, Chabowski M. Physical activity and cancer care-a review. Cancers (Basel) 2022;14(17):4154 [FREE Full text] [doi: 10.3390/cancers14174154] [Medline: 36077690]
12. Lavery JA, Boutros PC, Scott JM, Tammela T, Moskowitz CS, Jones LW. Pan-cancer analysis of postdiagnosis exercise and mortality. J Clin Oncol 2023;41(32):4982-4992 [FREE Full text] [doi: 10.1200/JCO.23.00058] [Medline: 37651670]
13. Friedenreich C, Stone C, Cheung W, Hayes S. Physical activity and mortality in cancer survivors: a systematic review and meta-analysis. JNCI Cancer Spectr 2020;4(1):pkz080 [FREE Full text] [doi: 10.1093/jncics/pkz080] [Medline: 32337494]
14. Romero SAD, Brown JC, Bauml JM, Hay JL, Li QS, Cohen RB, et al. Barriers to physical activity: a study of academic and community cancer survivors with pain. J Cancer Surviv 2018;12(6):744-752 [FREE Full text] [doi: 10.1007/s11764-018-0711-y] [Medline: 30182150]
15. IJsbrandy C, Hermens RPMG, Boerboom LWM, Gerritsen WR, van Harten WH, Ottevanger PB. Implementing physical activity programs for patients with cancer in current practice: patients' experienced barriers and facilitators. J Cancer Surviv 2019;13(5):703-712 [FREE Full text] [doi: 10.1007/s11764-019-00789-3] [Medline: 31347009]
16. Kennedy MA, Bayes S, Newton RU, Zissiadis Y, Spry NA, Taaffe DR, et al. Correction to: Implementation barriers to integrating exercise as medicine in oncology: an ecological scoping review. J Cancer Surviv 2022;16(6):1504-1507 [FREE Full text] [doi: 10.1007/s11764-021-01116-5] [Medline: 35286583]
17. Avancini A, Trestini I, Tregnago D, Belluomini L, Sposito M, Insolda J, et al. Willingness, preferences, barriers, and facilitators of a multimodal supportive care intervention including exercise, nutritional and psychological approach in patients with cancer: a cross-sectional study. J Cancer Res Clin Oncol 2023;149(7):3435-3445 [FREE Full text] [doi: 10.1007/s00432-022-04232-6] [Medline: 35943598]
18. Desai K, Bao T, Li QS, Raghunathan NJ, Trevino K, Green A, et al. Understanding interest, barriers, and preferences related to yoga practice among cancer survivors. Support Care Cancer 2021;29(9):5313-5321 [FREE Full text] [doi: 10.1007/s00520-021-06083-2] [Medline: 33660077]
19. Shaffer KM, Turner KL, Siwik C, Gonzalez BD, Upasani R, Glazer JV, et al. Digital health and telehealth in cancer care: a scoping review of reviews. Lancet Digit Health 2023;5(5):e316-e327 [FREE Full text] [doi: 10.1016/S2589-7500(23)00049-3] [Medline: 37100545]
20. Aapro M, Bossi P, Dasari A, Fallowfield L, Gascón P, Geller M, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020;28(10):4589-4612 [FREE Full text] [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
21. Kemp E, Trigg J, Beatty L, Christensen C, Dhillon HM, Maeder A, et al. Health literacy, digital health literacy and the implementation of digital health technologies in cancer care: the need for a strategic approach. Health Promot J Austr 2021;32 Suppl 1:104-114. [doi: 10.1002/hpja.387] [Medline: 32681656]
22. Parikh RB, Basen-Enquist KM, Bradley C, Estrin D, Levy M, Lichtenfeld JL, et al. Digital health applications in oncology: an opportunity to seize. J Natl Cancer Inst 2022;114(10):1338-1339 [FREE Full text] [doi: 10.1093/jnci/djac108] [Medline: 35640986]
23. Sirintrapun SJ, Lopez AM. Telemedicine in cancer care. Am Soc Clin Oncol Educ Book 2018;38:540-545 [FREE Full text] [doi: 10.1200/EDBK_200141] [Medline: 30231354]
24. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016;34(6):557-565 [FREE Full text] [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
25. Head BA, Studts JL, Bumpous JM, Gregg JL, Wilson L, Keeney C, et al. Development of a telehealth intervention for head and neck cancer patients. Telemed J E Health 2009;15(1):44-52 [FREE Full text] [doi: 10.1089/tmj.2008.0061] [Medline: 19199847]
26. van den Brink JL, Moorman PW, de Boer MF, Hop WCJ, Pruyn JFA, Verwoerd CDA, et al. Impact on quality of life of a telemedicine system supporting head and neck cancer patients: a controlled trial during the postoperative period at home. J Am Med Inform Assoc 2007;14(2):198-205 [FREE Full text] [doi: 10.1197/jamia.M2199] [Medline: 17213498]

XSL•FO
RenderX

27. Daly B, Kuperman G, Zervoudakis A, Baldwin Medsker A, Roy A, Ro AS, et al. InSight care pilot program: redefining seeing a patient. JCO Oncol Pract 2020;16(10):e1050-e1059 [FREE Full text] [doi: 10.1200/OP.20.00214] [Medline: 32468925]
28. Sabesan S, Larkins S, Evans R, Varma S, Andrews A, Beuttner P, et al. Telemedicine for rural cancer care in North Queensland: bringing cancer care home. Aust J Rural Health 2012;20(5):259-264. [doi: 10.1111/j.1440-1584.2012.01299.x] [Medline: 22998200]
29. Jhaveri D, Larkins S, Kelly J, Sabesan S. Remote chemotherapy supervision model for rural cancer care: perspectives of health professionals. Eur J Cancer Care (Engl) 2016;25(1):93-98. [doi: 10.1111/ecc.12309] [Medline: 25871852]
30. Bobb MR, Van Heukelom PG, Faine BA, Ahmed A, Messerly JT, Bell G, et al. Telemedicine provides noninferior research informed consent for remote study enrollment: a randomized controlled trial. Acad Emerg Med 2016;23(7):759-765 [FREE Full text] [doi: 10.1111/acem.12966] [Medline: 26990899]
31. Clark JM, Heifetz LJ, Palmer D, Brown LM, Cooke DT, David EA. Telehealth allows for clinical trial participation and multimodality therapy in a rural patient with stage 4 non-small cell lung cancer. Cancer Treat Res Commun 2016;9:139-142 [FREE Full text] [doi: 10.1016/j.ctarc.2016.09.005] [Medline: 28580436]
32. Doolittle GC, Harmon A, Williams A, Allen A, Boysen CD, Wittman C, et al. A cost analysis of a tele-oncology practice. J Telemed Telecare 1997;3 Suppl 1:20-22. [doi: 10.1258/1357633971930850] [Medline: 9218371]
33. Li BT, Daly B, Gospodarowicz M, Bertagnolli MM, Brawley OW, Chabner BA, et al. Reimagining patient-centric cancer clinical trials: a multi-stakeholder international coalition. Nat Med 2022;28(4):620-626. [doi: 10.1038/s41591-022-01775-6] [Medline: 35440725]
34. Daly B, Cracchiolo J, Holland J, Ebstein AM, Flynn J, Duck E, et al. Digitally enabled transitional care management in oncology. JCO Oncol Pract 2024;20(5):657-665. [doi: 10.1200/OP.23.00565] [Medline: 38382002]
35. Leidong W, Monachino M, Lloyd-Williams D, Nguyen TLH, Dsouza B, Moreira JP. Mobile apps for cancer patients: identifying positive impacts and concerns. Digit Health 2025;11:20552076241305707 [FREE Full text] [doi: 10.1177/20552076241305707] [Medline: 40041393]
36. Gresham G, Hendifar AE, Spiegel B, Neeman E, Tuli R, Rimel BJ, et al. Wearable activity monitors to assess performance status and predict clinical outcomes in advanced cancer patients. NPJ Digit Med 2018;1:27 [FREE Full text] [doi: 10.1038/s41746-018-0032-6] [Medline: 31304309]
37. Keats MR, Yu X, Sweeney Magee M, Forbes CC, Grandy SA, Sweeney E, et al. Use of wearable activity-monitoring technologies to promote physical activity in cancer survivors: challenges and opportunities for improved cancer care. Int J Environ Res Public Health 2023;20(6):4784 [FREE Full text] [doi: 10.3390/ijerph20064784] [Medline: 36981693]
38. Emard N, Lynch KA, Liou KT, Atkinson T, Green AK, Daly B, et al. Virtual mind-body programming for patients with cancer during the COVID-19 pandemic: qualitative study. JMIR Cancer 2021;7(2):e27384 [FREE Full text] [doi: 10.2196/27384] [Medline: 33882018]
39. Trevino KM, Raghunathan N, Latte-Naor S, Polubriaginof FCG, Jensen C, Atkinson TM, et al. Rapid deployment of virtual mind-body interventions during the COVID-19 outbreak: feasibility, acceptability, and implications for future care. Support Care Cancer 2021;29(2):543-546 [FREE Full text] [doi: 10.1007/s00520-020-05740-2] [Medline: 32902712]
40. Mao JJ, Bryl K, Gillespie EF, Green A, Hung TKW, Baser R, et al. Randomized clinical trial of a digital integrative medicine intervention among patients undergoing active cancer treatment. NPJ Digit Med 2025;8(1):29 [FREE Full text] [doi: 10.1038/s41746-024-01387-z] [Medline: 39809874]
41. Lazarou I, Krooupa A, Nikolopoulos S, Apostolidis L, Sarris N, Papadopoulos S, et al. Cancer patients' perspectives and requirements of digital health technologies: a scoping literature review. Cancers (Basel) 2024;16(13):2293 [FREE Full text] [doi: 10.3390/cancers16132293] [Medline: 39001356]
42. Curry J, Patterson M, Greenley S, Pearson M, Forbes CC. Feasibility, acceptability, and efficacy of online supportive care for individuals living with and beyond lung cancer: a systematic review. Support Care Cancer 2021;29(11):6995-7011 [FREE Full text] [doi: 10.1007/s00520-021-06274-x] [Medline: 34008080]
43. Madanian S, Nakarada-Kordic I, Reay S, Chetty T. Patients' perspectives on digital health tools. PEC Innov 2023;2:100171 [FREE Full text] [doi: 10.1016/j.pecinn.2023.100171] [Medline: 37384154]
44. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
45. Braun V, Clarke V. Thematic Analysis: A Practical Guide. London, UK: SAGE Publications; 2021.
46. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
47. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
48. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

XSL•FO
RenderX

49. Singh B, Olds T, Curtis R, Dumuid D, Virgara R, Watson A, et al. Effectiveness of physical activity interventions for improving depression, anxiety and distress: an overview of systematic reviews. Br J Sports Med 2023 Sep;57(18):1203-1209 [FREE Full text] [doi: 10.1136/bjsports-2022-106195] [Medline: 36796860]
50. Peng Y, Zhang K, Wang L, Peng Y, Liu Z, Liu L, et al. Effect of a telehealth-based exercise intervention on the physical activity of patients with breast cancer: a systematic review and meta-analysis. Asia Pac J Oncol Nurs 2022;9(12):100117 [FREE Full text] [doi: 10.1016/j.apjon.2022.100117] [Medline: 36276881]
51. Mishra SI, Scherer RW, Snyder C, Geigle PM, Berlanstein DR, Topaloglu O. Exercise interventions on health-related quality of life for people with cancer during active treatment. Cochrane Database Syst Rev 2012;2012(8):CD008465 [FREE Full text] [doi: 10.1002/14651858.CD008465.pub2] [Medline: 22895974]
52. Yi L, Tian X, Jin Y, Luo M, Jiménez-Herrera MF. Effects of yoga on health-related quality, physical health and psychological health in women with breast cancer receiving chemotherapy: a systematic review and meta-analysis. Ann Palliat Med 2021;10(2):1961-1975 [FREE Full text] [doi: 10.21037/apm-20-1484] [Medline: 33440970]
53. Xunlin NG, Lau Y, Klainin-Yobas P. The effectiveness of mindfulness-based interventions among cancer patients and survivors: a systematic review and meta-analysis. Support Care Cancer 2020;28(4):1563-1578. [doi: 10.1007/s00520-019-05219-9] [Medline: 31834518]
54. Atkinson TM, Liou KT, Borten MA, Li QS, Popkin K, Webb A, et al. Association between music therapy techniques and patient-reported moderate to severe fatigue in hospitalized adults with cancer. JCO Oncol Pract 2020;16(12):e1553-e1557 [FREE Full text] [doi: 10.1200/OP.20.00096] [Medline: 32639926]
55. Wayne PM, Lee MS, Novakowski J, Osypiuk K, Ligibel J, Carlson LE, et al. Tai chi and qigong for cancer-related symptoms and quality of life: a systematic review and meta-analysis. J Cancer Surviv 2018;12(2):256-267 [FREE Full text] [doi: 10.1007/s11764-017-0665-5] [Medline: 29222705]
56. Mustian KM, Sprod LK, Palesh OG, Peppone LJ, Janelsins MC, Mohile SG, et al. Exercise for the management of side effects and quality of life among cancer survivors. Curr Sports Med Rep 2009;8(6):325-330 [FREE Full text] [doi: 10.1249/JSR.0b013e3181c22324] [Medline: 19904073]
57. Gerritsen JKW, Vincent AJPE. Exercise improves quality of life in patients with cancer: a systematic review and meta-analysis of randomised controlled trials. Br J Sports Med 2016;50(13):796-803. [doi: 10.1136/bjsports-2015-094787] [Medline: 26719503]
58. Rodríguez-Cañamero S, Cobo-Cuenca AI, Carmona-Torres JM, Pozuelo-Carrascosa DP, Santacruz-Salas E, Rabanales-Sotos JA, et al. Impact of physical exercise in advanced-stage cancer patients: systematic review and meta-analysis. Cancer Med 2022;11(19):3714-3727 [FREE Full text] [doi: 10.1002/cam4.4746] [Medline: 35411694]
59. Bailey LE, Morris MA. Mind-body therapies adjuvant to chemotherapy improve quality of life and fatigue in top cancers: a systematic review and meta-analysis. Complement Ther Clin Pract 2024;54:101811 [FREE Full text] [doi: 10.1016/j.ctcp.2023.101811] [Medline: 38029633]
60. Carlson LE, Zelinski E, Toivonen K, Flynn M, Qureshi M, Piedalue K, et al. Mind-body therapies in cancer: what is the latest evidence? Curr Oncol Rep 2017;19(10):67. [doi: 10.1007/s11912-017-0626-1] [Medline: 28822063]
61. Nipp RD, El-Jawahri A, Fishbein JN, Eusebio J, Stagl JM, Gallagher ER, et al. The relationship between coping strategies, quality of life, and mood in patients with incurable cancer. Cancer 2016;122(13):2110-2116 [FREE Full text] [doi: 10.1002/cncr.30025] [Medline: 27089045]
62. McCorkle R, Ercolano E, Lazenby M, Schulman-Green D, Schilling LS, Lorig K, et al. Self-management: enabling and empowering patients living with cancer as a chronic illness. CA Cancer J Clin 2011;61(1):50-62 [FREE Full text] [doi: 10.3322/caac.20093] [Medline: 21205833]
63. Dev R, Agosta M, Fellman B, Reddy A, Baldwin S, Arthur J, et al. Coping strategies and associated symptom burden among patients with advanced cancer. Oncologist 2024;29(2):166-175 [FREE Full text] [doi: 10.1093/oncolo/oyad253] [Medline: 37669020]
64. Maita KC, Maniaci MJ, Haider CR, Avila FR, Torres-Guzman RA, Borna S, et al. The impact of digital health solutions on bridging the health care gap in rural areas: a scoping review. Perm J 2024;28(3):130-143 [FREE Full text] [doi: 10.7812/TPP/23.134] [Medline: 39135461]
65. Gorzelitz JS, Bouji N, Stout NL. Program barriers and facilitators in virtual cancer exercise implementation: a qualitative analysis. Transl J Am Coll Sports Med 2022;7(3) [FREE Full text] [doi: 10.1249/tjx.0000000000000199] [Medline: 35601444]
66. Shah A, Hussain-Shamsy N, Strudwick G, Sockalingam S, Nolan RP, Seto E. Digital health interventions for depression and anxiety among people with chronic conditions: scoping review. J Med Internet Res 2022;24(9):e38030 [FREE Full text] [doi: 10.2196/38030] [Medline: 36155409]
67. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. N Engl J Med 2017;377(16):1585-1592. [doi: 10.1056/NEJMsr1503323] [Medline: 29045204]
68. Pasek M, Suchocka L, Gąsior K. Model of social support for patients treated for cancer. Cancers (Basel) 2021;13(19):4786 [FREE Full text] [doi: 10.3390/cancers13194786] [Medline: 34638270]

XSL•FO
RenderX

69. Ruiz-Rodríguez I, Hombrados-Mendieta I, Melguizo-Garín A, Martos-Méndez MJ. The importance of social support, optimism and resilience on the quality of life of cancer patients. Front Psychol 2022;13:833176 [FREE Full text] [doi: 10.3389/fpsyg.2022.833176] [Medline: 35356348]
70. Montalescot L, Baussard L, Charbonnier E. Factors associated with digital intervention engagement and adherence in patients with cancer: systematic review. J Med Internet Res 2024;26:e52542 [FREE Full text] [doi: 10.2196/52542] [Medline: 39661976]
71. Ellen ME, Shach R, Balicer RD. Helping patients help themselves: supporting the healthcare journey. Patient Educ Couns 2018 Sep;101(9):1708-1711. [doi: 10.1016/j.pec.2018.04.005] [Medline: 29691111]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
HIPAA: Health Insurance Portability and Accountability Act
IM@Home: Integrative Medicine at Home
IMPROVE: Integrative Medicine for Patient-reported Outcomes, Values, and Experience
IRB: Institutional Review Board
MSK: Memorial Sloan Kettering Cancer Center

Edited by N Cahill; submitted 03.Jul.2025; peer-reviewed by C Donachie; comments to author 11.Nov.2025; revised version received 19.Nov.2025; accepted 16.Dec.2025; published 16.Jan.2026.

Please cite as:
Bryl KL, Silverwood S, Desai K, Schobert K, Li X, Chimonas S, Mao JJ, Gillespie EF
Benefits and Challenges of a Digital Exercise and Mind-Body Program During Active Cancer Treatment: Qualitative Study of Patients' Perceptions
JMIR Cancer 2026;12:e80075
URL: https://cancer.jmir.org/2026/1/e80075
doi:10.2196/80075
PMID:

©Karolina L Bryl, Sierra Silverwood, Krupali Desai, Kelsey Schobert, Xiaotong Li, Susan Chimonas, Jun J Mao, Erin F Gillespie. Originally published in JMIR Cancer (https://cancer.jmir.org), 16.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Surrogate Decision-Making by Family Caregivers for Hyperthermic Intraperitoneal Chemotherapy in Gastric Cancer: Qualitative Study in a High-Volume Chinese Center

Zheng-Ke-Ke Tan[1], MSN; Dan-Ni Li[1], MSN; Kui Jia[2], MSN; Wen-Zhen Tang[1], PhD; Xin Chen[1], MSN; Li Yang[1], BS

[1]Nursing Department, First Affiliated Hospital of GuangXi Medical University, Nanning, China

[2]Gastrointestinal Surgery Department, First Affiliated Hospital of GuangXi Medical University, Nanning, China

Corresponding Author:
Li Yang, BS
Nursing Department
First Affiliated Hospital of GuangXi Medical University
Number 6, Shuangyong Road, Qingxiu District
Nanning
China
Phone: 86 13607814589
Email: m15177776539@163.com

Abstract

Background: Hyperthermic intraperitoneal chemotherapy (HIPEC) has been integrated into the management of gastric cancer (GC) as a combined approach for addressing peritoneal metastasis, serving both prophylactic and therapeutic roles following GC surgery. The pivotal decision regarding HIPEC administration typically arises intraoperatively, creating a complex clinical scenario where family caregivers must act as surrogate decision-makers under substantial time constraints. This decision-making process proves particularly challenging due to limited understanding of the procedure's risk-benefit profile and long-term outcomes among nonmedical surrogates, challenges often exacerbated by the acute stress of the surgical environment.

Objective: This qualitative study aims to explore how family caregivers of patients with GC navigate the HIPEC decision-making process, specifically examining the facilitators, challenges, and the role of information acquisition that shape the shared decision-making mode.

Methods: This study adopted a qualitative approach using semistructured interviews; 15 family caregivers of patients with GC in a major tertiary hospital in Guangxi Province were selected as research objects through a purposive sampling method. Participants were asked to comment on their experience of surrogate decision-making for the HIPEC process. The Colaizzi 7-step method was used to analyze and summarize the themes.

Results: The mean age of the 15 participants was 39.8 (SD 13.29, range 20-68) years, and all patients were on average aged 56.7 (SD 10.78, range 36-74) years. The relationship to the patient was distributed as follows: 33% (5/15) spouses, 60% (9/15) children, and 6% (1/15) other relatives. Four major themes emerged from the data analysis: (1) shared decision-making participation mode (doctor-led passive decision-making and doctor-family shared decision-making); (2) decision-information sources (decision-making information came from medical-care personnel, decision-making information came from the internet, and decision-making information came from acquaintances); (3) challenges in the decision-making process (financial burden and anticipated therapeutic efficacy); and (4) facilitator in the decision-making process (positive health beliefs and cultural dimensions of perceived responsibility: a Confucian perspective).

Conclusions: HIPEC decision-making by family caregivers of patients with GC was primarily passive decision-making, and many obstacles and facilitators were encountered in the process. Medical staff should share information and encourage and guide family caregivers to participate in the decision-making process through decision assistance or decision support.

(JMIR Cancer 2026;12:e80471) doi:10.2196/80471

KEYWORDS

hyperthermic intraperitoneal chemotherapy; gastric cancer; family caregivers; surrogate decision-making; qualitative study

XSL•FO
RenderX

Introduction

Gastric cancer (GC) is 1 of the 5 most common malignant tumors in the world [1], and the cumulative risk of peritoneal metastasis (PM) is nearly 30%. It is the most common form of GC metastasis [2,3]. Systemic chemotherapy drugs have difficulty penetrating the peritoneal barrier; PM treatment is difficult, and the disease progresses rapidly. Once PM occurs, the mortality of patients is very high, which has become the main cause of death in the literature.

Developments have been made in the field of peritoneal tumors over the last decade, and hyperthermic intraperitoneal chemotherapy (HIPEC) is an emerging treatment technique. It heats an infusion-containing chemotherapy drug to the therapeutic temperature and is reinjected into the patient's abdominal cavity for a certain period of time. The main principle of HIPEC is hyperthermia, chemotherapy, and mechanical scour, which can break through the plasma-peritoneal barrier, remove the residual tissue, and free cancer cells after surgery [4,5]. HIPEC prevents PM based on radical gastrectomy and can also be combined with cytoreductive surgery to treat selected patients with resectable primary or secondary peritoneal malignancy [6,7]. The median survival of patients with advanced GC can reach 18.1 months after HIPEC [8]. However, there is some controversy about the safety and effectiveness of HIPEC. A meta-analysis reported that HIPEC failed to improve the 1-3 year overall survival and gastrointestinal complication rates, but improved long-term overall survival [9]. From the perspective of health literacy popularization, a small percentage of the public has heard of or understood HIPEC, which is a complex operation [10]. Treatment decisions in HIPEC are individualized and require patient and physician involvement. Therefore, the difficulty for patients is to question whether the disease can be successfully treated, as HIPEC is unknown. These concerns and questions can run from decision-making to treatment.

Surrogate decision-makers are more common in intensive care units, ignoring the fact that patients with cancer may also need family caregivers to make surrogate decisions [11,12]. Family caregivers often participate in decision-making, and even become the main decision-maker, when some patients with GC may have adverse psychological conditions, physical discomfort resulting in reduced decision-making ability, or surgeons need to put forward HIPEC-related treatment recommendations according to intraoperative conditions. Although the safety of HIPEC has been proven, it is associated with higher mortality, lower overall treatment morbidity, and reduced quality of life in the short term [13]. People sometimes make different medical decisions for others than they would make for themselves [14]. Family caregivers are usually concerned that surrogate decision-making will influence the patient's prognosis and lack of HIPEC experience, thus they may fall into decision-making difficulties and conflicts, resulting in possible psychological pressures such as regret, anxiety, contradiction, and unease [11]. Negative emotions will again affect the implementation of decision-making and reduce the satisfaction of patients and caregivers with medical treatment, forming a vicious circle [15].

Previous studies have been conducted to explore the surrogate decision-making process from diagnosis to treatment in patients' relatives with digestive tract tumors because their lack of understanding of treatment information leads to passive decision-making [16]. HIPEC is a special treatment for gastrointestinal tumors, and most studies focus on prognosis and adverse reactions [13]. Conversely, relatively few studies have focused on the surrogate decision-making experience of HIPEC, which may be due to the underrepresentation of families of patients with GC in HIPEC clinical trials. Different surrogate decision-makers have different views on health, disease, and participation mode of treatment, which are important for whether patients can obtain better treatment plans in time [12].

This study aims to identify the decision-making participation mode of surrogate decision-makers for patients with GC considering HIPEC, particularly their stance on the spectrum of paternalistic, physician-led decision-making and collaborative, shared decision-making, while elucidating the multilevel challenges and facilitators underlying this process. Our results can serve as a reference for formulating targeted decision-support strategies and guiding doctor-patient communication, thereby improving HIPEC decision-making efficiency, safeguarding patients' physical and mental health, and enhancing patient satisfaction with medical treatment.

Methods

Design

An exploratory qualitative study involved semistructured interviews with family caregivers of patients with GC treated with HIPEC who acted as the primary decision-makers. From January to April 2024, family caregivers of patients with GC undergoing abdominal HIPEC after gastrointestinal gland surgery in a major tertiary hospital in Guangxi were selected as the research subjects. We used face-to-face interviews to understand the real experience of family caregivers' HIPEC decision-making process. This study followed the comprehensive criteria for the reporting of qualitative research guidelines (Multimedia Appendix 1) [17].

Study Setting

This study was conducted in the Gastrointestinal Surgery Department of the First Affiliated Hospital of Guangxi Medical University, a national key clinical construction specialty, a key medical and health discipline of Guangxi, and the chief unit of the Guangxi Gastric Cancer Alliance, focusing on the diagnosis and treatment of GC. Through the combination of multiple disciplines, the department carries out accelerated rehabilitation surgery for GC and promotes the rapid rehabilitation of patients with GC. The related technology of accelerated rehabilitation surgery is in a leading position in the country. Doctors evaluate the need for HIPEC before or during surgery, and the patient or family decides whether to proceed with HIPEC after surgery. All patients underwent HIPEC within 48 h after closing the abdominal cavity. The perfused temperature was 43 °C, and each operation time was 60-90 minutes.

XSL•FO
RenderX

Research Team and Interviewer Characteristics

This study's team comprised 6 health care providers, namely, 3 postgraduate students (ZKKT, DNL, and XC), 2 specialists in gastrointestinal surgery nursing (KJ and WZT), and 1 expert in the field of oncology nursing (LY). All interviews were conducted by 2 primary researchers (ZKKT and DNL), both are female and holding MSc Nursing degrees, with over 3 years of dedicated research and academic training in oncology. Before data collection, the interviewers completed formal training in qualitative methodology, including specialized instruction in interview techniques, probing strategies, and ethical considerations. To minimize potential bias, neither researcher had established clinical relationships with any of the participants' families. This absence of prior therapeutic engagement was explicitly disclosed to all participants at the commencement of each interview session [17]. This study was supervised by KJ, a professor of nursing with extensive research experience in gastrointestinal oncology. The project was led by author LY, Director of Nursing at a major tertiary hospital in Southern China.

Recruitment and Sample

The purposive sampling method was used to recruit the participants. Participants were surrogates of adult patients with GC. We approached surrogates of nondecisional patients who required or were expected to require HIPEC after surgery. To be eligible to participate, surrogates needed to be at least 18 years of age, be fluent in the Chinese language, have clear awareness with sufficient communication skills, and self-identify as a primary decision-maker for the patient. Exclusion criteria included extraperitoneal disease, ascites >500 mL, physiologic inability to undergo HIPEC [5], surrogates of patients with recurrent disease, or those who were not involved in the treatment decision. This approach ensured that all participants had direct experience with the phenomenon of interest. Potential eligible surrogates were screened through electronic health records. After the patient completed the HIPEC treatment, the primary nurses of the patients in charge invited the respondent and confirmed eligibility. Participants were informed that their participation was entirely voluntary and that they could opt out at any time. All participant records were kept confidential.

Sample size was based on data saturation, and the calculation and evaluation methods included 3 elements, namely, base size, run length, and new information threshold [18]. Recruitment was stopped when the 2 researchers considered that the content was saturated and without new aspects emerging. Base size as denominator, the run length was the number of interviews within which we looked for and calculated new information. The new information threshold represents the proportion of new information. We will prospectively calculate saturation using a base size of 4 interviews and a run length of 2 interviews. We have selected a new information threshold of ≤5% to indicate that we have reached adequate saturation.

Data Collection

Interviews were conducted in a quiet hospital room, a doctor's office, or a classroom. The interviewers first introduced the purpose, significance, and content of the interview to the interviewees and obtained their consent. Before the interview, interviewees were required to fill in a self-made questionnaire, which included demographic variables (age, gender, education level, occupation, relationship between family caregivers, and patients) of patients and their families, as well as disease-related information (HIPEC protocol and tumor stage) of patients.

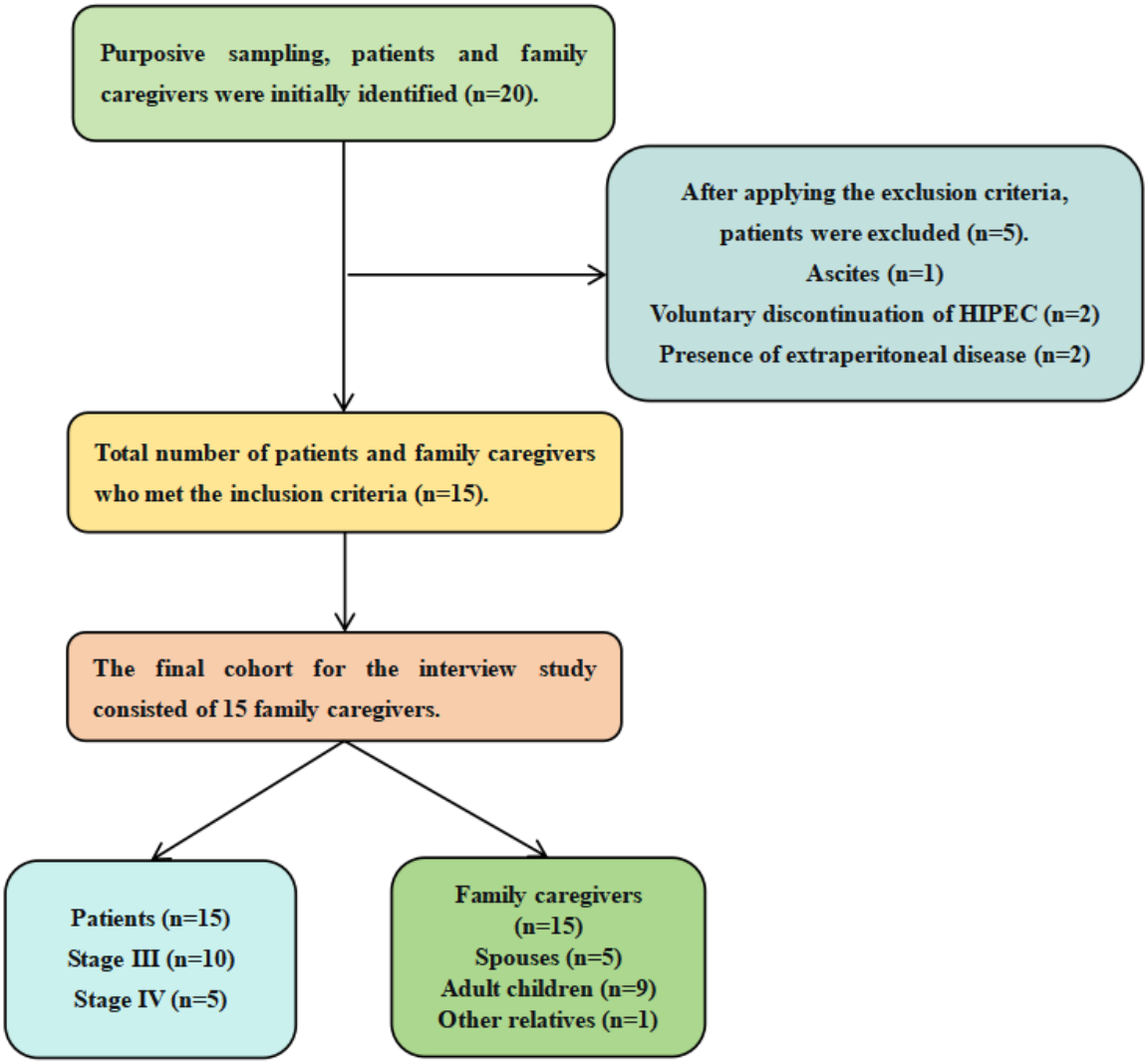
They conducted a one-to-one semistructured interview with the interviewees according to the final guide of the interview (Textbox 1), which was developed and tested in 2 preinterviews whose interview data were not included in the data analysis. During the process, they paid attention to observe and record the changes in the nonverbal communication of the interviewees. They were encouraged to express their real thoughts and ask questions. All interviews were audio recorded, transcribed verbatim, and compared with digital audio recordings to ensure the accuracy of the content [17]. Unclear or doubtful contents expressed by the interviewees were confirmed through retelling, questioning, and other ways. Induction, intervention, or judgment was prohibited. The interview duration was limited to approximately 30 minutes per participant. A recruitment diagram appears in Figure 1.

Textbox 1. Qualitative interview guide.

- I'd like to start by asking how did you decide to receive hyperthermic intraperitoneal chemotherapy (HIPEC) on behalf of the patient.
- Can you recall the process of treatment decision?
- What do you know about HIPEC?
- What made you decide to receive HIPEC?
- What concerns did you have during the decision-making process?
- Is there anything you'd like to know or want the medical staff to help you with?

XSL•FO
RenderX

Figure 1. Study enrollment diagram. HIPEC: hyperthermic intraperitoneal chemotherapy.



Data Analysis

The audio recordings from each interview were transcribed verbatim. Within 24 h after the interview, 2 researchers (ZKKT and DNL) independently organized the interview data, transcribed it into text, coded transcripts, and analyzed the data with the Colaizzi 7-step analysis method [19]: The interview data were read repeatedly to deeply understand the content information; code recurring, meaningful statements; form a categorized coded view; write a detailed description; have sublimation view as the theme; and then to return the written text to the interviewee for verification. The 2 researchers were fully immersed in the interview content, independently reading and coding in detail. It was summarized after completion, and if any objection was raised, they were discussed and determined within the team (ZKKT, DNL, XC, and WZT). During this process, the research team adopted triangulation, randomly selected 20% transcripts and codes, and then submitted them to the noninterviewers for reviewing and improving the codes [20]. Then, regular meetings were held within the team to generate initial themes (ZKKT, DNL, XC, and WZT).

Ethical Considerations

This study was approved by the Ethics Committee of the First Affiliated Hospital of Guangxi Medical University (2024-E055-01). Before the interview, the interviewer introduced the purpose, significance, and content of the research to the subjects. All data were used only for research purposes. All participants signed informed consent forms after expressing their willingness to participate. During the interview, participants were free to interrupt the experiment if they chose to do so. Immediately following data collection, the interviewer engaged all participants in a standardized debriefing process to disclose the sham trial. Participants were not provided compensation for participation. All personal identifiers of the participants were removed before analysis to ensure privacy. This study strictly adheres to Beauchamp and Childress's Four Principles of Biomedical Ethics [21] to establish the theoretical framework for our ethical approach, recognizing that decision-making during a patient's surgical procedure inherently precludes the preservation of their decision-making capacity.

XSL•FO
RenderX

Results

Patient and Surrogate Decision-Maker Characteristics

The analysis sample for data saturation included 15 patients and their surrogate decision-makers. All participants were on average 39.8 (SD 13.29, range 20-68) years of age at the time of interview. The majority were male (11/15, 73%), and 27% (4/15) were female. The relationship to the patient was distributed as follows: 33% (5/15) spouses, 60% (9/15) children, and 6% (1/15) other relatives. All patients were on average 56.7 (SD 10.78, range 36-74) years of age at the time of interview. The demographic and clinical characteristics of these participants and patients are presented in Table 1.

Table 1. Characteristics of participants.

Number	Surrogate decision-maker					Patient					
	Gender	Age (years)	Education	Occupation	Relationship with patients	Gender	Age (years)	Medical insurance	HIPEC[a], n	Stage	Chemotherapeutics drugs
1	Male	43	Middle school	Farmer	Husband	Female	46	Medical insurance for residents	3	III[b]	Lobaplatin
2	Female	36	Junior college	Office clerk	Wife	Male	36	Medical insurance for residents	3	III	Docetaxel
3	Male	20	College	Student	Son	Male	53	No medical insurance	5	IV[c]	Lobaplatin
4	Male	55	Middle school	Farmer	Husband	Female	51	Medical insurance for residents	3	III	Oxaliplatin
5	Male	45	Master	Construction industry manager	Son	Female	67	Medical insurance for residents	3	IV	Lobaplatin
6	Male	32	College	Office clerk	Son	Male	60	Medical insurance for residents	3	IV	Docetaxel
7	Male	33	College	Civil servant	Son	Male	60	Medical insurance for residents	2	III	Oxaliplatin
8	Female	45	Primary school	Farmer	Wife	Male	47	Medical insurance for residents	3	IV	Docetaxel
9	Male	26	Junior college	Office clerk	Son	Male	54	Medical insurance for residents	1	III	Docetaxel
10	Female	49	College	Civil servant	Daughter	Male	74	Medical insurance for urban employees	3	III	Docetaxel
11	Male	51	High school	Worker	Son	Male	73	Medical insurance for residents	3	III	Docetaxel
12	Male	45	Junior college	Individual	Full brother	Male	51	Medical insurance for residents	3	III	Lobaplatin
13	Male	26	High school	Farmer	Son	Male	58	Medical insurance for residents	1	IV	Oxaliplatin
14	Female	24	Junior college	Individual	Daughter	Male	51	Medical insurance for residents	3	III	Oxaliplatin
15	Male	68	Junior college	Retiree	Husband	Female	70	Medical insurance for urban employees	1	III	Docetaxel

[a]HIPEC: hyperthermic intraperitoneal chemotherapy.

[b]III: stage III gastric cancer is categorized as locally advanced disease. In this stage, the primary tumor has invaded deeply through the gastric wall (potentially involving the serosa) and has extensively metastasized to the regional lymph nodes.

[c]IV: stage IV represents the most advanced phase of gastric cancer. It is defined by the presence of distant metastasis (M1), regardless of the extent of the primary tumor (Any T) or the involvement of regional lymph nodes (Any N).

Topic

The researchers identified four themes through their analysis: (1) shared decision-making participation mode, (2) decision-information sources, (3) challenges in the decision-making process, and (4) facilitators in the decision-making process. The themes and subthemes derived from the analysis are shown in Table 2.

XSL•FO
RenderX

Table 2. Themes and subthemes.

Themes	Subthemes
1. Shared decision-making participation mode	• Doctor-led passive decision-making • Doctor-family sharing decision-making
1. Decision-information sources	• Decision-making information comes from the medical-care personnel • Decision-making information comes from the internet • Decision-making information comes from acquaintances
1. Challenges in the decision-making process	• The financial burden • Anticipated therapeutic efficacy
1. Facilitator in the decision-making process	• Positive health beliefs • Cultural dimensions of perceived responsibility: a Confucian perspective

Theme 1: Shared Decision-Making Participation Mode

Overview

The way of shared decision-making was primarily based on the level of information mastery and education of decision-makers.

Doctor-Led Passive Decision-Making

Nine surrogate decision-makers expressed a preference for physician-dominated decision-making, indicating that their role was primarily to adhere to the physician's recommended treatment course.

The adoption of this decision-making approach demonstrates a significant correlation with educational attainment. Surrogates with limited formal education exhibited a substantial knowledge gap regarding HIPEC compared to clinical professionals, while simultaneously expressing stronger deference to hierarchical medical authority.

> *All of us live in rural areas. I, my wife, my younger brother and my younger sister all have only a middle school education. They seldom go to the hospital for medical treatment. I had never heard of HIPEC and had to rely on the doctor to make the decision. I'm at a loss right now.* [Number 4, aged 55 years, farmer, husband of patient]
>
> *I never considered challenging the doctor's decision. Since the treatment method they proposed was effective, I followed the expert's advice. After all, he was respected authorities whose opinions are widely accepted.* [Number 13, aged 26 years, farmer, son of patient]

From a cognitive perspective, many surrogates reported experiencing significant feelings of inadequacy and information overload. When participants acquired information but demonstrated either an unwillingness to engage in multisource verification or an inability to synthesize information from disparate channels, they consistently reverted to physician-dependent decision-making as a default strategy.

> *Several doctors came and told me a lot of information. I managed to understand a little bit, but it was too complicated. Without knowledge of HIPEC, the patient was lying on the bed, unable to move, unable to make decisions. We do what the doctor recommends completely at this time.* [Number 11, aged 50 years, worker, son of patient]
>
> *Although the doctor is obligated to inform us of this information, I don't think I can fully understand it. I don't worry about whether I can understand it or not. All I need to know is that the doctor's decision is correct, and I just need to follow his decision.* [Number 1, aged 43 years, farmer, husband of patient]

Doctor-Family Shared Decision-Making

Six surrogate decision-makers combined information obtained from various sources. Family caregivers who engaged in doctor-family shared decision-making typically possessed higher educational attainment, demonstrating not only comprehension of complex HIPEC information but also the capacity to critically evaluate disparate information sources, including both physician-provided guidance and internet-acquired content.

> *At the beginning, after being informed of HIPEC by the clinical doctor, first checked the Internet, but also asked acquaintances, to understand the thermal perfusion technology can prevent metastasis, avoid risk effect, we decided to carry out HIPEC.* [Number 5, aged 45 years, construction-industry manager, son of patient]
>
> *On the premise of knowing a lot about the principle, knowledge and effect of thermoperfusion, I took the initiative to choose the best hospital recognized for surgery and HIPEC.* [Number 6, aged 32 years, office clerk, son of patient]

These surrogate decision-makers did not passively await information but rather demonstrated a proactive communication stance, exemplified by actively seeking clarifications during consultations and engaging in iterative questioning with physicians. They perceived clinicians as collaborative partners in problem-solving rather than unidirectional authorities.

> *After consulting with acquaintances, I repeatedly confirmed with the doctor whether changing the treatment plan could cure gastric cancer and also determined whether the quality of life would change in the future.* [Number 5, aged 45 years, construction-industry manager, son of patient]

XSL•FO
RenderX

After I came to this hospital now, I will also ask the doctor to check the information I found online. Even though they urged me, I couldn't make up my mind easily. [Number 6, aged 32 years, office clerk, son of patient]

Theme 2: Decision-Information Sources

Decision-Making Information Came From Medical Care Personnel

All participants indicated that medical personnel were the initial source of information for their HIPEC decision-making. As a new therapeutic modality, HIPEC has a low popularity among the public, with less access, and obtaining relevant information initially in daily life was difficult for patients and their families.

I've never heard of HIPEC, and I don't know how it works. The first time I heard HIPEC, it was introduced by the surgeon. [Number 2, aged 36 years, office clerk, wife of patient]

I usually read disease-related information on social platforms, but rarely mention HIPEC, and only know it from doctors. [Number 10, aged 49 years, civil servant, daughter of patient]

The main source of HIPEC information for 10 (10/15, 66.6%) surrogate decision-makers of patients was medical staff. A total of 6 surrogate decision-makers did not subsequently seek information in other ways, so this became their only access to information. In particular, older decision-makers living in rural areas lack the competence and access to information.

We rural people do not know this, only through the doctor's recommendation, and do not look up other information. [Number 1, aged 43 years, farmer, husband of patient]

Born in a rural area and an elderly person, I rarely use a smartphone, so I only know about it from my doctor. [Number 4, aged 55 years, farmer, husband of patient]

Decision-Making Information Came From the Internet

Some surrogate decision-makers indicated that they actively sought information and confirmation through internet channels, these being by those with higher education or high internet usage. The primary information sought by participants concerned 3 key areas: the mechanism of HIPEC treatment, its efficacy in GC, and potential adverse effects.

After the doctor first told me HIPEC, I searched on the Internet myself to simply understand what therapeutic effect it has on GC. I looked it up through wechat public account and Baidu. [Number 2, aged 36 years, office clerk, wife of patient]

I searched for side effects (of HIPEC) through Baidu. There are other doctors on Baidu, they might give me different advice. [Number 7, aged 33 years, civil servant, son of patient]

Furthermore, the field of HIPEC is highly specialized, and some of the family caregivers of patients have difficulties understanding it.

But I have no knowledge reserve, even if I find relevant information, I cannot understand. [Number 7, aged 33 years, civil servant, son of patient]

Decision-Making Information Came From Acquaintances

In addition to seeking information through the internet, 3 decision-makers consulted expert acquaintances. For unknown professional treatment information, patients' families tended to consult other doctors through trusted relatives and friends. Surrogate decision-makers find it easier to hear the truth from close friends, to obtain better judgment, and ensure information credibility.

I have relatives and friends who are doctors, they asked someone who is an expert in the field of stomach cancer to know if HIPEC can help my mother's condition. [Number 5, aged 45 years, construction-industry manager, son of patient]

I asked the doctor at our local hospital about the effect of HIPEC and whether it was harmful to the patient's health. After all, it was the first time I heard about this treatment, and I was still a little afraid. [Number 6, aged 32 years, office clerk, son of patient]

Theme 3: Challenges in the Decision-Making Process

Financial Burden

An overarching theme from the interviews was that the significant economic pressure associated with postoperative HIPEC presented a major decisional impediment for nearly all participants. This financial strain was frequently cited as a direct barrier to completing the full course of recommended therapy.

As ordinary people, the economic pressure is relatively large, the thermoperfusion drainage tube is completely self-funded with thousands of yuan each time. I hope the society can be helpful in this respect. [Number 11, aged 51 years, worker, son of patient]

The price is very expensive and the medical insurance is not reimbursed, our average family salary is not high. what if the treatment results are not proportional to our efforts? [Number 9, aged 26 years, office clerk, son of patient]

Anticipated Therapeutic Efficacy

First, a notable uncertainty was noted among some participants concerning the expected benefits of HIPEC, particularly pertaining to the therapeutic protocol and its effectiveness in preventing metastasis and recurrence. This ambiguity was rooted in their nonprofessional background, which consequently hindered their comprehension of the proposed HIPEC regimen and led to profound decisional uncertainty.

The doctor and I have discussed whether to do it three times or four times before, but we still don't know what effect the number of times has on the efficacy, and we are not sure whether the patient can complete the radical cure or prevent metastasis after completing HIPEC, This makes me quite unsure

XSL•FO
RenderX

whether I really need it. [Number 2, aged 36 years, office clerk, wife of patient]

I mainly want to know what effect can be achieved by doing three times, although the doctor always said that the effect is not guaranteed, I hope you can take the initiative to give us a clear answer. [Number 5, aged 45 years, construction-industry manager, son of patient]

Second, some participants, after an in-depth understanding of HIPEC, were concerned that the new technology in chemotherapy had unknown risks of damaging health or aggravating conditions.

HIPEC is also a form of chemotherapy that kills cancer cells and seems to kill healthy cells, raising concerns that patients may suffer weakened immunity and other after-effects. [Number 2, aged 36 years, office clerk, wife of patient]

I'm still worried about any adverse reactions or accidents with this new technology. [Number 10, aged 49 years, civil servant, daughter of patient]

Theme 4: Facilitator in the Decision-Making Process

Positive Health Beliefs

The participants showed a desire to assist patients in their recovery and were the most important facilitators in decision-making. Such participants indicated that they would be more willing to consent to HIPEC, particularly for patients with GC at high risk of PM. They perceived the procedure as offering a promising therapeutic approach, citing its potential to prevent and treat metastatic dissemination, improve survival outcomes, and maintain a few side effects profile. Surrogate decision-makers considered that health was more important than wealth.

I know that the patient is not in good condition now. As long as the curative effect is good, the treatment should be treated, do not think about the problem of money, health is the most important. [Number 14, aged 24 years, individual, daughter of patient]

Health is the most important thing. I understand that HIPEC as early as possible after surgery can more thoroughly identify cancer cells, and then prolong life. As long as there is a effective treatment, we must actively coordinate. [Number 6, aged 32 years, office clerk, son of patient]

Cultural Dimensions of Perceived Responsibility: Confucian Perspective

A total of 60% (9/15) of surrogate decision-makers were children of patients who had the Chinese Confucian cultural values that maintained the integrity of the family and honored their parents. Having a parent with cancer triggered feelings of guilt, and participants felt it was their responsibility to help patients choose the best treatment. They were concerned that making the decision to forgo HIPEC would lead to adverse health outcomes for their parent.

My dad's situation of cancer was bad, but my parents have worked so hard for us since childhood, and as long as there is any hope, I want to try. [Number 10, aged 49 years, civil servant, daughter of patient]

My father knew that the cost of HIPEC was very expensive in a huff, and he wanted to give up the treatment. But he had worked hard all his life, and I, as a son, certainly would not give up the treatment easily. Now we comfort him every day and advise him to actively cooperate. [Number 9, aged 26 years, office clerk, son of patient]

Discussion

Principal Findings

In this qualitative study, 15 family caregivers of patients with GC as surrogate decision-makers were preliminarily studied in China. Applying thematic analysis to those that met our study inclusion criteria, we identified 4 themes, including the mode, facilitators, challenges, and decision-information sources of surrogate HIPEC decision-making.

Our findings identified 2 predominant modes of shared decision-making among HIPEC surrogate decision-makers. Those with lower educational attainment demonstrated reduced health literacy and stronger deference to hierarchical medical authority, making them more likely to rely on a health care provider for decision-making. This appears rooted in their limited capacity to comprehend medical concepts, compounded by inherent information asymmetry that predisposes clinical dominance. Lacking the ability and confidence to navigate this entrenched power structure, they often transferred final decision-making authority to clinicians, thereby mitigating the burden of decisional conflict [22]. Additionally, these decision-makers lack the foundational cognitive skills to assess information reliability and weigh risks against benefits—a finding consistent with Robertson et al [23]. This deficit impedes their ability to process highly specialized information. When confronted with unfamiliar and complex decisions, they frequently experience cognitive and informational overload. Such overload precipitates decisional conflict, characterized by internal contradiction and profound uncertainty, ultimately precluding meaningful participation in either independent or shared decision-making.

Surrogate decision-makers who used diverse information sources were more likely to engage in doctor-family shared decision-making, a finding consistent with Bakke et al [24]. From a personal capacity perspective, their cognitive and educational foundations translated into superior health literacy, enabling them to evaluate the credibility of web-based information and synthesize disparate sources into a coherent cognitive framework. This capacity fostered greater autonomy in assessing treatment efficacy and survival rates through independent verification [25]. Notably, during our interviews, these individuals voluntarily demonstrated their understanding of HIPEC specifics. Regarding behavioral strategies, these surrogates proactively managed information flow to ensure comprehensive understanding and approached decision-making as a process requiring their ultimate accountability. When

XSL•FO
RenderX

physicians and patients communicated effectively, they collectively arrived at treatment decisions [26].

The participation of patients and their families in decision-making conformed to the ethical principle of autonomy and should not play a passive decision-making role in this process [27]. Our findings underscore that facilitating effective shared decision-making requires the integrated provision of essential information alongside careful consideration of patient and family values throughout the decision-making continuum. Patient decision AIDS (PDAs) are popular tools in recent years that help patients make the best choices in health care options based on their personal values and goals, based on evidence-based medicine [28]. Online PDA provides treatment options according to the preferences of patients and their families, including related benefits, risks, and diagrams [29]. PDA can realize centralized information and effectively reduce decision-making conflicts in patients with cancer. From the health care team's perspective, we recommend establishing a dedicated HIPEC decision-support team comprising surgeons and specialized oncology nurses to provide structured decision guidance. This team would be responsible for implementing and systematically evaluating surrogate decision-makers' comprehension and use of decision aid materials [30]. Whole-process and dynamic decision-making support was provided to improve the health literacy of patients and their families and thus reduce decision-making difficulties and conflicts.

We found that participants lacked the basic knowledge of HIPEC and the ability to identify and understand information. Consistent with the result of Hart et al [31], surrogate decision-makers spent more time understanding decision information, with less information, they often misinterpret and undermine decisions. We found that participants with low education indicated that they were not interested in researching decision information. Xie et al [32] similarly found that caregivers' desire for information was lower than for decision-making participation. When the family caregivers, as nonprofessionals, were faced with highly specialized information such as HIPEC, they had difficulty understanding it. This inability even destroyed the positive degree of surrogate decision-makers who obtain further information, especially among the less educated [33]. These findings underscore that health care providers must prioritize the dissemination of HIPEC knowledge and enhance health education for patients with GC and their families, as this constitutes a fundamental prerequisite for facilitating informed decision-making. First, implementation should begin by advancing health care providers' understanding of current HIPEC evidence, enabling them to coherently translate and disseminate updated research insights to surrogates during decision-making consultations. Second, a supportive environment should be created. We recommend that health care institutions implement PDAs containing structured educational modules, which should provide an overview of GC natural history for HIPEC, comparative outcomes of treatment options, and so on. As demonstrated by Jayakumar et al [34], PDAs with professional guidance yield superior decision quality compared to merely distributing educational materials passively. With guidance from health care providers, these tools help laypersons establish fundamental disease and treatment knowledge, creating essential cognitive foundations for decision-making. Furthermore, providing educational videos at hospital admission significantly enhances engagement and health literacy among patients and their families, thereby improving both the decision-making experience and overall satisfaction with care [35].

The participants expressed diverse challenges facing HIPEC decisions. Financial burden emerged as the primary and most formidable challenge during the decision-making process. Consistent with the result of Graves et al [36], the lower-income patients and caregivers with HIPEC were associated with higher rates of decision regret. Some participants said that the single-use perfusion catheters required for HIPEC were a substantial out-of-pocket expense. During initial discussions of HIPEC, clinicians informed patients of the treatment costs but did not sufficiently underscore the substantial financial implications. As a result, families proceeded with the procedure only to discover—after 1 or 2 cycles—that the cumulative expenses far exceeded their anticipated financial burden. Consistent with reports from Western health systems [37], it was compounded in China as HIPEC was currently not reimbursable through national health insurance, representing a significant financial burden borne directly by families. Consequently, most treatment costs must be borne directly by families rather than through social insurance mechanisms. This financial burden compromises the ability of surrogate decision-makers to pursue clinically optimal decisions and develop treatment plans aligned with optimal clinical guidelines [38]. Uninsured patients with PM will be less likely to consider HIPEC or adjuvant chemotherapy, resulting in people with low economic and social status still at a disadvantage in terms of health [37]. Hamilton et al [39] proved that although the cost of HIPEC for patients was higher than that of systemic chemotherapy, the quality and the overall cost-effectiveness are better. Therefore, according to China's medical conditions, the medical security system should be improved, and the compensation ratio related to HIPEC should be refined to accurately solve the economic burden of patients with GC.

Moreover, the unpredictability of treatment outcomes significantly contributes to this complex decision-making situation. First, most significantly, while the combination of cytoreductive surgery and HIPEC demonstrates efficacy in both treating and preventing peritoneal metastases in advanced GC, this approach remains controversial and continues to evolve within the oncology community [40]. Based on their concern for the patient, participants were concerned about whether a previously unknown treatment can actually improve the cancer situation and inflict new damage to the patient. A joint multicenter study from the United Kingdom and Australia demonstrated significant variations in post-HIPEC survival outcomes and length of hospital stay across different geographic populations [41]. As highlighted by Marano et al [42], even within high-volume specialized institutions, the application of cytoreductive surgery with HIPEC varies substantially in terms of indication, technical execution, and outcome expectations, underscoring the need for clearer communication strategies with patients and their families. Second, our findings indicate that

XSL•FO
RenderX

participants who engaged with multiple information channels and possessed some prior understanding of HIPEC experienced particularly pronounced uncertainty, often stemming from encountering conflicting or inconsistent information across these sources. Different from the results of Crijns et al [26], higher health literacy and satisfaction with the decision of decision-makers were associated with sufficient information support for them. We analyzed that the possible cause was cancer information overload, which can be defined as feeling overwhelmed by the amount of cancer-related material in the information environment. Shi et al [43] show that cancer information overload is associated with perceived perception of illness and treatment uncertainty can lead to anxiety, depression, and emotional fatigue, and, in turn, impaired decision-making ability [30,44]. Effective clinician-family communication serves as a fundamental prerequisite for mitigating therapeutic uncertainty among decision-makers, aiming to support their psychological well-being and maintain alignment with treatment goals. In this context, physicians should engage in transparent dialogue with surrogate decision-makers, recognizing that artificial intelligence-based decision aids cannot substitute for essential human interaction [45].

Regarding the facilitator underlying the surrogate decision-making process for HIPEC, participants indicated their choices were informed by positive health perceptions as well as socioculturally informed perceived responsibilities. First, a participant believed that the “hope” brought by HIPEC can restore a healthy life status. Consistent with the results of Boegle et al [46], when decision-makers perceive that HIPEC is meeting their needs, evoked hope and trust, it benefits the adherence to their treatment. To be precise, they are decision-makers who may care more about longer survival and relapse in the future [45]. Second, consistent with Nissen et al [47], family functionality support is also conducive to improving the disease pressure of patients with cancer and their families and to reducing the physical and psychological burden on caregivers. Furthermore, within traditional Chinese cultural norms, adult children serving as surrogate decision-makers typically adhere to the Confucian virtue of filial piety. This ethical framework mandates comprehensive care for elderly parents, and such filial behavior has been associated with mitigated caregiver burden among adult children [48,49]. Seeking aggressive treatment modalities is often viewed as an expression of loyalty and familial responsibility. However, as demonstrated by Wu et al [50], children with strong filial values may feel compelled to pursue intensive treatment options even when facing substantial financial costs, which can paradoxically lead to decision-making conflicts and psychological distress. In summary, we recommend implementing family centric support protocols for individuals considering HIPEC, delivered through multidisciplinary shared decision-making teams that integrate specialized psychotherapists and oncology nurses to provide sustained psychosocial support. This structured approach enhances decision-making quality by systematically leveraging family caregivers’ capacity for optimism as a constructive psychological resource [30,44]. Given the significant role of filial piety in surrogate decision-making, health care providers should cultivate a balanced understanding of Confucian cultural values among adult children and proactively identify decision-making conflicts during clinical consultations.

Strengths and Limitations

Strengths and limitations of our study warrant discussion. First, our study focuses on the decision-making process of family caregivers of patients with GC, considering HIPEC, an understudied population (to the best of our knowledge). By adopting a qualitative approach, we delve into their lived experiences, addressing a frequently overlooked dimension of patient-centered care. Second, the systematic methodology used enhances the methodological rigor of the inquiry.

Our study has several limitations that should be considered when interpreting the findings. First, our application of a purposive sampling strategy, recruiting participants from a single health care institution in Southern China, means our findings may not be fully representative of all family caregivers of patients with GC considering HIPEC. Second, our cohort consisted predominantly of male participants, reflecting the sociocultural context in China where male family caregivers often assume the role of primary decision-maker. Yet, the framework of culture could also be a strength: this research is contextualized within the framework of Chinese culture, examining the underlying logic and sensitivities of relevant sociocultural norms for the understanding of decision-making processes. Finally, this study was limited by its exclusive focus on the perspectives of surrogate decision-makers. The findings do not incorporate the views of treating clinicians, the patients themselves, or direct observation of clinical interactions, which represent critical avenues for future research to triangulate and enrich our understanding of the decision-making dynamics.

Future Directions

Regarding future research, we propose several directions. Subsequent studies should aim to disentangle the specific factors involved in both information acquisition and the decision-making process for HIPEC. This includes investigating how individuals receive and process information from various channels and how they navigate psychosocial challenges during decision-making. We further propose large-scale, multicenter studies that use a mixed methods approach engaging multiple stakeholders—including surgeons, oncologists, and patients—to develop an integrated model of HIPEC decision-making. Finally, we recommend that future research intentionally enroll more female surrogate decision-makers to facilitate a gender-based analysis and elucidate its influence on decision-making outcomes.

Conclusions

This qualitative study explored the real experiences of GC patients’ family caregivers in making decisions about surrogate HIPEC treatment, including the participation mode, decision-making information sources, facilitators, and challenges. Our findings suggested that providing necessary and sufficient information support with regard to HIPEC was the key to achieving shared decision-making by surrogate decision-makers and improving the decision-making experience. Our findings highlighted areas that have not been previously addressed. Health care providers must intensify their efforts to

XSL•FO
RenderX

refine the shared decision-making framework for HIPEC and develop culturally sensitive PDA tailored to the sociocultural and health care context. Such instruments are essential to support GC families in achieving optimal surrogate decision-making aligned with patient values and clinical evidence.

Funding

Funding for LY was provided by Guangxi Medical and Health Appropriate Technology Development and Promotion Application Project (S2022069) and the First Affiliated Hospital of Guangxi Medical University (YYZS2020030).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Tables of a qualitative interview guide, participant characteristics, and themes and subthemes.
[DOCX File , 20 KB - cancer_v12i1e80471_app1.docx]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Koemans WJ, Lurvink RJ, Grootscholten C, Verhoeven RHA, de Hingh IH, van Sandick JW. Synchronous peritoneal metastases of gastric cancer origin: incidence, treatment and survival of a nationwide Dutch cohort. Gastric Cancer 2021;24(4):800-809. [doi: 10.1007/s10120-021-01160-1] [Medline: 33495964]
3. ElKordy MA, Soliman RM, ElTohamy MI, Mohamed DNE, Mustafa AM. Predictors of peritoneal metastasis of gastric origin. J Egypt Natl Canc Inst 2022;34(1):53. [doi: 10.1186/s43046-022-00155-y] [Medline: 36529760]
4. Hübner M, Kusamura S, Villeneuve L, Al-Niaimi A, Alyami M, Balonov K, et al. Guidelines for perioperative care in cytoreductive surgery (CRS) with or without hyperthermic intraperitoneal chemotherapy (HIPEC): Enhanced Recovery After Surgery (ERAS®) society recommendations — part II: postoperative management and special considerations. Eur J Surg Oncol 2020;46(12):2311-2323. [doi: 10.1016/j.ejso.2020.08.006] [Medline: 32826114]
5. Hübner M, Kusamura S, Villeneuve L, Al-Niaimi A, Alyami M, Balonov K, et al. Guidelines for perioperative care in cytoreductive surgery (CRS) with or without hyperthermic intraperitoneal chemotherapy (HIPEC): Enhanced Recovery After Surgery (ERAS®) society recommendations - part I: preoperative and intraoperative management. Eur J Surg Oncol 2020 Dec;46(12):2292 [FREE Full text] [doi: 10.1016/j.ejso.2020.07.041] [Medline: 32873454]
6. McBride KE, Steffens D, Solomon MJ, Koh C, Ansari N, Young CJ, et al. Cost-analysis of cytoreductive surgery and hyperthermic intraperitoneal chemotherapy in patients with peritoneal malignancy: an Australian perspective with global application. Eur J Surg Oncol 2021;47(4):828-833. [doi: 10.1016/j.ejso.2020.09.010] [Medline: 32972815]
7. Read MD, Drake J, Hashemipour G, Powers BD, Mehta R, Sinnamon A, et al. Initial experience using laparoscopic HIPEC for gastric cancer with peritoneal metastasis: safety and outcomes. Ann Surg Oncol 2024 Jun;31(6):3750-3757. [doi: 10.1245/s10434-024-15102-5] [Medline: 38430428]
8. Khomiak A, Ghaffar S, Franco SR, Ziogas IA, Yee EJ, Franklin O, et al. Survival outcomes of cytoreductive surgery with HIPEC in gastric adenocarcinoma: a national cancer database study. Ann Surg Oncol 2024;31(13):8549-8559. [doi: 10.1245/s10434-024-16142-7] [Medline: 39237827]
9. Deng H, Li B, Qin X. The short- and long-term survival of hyperthermic intraperitoneal chemotherapy (HIPEC) in the advanced gastric cancer with/without peritoneal carcinomatosis: a systematic review and meta-analysis of randomized controlled trials. Updates Surg 2022;74(6):1805-1816. [doi: 10.1007/s13304-022-01376-5] [Medline: 36116077]
10. Mazurek M, Szlendak M, Forma A, Baj J, Maciejewski R, Roviello G, et al. Hyperthermic intraperitoneal chemotherapy in the management of gastric cancer: a narrative review. Int J Environ Res Public Health 2022;19(2):681 [FREE Full text] [doi: 10.3390/ijerph19020681] [Medline: 35055500]
11. Dionne-Odom JN, Ejem D, Azuero A, Taylor RA, Rocque GB, Turkman Y, et al. Factors associated with family caregivers' confidence in future surrogate decision making for persons with cancer. J Palliat Med 2018;21(12):1705-1712 [FREE Full text] [doi: 10.1089/jpm.2018.0148] [Medline: 30129873]
12. Bibas L, Peretz-Larochelle M, Adhikari NK, Goldfarb MJ, Luk A, Englesakis M, et al. Association of surrogate decision-making interventions for critically ill adults with patient, family, and resource use outcomes: a systematic review and meta-analysis. JAMA Netw Open 2019;2(7):e197229 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.7229] [Medline: 31322688]
13. Strijker D, Meijerink W, Bremers A, de Reuver P, van Laarhoven C, van den Heuvel B. Prehabilitation to improve postoperative outcomes in patients with peritoneal carcinomatosis undergoing hyperthermic intraperitoneal chemotherapy

XSL•FO
RenderX

(HIPEC): a scoping review. Eur J Surg Oncol 2022;48(3):657-665 [FREE Full text] [doi: 10.1016/j.ejso.2021.10.006] [Medline: 34702591]
14. Strough J, Stone ER, Parker AM, Bruine de Bruin W. Perceived social norms guide health care decisions for oneself and others: a cross-sectional experiment in a US online panel. Med Decis Making 2022;42(3):326-340 [FREE Full text] [doi: 10.1177/0272989X211067223] [Medline: 34961398]
15. Nishioka M, Okuyama T, Uchida M, Aiki S, Ito Y, Osaga S, et al. What is the appropriate communication style for family members confronting difficult surrogate decision-making in palliative care?: A randomized video vignette study in medical staff with working experiences of clinical oncology. Jpn J Clin Oncol 2019;49(1):48-56. [doi: 10.1093/jjco/hyy178] [Medline: 30508206]
16. Larsen MK, Schultz H, Mortensen MB, Birkelund R. Decision making in the oesophageal cancer trajectory a source of tension and edginess to patients and relatives: a qualitative study. Int J Qual Stud Health Well-Being 2023;18(1):2170018 [FREE Full text] [doi: 10.1080/17482631.2023.2170018] [Medline: 36727528]
17. Tong A, Sainsbury P, Craig J. Consolidated Criteria For Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
18. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020;15(5):e0232076 [FREE Full text] [doi: 10.1371/journal.pone.0232076] [Medline: 32369511]
19. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol Health 2010;25(10):1229-1245. [doi: 10.1080/08870440903194015] [Medline: 20204937]
20. Renz SM, Carrington JM, Badger TA. Two strategies for qualitative content analysis: an intramethod approach to triangulation. Qual Health Res 2018;28(5):824-831. [doi: 10.1177/1049732317753586] [Medline: 29424274]
21. Rus M, Groselj U. Ethics of vaccination in childhood-a framework based on the four principles of biomedical ethics. Vaccines (Basel) 2021;9(2):113 [FREE Full text] [doi: 10.3390/vaccines9020113] [Medline: 33540732]
22. Kinchen E, Lange B, Newman D. Patient and provider decision-making experiences: a qualitative study. West J Nurs Res 2021;43(8):713-722. [doi: 10.1177/0193945920977476] [Medline: 33353517]
23. Robertson EG, Wakefield CE, Shaw J, Darlington A, McGill BC, Cohn RJ, et al. Decision-making in childhood cancer: parents' and adolescents' views and perceptions. Support Care Cancer 2019;27(11):4331-4340. [doi: 10.1007/s00520-019-04728-x] [Medline: 30880372]
24. Bakke BM, Feuz MA, McMahan RD, Barnes DE, Li B, Volow AM, et al. Surrogate decision makers need better preparation for their role: advice from experienced surrogates. J Palliat Med 2022;25(6):857-863 [FREE Full text] [doi: 10.1089/jpm.2021.0283] [Medline: 35076297]
25. Cincidda C, Pizzoli SFM, Ongaro G, Oliveri S, Pravettoni G. Caregiving and shared decision making in breast and prostate cancer patients: a systematic review. Curr Oncol 2023 Jan 06;30(1):803-823 [FREE Full text] [doi: 10.3390/curroncol30010061] [Medline: 36661710]
26. Crijns TJ, Segina P, Kortlever JTP, Thomas JE, Ring D, Reichel L, et al. Moderators and mediators of the relationship of shared decision-making and satisfaction. J Patient Exp 2023;10:23743735231171563 [FREE Full text] [doi: 10.1177/23743735231171563] [Medline: 37138951]
27. Corbett RW, Brown EA. Conventional dialysis in the elderly: how lenient should our guidelines be? Semin Dial 2018;31(6):607-611. [doi: 10.1111/sdi.12744] [Medline: 30239040]
28. Sepucha KR, Abhyankar P, Hoffman AS, Bekker HL, LeBlanc A, Levin CA, et al. Standards for Universal Reporting of Patient Decision Aid Evaluation studies: the development of SUNDAE checklist. BMJ Qual Saf 2018;27(5):380-388 [FREE Full text] [doi: 10.1136/bmjqs-2017-006986] [Medline: 29269567]
29. Mardinger C, Steve A, Webb C, Sherman K, Temple-Oberle C. Breast reconstruction decision aids decrease decisional conflict and improve decisional satisfaction: a randomized controlled trial. Plast Reconstr Surg 2023;151(2):278-288. [doi: 10.1097/prs.0000000000009830] [Medline: 36696307]
30. Qu RJ, Cheng QH, Wang HB, Li H. Development and practice of an interprofessional shared decision-making program for middle-aged lung cancer patients undergoing chemotherapy. J Nurs Sci 2022;37(23):25-28 [FREE Full text] [doi: 10.3870/j.issn.1001-4152.2022.23.025] [Medline: 38586126]
31. Hart A, Schaffhausen CR, McKinney WT, Gonzales K, Perugini J, Snyder JJ, et al. "You don't know what you don't know": a qualitative study of informational needs of patients, family members, and living donors to inform transplant system metrics. Clin Transplant 2024;38(1):e15240. [doi: 10.1111/ctr.15240] [Medline: 38289894]
32. Xie B, Champion JD, Kwak J, Fleischmann KR. Mobile health, information preferences, and surrogate decision-making preferences of family caregivers of people with dementia in rural hispanic communities: cross-sectional questionnaire study. J Med Internet Res 2018;20(12):e11682 [FREE Full text] [doi: 10.2196/11682] [Medline: 30530450]
33. Jach HK, DeYoung CG, Smillie LD. Why do people seek information? The role of personality traits and situation perception. J Exp Psychol Gen 2022;151(4):934-959. [doi: 10.1037/xge0001109] [Medline: 34914411]
34. Jayakumar P, Moore MG, Furlough KA, Uhler LM, Andrawis JP, Koenig KM, et al. Comparison of an artificial intelligence-enabled patient decision aid vs educational material on decision quality, shared decision-making, patient

XSL•FO
RenderX

experience, and functional outcomes in adults with knee osteoarthritis: a randomized clinical trial. JAMA Netw Open 2021 Feb 01;4(2):e2037107 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.37107] [Medline: 33599773]
35. Shin S, Ho J, Francis-Taylor R, Wells K, Halliday G, Jacek S, et al. Effect of an educational video and information pamphlet on knowledge and attitudes about electroconvulsive therapy: a randomized, blind, controlled study. J ECT 2022;38(3):211-217. [doi: 10.1097/YCT.0000000000000848] [Medline: 35462384]
36. Graves A, Sadjadi J, Kosich M, Ward E, Sood D, Fahy B, et al. Decision regret in patients with appendiceal cancer undergoing cytoreductive surgery and hyperthermic intraperitoneal chemotherapy. Ann Surg Oncol 2024;31(12):7978-7986. [doi: 10.1245/s10434-024-16013-1] [Medline: 39192009]
37. Kubi B, Nudotor R, Fackche N, Rowe J, Cloyd JM, Ahmed A, et al. Influence of insurance status on the postoperative outcomes of cytoreductive surgery and HIPEC. J Surg Oncol 2023;127(4):706-715. [doi: 10.1002/jso.27147] [Medline: 36468401]
38. Greenup RA, Rushing C, Fish L, Campbell BM, Tolnitch L, Hyslop T, et al. Financial costs and burden related to decisions for breast cancer surgery. J Oncol Pract 2019;15(8):e666-e676 [FREE Full text] [doi: 10.1200/JOP.18.00796] [Medline: 31356147]
39. Hamilton TD, MacNeill AJ, Lim H, Hunink MGM. Cost-effectiveness analysis of cytoreductive surgery and HIPEC compared with systemic chemotherapy in isolated peritoneal carcinomatosis from metastatic colorectal cancer. Ann Surg Oncol 2019;26(4):1110-1117. [doi: 10.1245/s10434-018-07111-y] [Medline: 30690682]
40. Patel M, Arora A, Mukherjee D, Mukherjee S. Effect of hyperthermic intraperitoneal chemotherapy on survival and recurrence rates in advanced gastric cancer: a systematic review and meta-analysis. Int J Surg 2023;109(8):2435-2450 [FREE Full text] [doi: 10.1097/JS9.0000000000000457] [Medline: 37158149]
41. Steffens D, Alves S, Koh C, Ansari N, Solomon MJ, Cecil TD, et al. International prospective quality of life in patients with colorectal peritoneal metastases treated by cytoreductive surgery and hyperthermic intraperitoneal chemotherapy in the United Kingdom and Australia. Dis Colon Rectum 2025;68(9):1074-1084. [doi: 10.1097/DCR.0000000000003823] [Medline: 40492567]
42. Marano L, Marrelli D, Sammartino P, Biacchi D, Graziosi L, Marino E, Italian Peritoneal Surface Malignancies Oncoteam (S.I.C.O.). Cytoreductive surgery and hyperthermic intraperitoneal chemotherapy for gastric cancer with synchronous peritoneal metastases: multicenter study of 'Italian peritoneal surface malignancies oncoteam-S.I.C.O.'. Ann Surg Oncol 2021;28(13):9060-9070 [FREE Full text] [doi: 10.1245/s10434-021-10157-0] [Medline: 34057569]
43. Shi W, Rothman AJ, Yzer MC, Nagler RH. Effects of exposure to conflicting information about mammography on cancer information overload, perceived scientists' credibility, and perceived journalists' credibility. Health Commun 2023;38(11):2481-2490. [doi: 10.1080/10410236.2022.2077163] [Medline: 35607276]
44. Torke AM, Varner-Perez SE, Burke ES, Taylor TA, Slaven JE, Kozinski KL, et al. Effects of spiritual care on well-being of intensive care family surrogates: a clinical trial. J Pain Symptom Manage 2023;65(4):296-307 [FREE Full text] [doi: 10.1016/j.jpainsymman.2022.12.007] [Medline: 36526251]
45. Dhawale T, Steuten LM, Deeg HJ. Uncertainty of physicians and patients in medical decision making. Biol Blood Marrow Transplant 2017;23(6):865-869 [FREE Full text] [doi: 10.1016/j.bbmt.2017.03.013] [Medline: 28302584]
46. Boegle K, Bassi M, Comanducci A, Kuehlmeyer K, Oehl P, Raiser T, et al. Informal caregivers of patients with disorders of consciousness: a qualitative study of communication experiences and information needs with physicians. Neuroethics 2022;15(3):24 [FREE Full text] [doi: 10.1007/s12152-022-09503-0] [Medline: 35912377]
47. Nissen KG, Trevino K, Lange T, Prigerson HG. Family relationships and psychosocial dysfunction among family caregivers of patients with advanced cancer. J Pain Symptom Manage 2016;52(6):841-849.e1 [FREE Full text] [doi: 10.1016/j.jpainsymman.2016.07.006] [Medline: 27521285]
48. Wamsiedel M, Li B. Why the oldest old in China bypass primary care: culture, family, and health system limitations. Int J Qual Stud Health Well-being 2025;20(1):2536103 [FREE Full text] [doi: 10.1080/17482631.2025.2536103] [Medline: 40677222]
49. Pan Y, Chen R, Yang D. The relationship between filial piety and caregiver burden among adult children: a systematic review and meta-analysis. Geriatr Nurs 2022;43:113-123 [FREE Full text] [doi: 10.1016/j.gerinurse.2021.10.024] [Medline: 34864295]
50. Wu X, Tang Y, Chen J. The association of filial piety and financial toxicity with surrogate decisional conflict among adult children of ICU patients with cancer. J Pain Symptom Manage 2025;70(5):437-446. [doi: 10.1016/j.jpainsymman.2025.07.023] [Medline: 40721047]

Abbreviations

GC: gastric cancer
HIPEC: hyperthermic intraperitoneal chemotherapy
PDA: patient decision AIDS
PM: peritoneal metastasis

XSL•FO
RenderX

Edited by A Schwartz, M Balcarras, N Cahill; submitted 12.Jul.2025; peer-reviewed by L Marano, H Veenendaal; comments to author 06.Oct.2025; revised version received 21.Nov.2025; accepted 19.Dec.2025; published 19.Jan.2026.

Please cite as:
Tan ZKK, Li DN, Jia K, Tang WZ, Chen X, Yang L
Surrogate Decision-Making by Family Caregivers for Hyperthermic Intraperitoneal Chemotherapy in Gastric Cancer: Qualitative Study in a High-Volume Chinese Center
JMIR Cancer 2026;12:e80471
URL: https://cancer.jmir.org/2026/1/e80471
doi:10.2196/80471
PMID:

©Zheng-Ke-Ke Tan, Dan-Ni Li, Kui Jia, Wen-Zhen Tang, Xin Chen, Li Yang. Originally published in JMIR Cancer (https://cancer.jmir.org), 19.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cancer Information Burden Among US Cancer Survivors Using Health Information National Trends Survey 2022 and 2024: Cross-Sectional Study

Jingjing Gao[1], PhD; Sharon V Munroe[2], PhD; Yue Zhang[3]; Jason H Windett[4], PhD

[1]Department of Management, Policy & Community Health, UTHealth Houston School of Public Health, 1200 Pressler Street, Houston, TX, United States

[2]Department of Health and Human Performance, College of Health Professions, University of Texas Rio Grande Valley, Houston, TX, United States

[3]Department of Biostatistics and Data Science, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, United States

[4]School of Data Science, University of North Carolina at Charlotte, Charlotte, NC, United States

Corresponding Author:
Jingjing Gao, PhD
Department of Management, Policy & Community Health, UTHealth Houston School of Public Health, 1200 Pressler Street, Houston, TX, United States

Abstract

Background: Digital health platforms have become increasingly important sources of health information for cancer survivors. Cancer survivors frequently rely on digital resources to obtain information related to diagnosis, treatment, survivorship, and follow-up care. In this study, cancer information burden is defined as the perceived challenges associated with seeking, understanding, and evaluating cancer-related information, including the effort required to obtain information, frustration during information seeking, difficulty understanding information, and concern regarding information quality. As the online health information environment becomes increasingly complex, some survivors may experience difficulty locating, evaluating, and understanding cancer-related information. However, limited nationally representative research has examined how cancer information burden differs across demographic and socioeconomic groups.

Objective: This study examined cancer information burden and associated demographic and socioeconomic factors among US cancer survivors using nationally representative data from the Health Information National Trends Survey (HINTS) 2022 and 2024.

Methods: Data were drawn from pooled cycles of the HINTS conducted in 2022 and 2024. Survey-weighted regression analyses were used to evaluate factors associated with cancer information burden among cancer survivors. Outcomes included the perceived effort required to obtain cancer information, frustration during information seeking, difficulty understanding cancer information, and concerns regarding information quality. Composite measures of cancer information burden were also examined. Models were adjusted for demographic and socioeconomic characteristics, including age, sex, race and ethnicity, educational attainment, household income, geographic region, and survey cycle.

Results: Hispanic cancer survivors reported a significantly greater cancer information burden, including higher levels of perceived effort and frustration associated with obtaining and understanding cancer-related information, compared with non-Hispanic White survivors. Non-Hispanic Asian survivors also demonstrated elevated levels of information burden, although estimates were less precise because of smaller subgroup sample sizes. Postbaccalaureate education was associated with lower frustration during cancer information seeking. Female survivors reported lower levels of difficulty understanding cancer-related information and lower overall information burden than male survivors. Differences between the 2022 and 2024 survey cycles were not statistically significant after adjustment.

Conclusions: Differences in cancer information burden persist among US cancer survivors. Hispanic and non-Hispanic Asian survivors, and individuals with lower educational attainment, may experience greater challenges obtaining and navigating cancer-related information. These findings support the development of accessible, culturally responsive, and health literacy–sensitive cancer communication strategies.

(JMIR Cancer 2026;12:e101310) doi:10.2196/101310

KEYWORDS

cancer survivors; cancer information burden; digital health; health literacy; generative AI

XSL•FO
RenderX

Introduction

Cancer survivors are increasingly turning to online health resources to manage diagnosis, treatment, survivorship, symptoms, and long-term care. Digital health tools like patient portals, educational websites, telehealth services, and search engines have become essential parts of today's cancer information landscape [1-3]. As health care systems expand their digital communication infrastructure, survivors are likely to encounter more complex online information environments. Emerging technologies, including conversational agents and algorithmically supported search tools, represent 1 component of the changing digital information environment [4,5]. However, the present study did not assess awareness, use, or exposure to these technologies.

Cancer survivors may be especially vulnerable to these changes because cancer-related information is often medically complex, emotionally distressing, and hard to interpret. Previous studies have shown that survivors commonly face information overload, frustration, and uncertainty when looking for cancer-related information online [6-11]. In this study, cancer information burden is defined as the perceived difficulty associated with seeking, understanding, and evaluating cancer-related information, including effort, frustration, difficulty understanding, and concern regarding information quality. Differences in digital access, education levels, health literacy, and trust in health information sources can further affect individuals' capacity to take advantage of changing digital health environments [12-14].

Prior studies have documented demographic and socioeconomic differences in cancer information seeking, information-source preferences, patient-centered communication, and digital information use among cancer survivors. Earlier Health Information National Trends Survey (HINTS)–based and population-based studies found that age, race and ethnicity, educational attainment, income, cancer history, and perceived health were associated with information-seeking behaviors and experiences [15-18]. These studies establish that cancer information experiences are not uniform across survivor populations. However, less is known about whether these patterns remain evident in the most recent HINTS cycles and how multiple dimensions of perceived information-seeking burden perform when examined separately and as a validated composite measure. Older, Black, and Latino seniors, and seniors from other minority groups have less access to digital tools, less online experience, and are less confident using web-based health resources, making them less likely to use eHealth or mobile health tools [19-21]. Nonetheless, limited nationally representative research has examined cancer information burden among survivors during the 2022 and 2024 period or identified the population groups most likely to report difficulty obtaining and understanding cancer information. Identifying population groups that report greater difficulty accessing and understanding cancer information may inform the development of more accessible digital communication strategies. It may also provide a baseline for future studies that directly assess how emerging digital technologies affect information access and navigation.

Using pooled HINTS 2022 and 2024 data, this study provides an updated, nationally representative assessment of cancer information-seeking burden among US cancer survivors. The analysis examined 4 related but distinct experiences: effort in obtaining information, frustration during information seeking, difficulty understanding information, and concern regarding information quality. We also evaluated the internal consistency and dimensional structure of a composite burden measure and examined its demographic and socioeconomic correlates. The study was intended to update and extend prior evidence rather than attribute observed patterns to emerging technologies. The primary research question was as follows: which demographic and socioeconomic characteristics are associated with perceived effort, frustration, difficulty understanding, concern regarding information quality, and overall cancer information-seeking burden among US cancer survivors? As a secondary analysis, we asked the following: does the association between age and cancer information-seeking burden differ between HINTS 2022 and HINTS 2024? Age was selected because older and younger cancer survivors may differ in digital access, familiarity with online health information, and experience navigating digital health resources.

This study was informed by the Comprehensive Model of Information Seeking, which proposes that health information-seeking experiences are shaped by antecedent characteristics, including demographic factors, personal experience, salience, and beliefs, as well as characteristics of information channels and sources [22-25]. In the context of cancer survivorship, age, sex, race and ethnicity, educational attainment, household income, geographic context, and prior cancer-related experience may influence how individuals seek, understand, and evaluate information. The present analysis focused on demographic and socioeconomic antecedents available consistently in HINTS 2022 and 2024. Because the dataset did not include all constructs required for a complete test of the framework, including detailed cancer characteristics, digital access, eHealth literacy, prior health care experiences, information-source characteristics, and trust, the framework was used to guide interpretation rather than to test the full model.

Methods

Data Source and Study Population

This study used pooled cross-sectional data from the 2022 and 2024 cycles of the HINTS, a nationally representative survey conducted by the National Cancer Institute that explores health communication behaviors, health information seeking, and digital health engagement among US adults. These were the 2 most recent publicly available nationally representative HINTS datasets at the time of analysis. A standard HINTS survey cycle was not conducted in 2023; therefore, no nationally representative HINTS 2023 dataset was available for inclusion. Cancer survivors were identified using the HINTS question that asked whether a doctor or other health professional had ever told the respondent that they had cancer. Respondents who answered "yes" were classified as cancer survivors. This definition captured a self-reported cancer diagnosis at any time and was not restricted to diagnoses occurring within the previous

XSL•FO
RenderX

year. Respondents who answered "no" were excluded from the cancer-survivor sample. No additional cancer-specific eligibility criteria were applied based on cancer type, stage, treatment status, recurrence, or time since diagnosis. Pooling the 2 survey cycles yielded 1972 cancer survivors. For outcome-specific descriptive analyses, respondents were required to have a valid response to the corresponding cancer information burden item. The composite score was calculated as the mean of available responses across the 4 burden items. Fully adjusted regression models were further restricted to respondents with complete data for the relevant outcome and all included covariates. Survey cycle was not treated as a measure or proxy of generative AI (GenAI) exposure because HINTS did not assess respondents' awareness, use, or exposure to GenAI tools.

Measures

Cancer Information Burden Outcomes

Four HINTS items were used to assess cancer information-seeking burden, defined as the perceived effort, frustration, difficulty understanding, and concern regarding information quality experienced during the search for cancer-related information. Respondents were asked, "based on the results of your most recent search for information about cancer, how much do you agree or disagree with each of the following statements?" The statements were as follows: "it took a lot of effort to get the information you needed," "you felt frustrated during your search for the information," "you were concerned about the quality of the information," and "the information you found was hard to understand." These items represent perceived effort, frustration, concern regarding information quality, and difficulty understanding cancer information, respectively. The original HINTS response options were strongly disagree, somewhat disagree, somewhat agree, and strongly agree.

Each item was coded on a 4-point ordinal scale ranging from 0 (strongly disagree) to 3 (strongly agree), with higher values indicating greater perceived cancer information burden. Responses that were coded as inapplicable, filter missing, not ascertained, or answered in error in the original HINTS data were treated as missing. The 4 items were analyzed separately as ordinal outcomes.

An overall cancer information burden score was calculated as the arithmetic mean of the available responses across the 4 coded burden items. The resulting score ranged from 0 to 3, with higher scores indicating greater overall perceived cancer information burden.

The 4-item composite demonstrated high internal consistency, with a Cronbach α of 0.881. Exploratory factor analysis supported a single-factor structure. The first factor had an eigenvalue of 2.948 and explained 73.7% of the total variance. Factor loadings were 0.876 for perceived effort, 0.898 for frustration, 0.820 for concern regarding information quality, and 0.838 for difficulty understanding. These findings support combining the 4 items into a composite score while retaining the individual items for dimension-specific analyses.

Covariates

Covariates were selected using a theory-informed approach based on the Comprehensive Model of Information Seeking and prior cancer communication literature. Demographic and socioeconomic variables were treated as antecedent characteristics that may shape cancer information-seeking experiences. Demographic and socioeconomic covariates included age, sex, race and ethnicity, educational attainment, household income, geographic census division, and survey cycle. Race and ethnicity categories included non-Hispanic White, non-Hispanic Black, Hispanic, non-Hispanic Asian, and non-Hispanic Other populations. Educational attainment was categorized as less than high school, high school graduate, some college, bachelor's degree, and postbaccalaureate degree. Household income was categorized into 6 groups ranging from less than US $20,000 annually to US $100,000 or greater. Geographic variation was assessed using US census divisions. Survey cycle was included as a temporal covariate to evaluate whether cancer information burden differed between HINTS 2022 and HINTS 2024; it was not interpreted as a proxy for GenAI exposure.

Missing Data and Analytic Samples

Of the 1972 cancer survivors in the pooled HINTS 2022 and 2024 samples, the number of respondents with valid data differed across the cancer information burden outcomes. Valid outcome data were available for 1373 respondents for perceived effort in obtaining cancer information, 1366 for difficulty understanding cancer information, 1367 for frustration during information seeking, and 1368 for concern regarding information quality. The composite cancer information burden score was available for 1389 respondents because it was calculated as the mean of available responses to the 4 burden items.

Adjusted regression models used complete-case estimation. Respondents were included in a given model only when they had valid data for the corresponding outcome and all covariates included in that model. Therefore, respondents with valid outcome data could contribute to descriptive analyses but were excluded from adjusted analyses when age, sex, race and ethnicity, education, household income, census division, or survey-cycle information were missing. Household income had the largest amount of missing covariate data and accounted for much of the reduction in analytic sample size.

The final complete-case regression samples included 957 respondents for perceived effort, 952 for difficulty understanding, 953 for frustration, 956 for concern regarding information quality, and 969 for the composite score. No values were imputed in the primary analyses.

Statistical Analysis

Survey-weighted analyses were conducted to account for the complex HINTS sampling design, including stratification, clustering, and sampling weights. Covariates were selected using the Comprehensive Model of Information Seeking and prior cancer communication literature. Age, sex, race and ethnicity, educational attainment, household income, US census division, and survey cycle were included as antecedent demographic and socioeconomic characteristics potentially

XSL•FO
RenderX

associated with cancer information-seeking burden. Missing data were handled using complete-case analysis. Each regression model included only respondents with complete data for the outcome of interest and all covariates. Consequently, analytic sample sizes varied slightly across outcomes because of differences in item nonresponse. No imputation procedures were performed.

Survey-weighted proportional-odds ordinal logistic regression models were estimated for the 4 ordinal cancer information burden outcomes. Results are presented as odds ratios (ORs) with 95% CIs. An OR greater than 1 indicates higher odds of reporting a greater level of cancer information burden, whereas an OR below 1 indicates lower odds.

The proportional-odds assumption was assessed for each ordinal outcome using parallel-lines diagnostic tests. Evidence of nonproportionality was identified for perceived effort in obtaining cancer information and concern regarding information quality, while findings for difficulty understanding cancer information were mixed across tests. No evidence of nonproportionality was identified for frustration during information seeking. For outcomes showing evidence of assumption violations, generalized ordered logistic regression models with partial proportional odds were estimated as sensitivity analyses, allowing selected coefficients to vary across outcome thresholds. These sensitivity analyses yielded conclusions that were substantively consistent with the primary models.

Survey-weighted linear regression was used to examine the composite cancer information burden score. Cronbach α and exploratory factor analysis were used to assess the internal consistency and dimensional structure of the 4-item composite measure.

As a secondary analysis, an age × survey cycle interaction term was added to each model to assess whether age-related associations with cancer information burden differed between HINTS 2022 and HINTS 2024. Age was selected as the interaction of interest because it may be associated with access to, familiarity with, and navigation of digital health information environments. Additional interactions were not systematically examined because the available sample size limited statistical power for multiple subgroup analyses, and extensive interaction testing could increase the likelihood of chance findings. Statistical significance was assessed using 2-sided tests with an α level of .05. All analyses were conducted using Stata (version 19.5; StataCorp LLC).

Ethical Considerations

This study was a secondary analysis of publicly available, deidentified data from the National Cancer Institute's HINTS. Because the dataset is publicly available and contains no identifiable participant information, a separate institutional review board review was not required for this study. The original HINTS data collection protocols were reviewed and approved by the Westat Institutional Review Board, and all participants provided informed consent before participating.

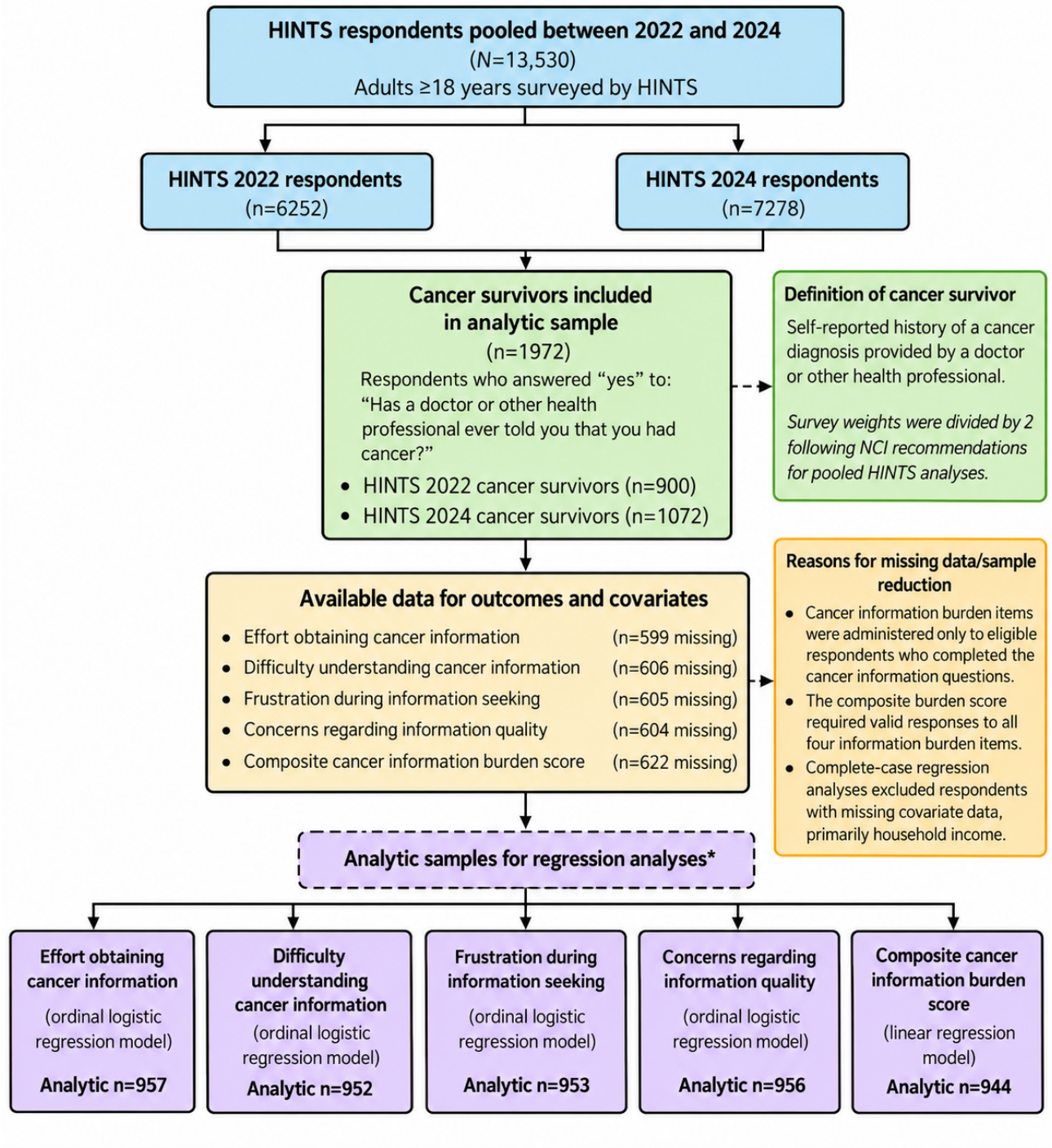
Results

Sample Characteristics

Figure 1 presents the study population selection process, while Table 1 summarizes the weighted descriptive characteristics of US cancer survivors included in the analytic sample. The pooled analytic sample included 1972 US cancer survivors from the 2022 and 2024 HINTS cycles. The unweighted sample included 900 (45.6%) respondents from HINTS 2022 and 1072 (54.4%) from HINTS 2024. After application of survey weights, the corresponding population estimates were 50.5% (n=997) for HINTS 2022 and 49.5% (n=976) for HINTS 2024. The weighted mean age of respondents was 65.8 (95% CI 64.5-67.1) years. Approximately 52.1% (95% CI 48.4%-55.6%) of respondents were female survivors. The sample was predominantly non-Hispanic White (77.8%, 95% CI 74.6%-80.7%) survivors, although Hispanic (8.4%, 95% CI 6.7%-10.5%), non-Hispanic Black (7.3%, 95% CI 5.7%-9.2%), non-Hispanic Asian (2.0%, 95% CI 1.3%-2.9%), and non-Hispanic Other (4.6%, 95% CI 3.1%-6.8%) populations were also represented. Educational attainment varied substantially across respondents. Approximately 42.3% (95% CI 39.2%-45.6%) reported some college education, while 16.9% (95% CI 14.8%-19.2%) reported a bachelor's degree and 16.8% (95% CI 14.7%-19.3%) reported postbaccalaureate education. Approximately 6.0% (95% CI 4.5%-7.8%) of respondents reported less than a high school education. Household income levels also varied across the sample. Approximately 22.0% (95% CI 18.7%-25.7%) of respondents reported annual household incomes between US $50,000 and US $74,999, while 21.6% (95% CI 18.6%-25.0%) reported incomes of at least US $100,000 annually. Approximately 16.3% (95% CI 13.8%-19.3%) reported annual household incomes below US $20,000. Of the 1972 cancer survivors included in the pooled dataset, between 1366 and 1373 respondents provided valid responses to the individual cancer information burden items. The composite cancer information burden score was available for 1389 respondents because it was calculated as the mean of available responses across the 4 burden items. After complete-case exclusion of respondents with missing covariate information, primarily household income, the final regression samples ranged from 952 to 957 participants. The composite burden analysis included 969 respondents.

XSL•FO
RenderX

Figure 1. Study population selection and analytic sample flowchart. Health Information National Trends Survey (HINTS) 2022 and HINTS 2024 were the most recent publicly available survey cycles. No standard HINTS survey was conducted in 2023. N and n represent unweighted sample sizes. The composite cancer information burden score was calculated only for respondents with valid responses to all 4 burden items (n=1350). *Analytic n represents the complete-case sample used for each primary regression model after excluding respondents with missing covariate data. NCI: National Cancer Institute.



XSL•FO
RenderX

Table . Weighted descriptive characteristics of US cancer survivors included in the analytic sample[a].

Characteristic	Unweighted, n	Weighted
Age (y), mean (95% CI)	1934	65.8 (64.5-67.1)
Composite cancer information burden score, mean (95% CI)	1389	1.24 (1.17-1.31)
Sex, % (95% CI)		
Female	1118	52.1 (48.4-55.6)
Male	828	47.9 (44.4-51.6)
Race and ethnicity, % (95% CI)		
Non-Hispanic White	1338	77.8 (74.6-80.7)
Non-Hispanic Black	207	7.3 (5.7-9.2)
Hispanic	184	8.4 (6.7-10.5)
Non-Hispanic Asian	44	2.0 (1.3-2.9)
Non-Hispanic Other	58	4.6 (3.1-6.8)
Educational attainment, % (95% CI)		
Less than high school	117	6.0 (4.5-7.8)
High school graduate	315	18.0 (15.6-20.6)
Some college	610	42.3 (39.2-45.6)
Bachelor’s degree	476	16.9 (14.8-19.2)
Postbaccalaureate degree	427	16.8 (14.7-19.3)
Household income (US $), % (95% CI)		
Less than 20,000	308	16.3 (13.8-19.3)
20,000-34,999	245	17.6 (13.9-22.0)
35,000-49,999	241	16.4 (13.8-19.3)
50,000-74,999	294	22.0 (18.7-25.7)
75,000-99,999	115	6.1 (4.5-8.1)
100,000 or greater	284	21.6 (18.6-25.0)
Survey cycle, % (95% CI)		
HINTS[b] 2022	900	50.5 (47.3-53.6)
HINTS 2024	1072	49.5 (46.4-52.7)
Effort in obtaining cancer information, % (95% CI)		
Strongly disagree	355	25.2 (21.7-28.9)
Somewhat disagree	470	35.1 (31.4-38.9)
Somewhat agree	411	30.4 (26.6-34.5)
Strongly agree	137	9.4 (7.2-12.2)
Cancer information is hard to understand, % (95% CI)		
Strongly disagree	414	30.2 (26.4-34.3)
Somewhat disagree	487	34.8 (31.1-38.7)
Somewhat agree	367	28.1 (24.3-32.1)
Strongly agree	98	6.9 (5.2-9.2)
Cancer information seeking is frustrating, % (95% CI)		
Strongly disagree	446	31.5 (27.8-35.4)
Somewhat disagree	443	32.5 (29.0-36.2)
Somewhat agree	355	27.8 (24.4-31.5)

XSL•FO
RenderX

Characteristic	Unweighted, n	Weighted
Strongly agree	123	8.2 (6.3-10.6)
Concerned about cancer information quality, % (95% CI)		
Strongly disagree	323	22.3 (19.1-25.8)
Somewhat disagree	365	26.9 (23.5-30.7)
Somewhat agree	451	35.9 (32.3-39.7)
Strongly agree	229	14.9 (12.2-18.0)

[a]Unweighted counts are presented alongside survey-weighted percentages or means and their 95% CIs. Survey-weighted estimates account for the complex HINTS sampling design. Percentages may not sum to 100% because of rounding.

[b]HINTS: Health Information National Trends Survey.

Regarding measures of cancer information burden, approximately 39.8% of respondents somewhat or strongly agreed that obtaining cancer information required substantial effort. Similarly, 35.0% somewhat or strongly agreed that cancer information was difficult to understand, while 36.0% somewhat or strongly agreed that seeking cancer information was frustrating. Concerns regarding cancer information quality were also common, with approximately 50.8% somewhat or strongly agreeing that they were concerned about the quality of cancer-related information. The 4 cancer information burden items demonstrated high internal consistency (Cronbach α=0.881). Exploratory factor analysis identified 1 dominant factor with an eigenvalue of 2.948, accounting for 73.7% of the total variance. All 4 items loaded strongly on this factor, with loadings ranging from 0.820 to 0.898. These findings supported the use of the mean of the 4 items as an overall cancer information burden score.

Factors Associated With Cancer Information Burden

Parallel-lines diagnostics indicated evidence of nonproportionality for perceived effort in obtaining cancer information and concern regarding information quality. Diagnostic findings for difficulty understanding cancer information were mixed, whereas no evidence of nonproportionality was identified for frustration during information seeking. Partial proportional-odds sensitivity models were therefore estimated for outcomes showing evidence of assumption violations. These sensitivity analyses produced conclusions that were substantively consistent with those of the primary proportional-odds models. Table 2 presents survey-weighted regression models examining factors associated with cancer information burden among US cancer survivors. Several significant demographic and socioeconomic differences emerged across perceived effort in obtaining cancer information, difficulty understanding cancer information, frustration during information seeking, concern regarding information quality, and the composite cancer information burden score.

Hispanic cancer survivors experienced greater cancer information burden than non-Hispanic White survivors for selected outcomes. Hispanic respondents had significantly higher odds of reporting greater effort in obtaining cancer information (OR 2.611, 95% CI 1.387-4.913, P=.003) and greater frustration during cancer information seeking (OR 2.087, 95% CI 1.154-3.773, P=.02). Hispanic survivors also had higher composite cancer information burden scores, although this association did not reach statistical significance after adjustment for covariates (β=0.265, 95% CI −0.002 to 0.532, P=.051). These findings indicate that Hispanic cancer survivors reported greater cancer information burden for selected outcomes. Because the data are cross-sectional, the observed associations do not establish why these differences occurred.

XSL•FO
RenderX

Table . Survey-weighted ordinal logistic and linear regression models of cancer information burden among US cancer survivors[a].

Variable	Effort in obtaining cancer information, OR[b] (95% CI)	Difficulty understanding cancer information, OR (95% CI)	Frustration during information seeking, OR (95% CI)	Concern regarding information quality, OR (95% CI)	Composite cancer information burden, β[c] (95% CI)
Age (y)	0.993 (0.977-1.010)	0.988 (0.973-1.004)	0.991 (0.977-1.006)	0.986 (0.971-1.001)	−0.005 (−0.011 to 0.002)
Sex (reference: male)					
Female	0.782 (0.554-1.106)	0.562 (0.383-0.824)[d]	0.683 (0.489-0.955)[e]	0.747 (0.539-1.034)	−0.177 (−0.313 to −0.041)[e]
Race and ethnicity (reference: non-Hispanic White)					
Non-Hispanic Black	1.381 (0.565-3.376)	0.809 (0.359-1.826)	0.811 (0.357-1.843)	1.153 (0.544-2.444)	0.047 (−0.257 to 0.351)
Hispanic	2.611 (1.387-4.914)[d]	1.681 (0.769-3.671)	2.087 (1.154-3.774)[e]	1.066 (0.573-1.981)	0.265 (−0.002 to 0.532)
Non-Hispanic Asian	5.495 (1.542-19.584)[d]	1.577 (0.385-6.465)	2.898 (0.687-12.230)	1.363 (0.367-5.062)	0.480 (−0.083 to 1.044)
Non-Hispanic other	2.974 (0.806-10.980)	0.392 (0.119-1.293)	1.324 (0.458-3.823)	2.798 (1.223-6.401)[e]	0.176 (−0.085 to 0.436)
Educational attainment (reference: less than high school)					
High school graduate	0.495 (0.132-1.852)	0.766 (0.223-2.630)	0.334 (0.088-1.268)	0.376 (0.090-1.568)	−0.338 (−0.824 to 0.149)
Some college	0.868 (0.250-3.017)	0.774 (0.244-2.461)	0.344 (0.095-1.247)	0.335 (0.082-1.372)	−0.306 (−0.767 to 0.155)
Bachelor's degree	0.591 (0.158-2.202)	0.480 (0.152-1.518)	0.270 (0.072-1.019)	0.438 (0.105-1.825)	−0.399 (−0.875 to 0.077)
Postbaccalaureate degree	0.594 (0.162-2.172)	0.506 (0.156-1.640)	0.256 (0.068-0.955)[e]	0.354 (0.084-1.493)	−0.440 (−0.919 to 0.038)
Household income (reference: less than US $20,000)					
20,000-34,999	0.774 (0.306-1.953)	0.791 (0.306-2.047)	0.518 (0.219-1.228)	0.865 (0.357-2.094)	−0.163 (−0.506 to 0.181)
35,000-49,999	0.675 (0.289-1.579)	1.713 (0.737-3.978)	0.933 (0.426-2.042)	1.323 (0.576-3.038)	0.039 (−0.279 to 0.357)
50,000-74,999	0.711 (0.318-1.589)	1.065 (0.480-2.366)	0.901 (0.436-1.862)	1.159 (0.518-2.595)	−0.023 (−0.321 to 0.276)
75,000-99,999	0.561 (0.179-1.760)	1.956 (0.610-6.273)	1.480 (0.517-4.231)	1.566 (0.544-4.510)	0.092 (−0.349 to 0.532)
100,000 or greater	0.507 (0.204-1.262)	0.855 (0.369-1.978)	0.819 (0.358-1.875)	1.025 (0.427-2.460)	−0.151 (−0.491 to 0.188)
Survey cycle (reference: HINTS[f] 2022)					
HINTS 2024	1.002 (0.625-1.604)	1.086 (0.671-1.756)	0.875 (0.550-1.392)	0.697 (0.451-1.077)	−0.045 (−0.222 to 0.131)
Observations, n	957	952	953	956	969
Overall model, *P* value	<.001	.006	.04	.40	.009
R^2	N/A[g]	N/A	N/A	N/A	.068

[a]ORs greater than 1 indicate higher odds of reporting a higher level of cancer information burden. The proportional-odds assumption was evaluated using parallel-lines tests, and partial proportional-odds sensitivity models were estimated for outcomes showing evidence of nonproportionality.

[b]OR: odds ratio.

[c]β: unstandardized linear regression coefficient.

[d]*P*<.01.

[e]*P*<.05.

[f]HINTS: Health Information National Trends Survey.

[g]N/A: not applicable.

Non-Hispanic Asian cancer survivors also demonstrated elevated cancer information burden for selected outcomes. In particular, Asian survivors had significantly higher odds of reporting greater effort in obtaining cancer information than non-Hispanic White survivors (OR 5.495, 95% CI 1.543-19.568, *P*=.009). Estimates for frustration and the composite burden score were

XSL•FO
RenderX

higher, but these associations did not reach statistical significance. The wide CIs likely reflect the smaller number of non-Hispanic Asian respondents in the analytic sample.

Female cancer survivors generally reported a lower cancer information burden than male survivors. Female survivors had significantly lower odds of reporting difficulty understanding cancer information (OR 0.562, 95% CI 0.383-0.824, P=.003) and greater frustration when seeking cancer information (OR 0.683, 95% CI 0.489-0.954, P=.03). They also had significantly lower overall cancer information burden scores (β=−0.177, 95% CI −0.313 to −0.041, P=.01). These associations may be related to differences in health information-seeking experience or familiarity with health care navigation, although these potential explanations were not directly measured in HINTS.

Educational attainment was associated with frustration during cancer information seeking. Respondents with postbaccalaureate education had significantly lower odds of reporting greater frustration than respondents with less than a high school education (OR 0.256, 95% CI 0.069-0.951, P=.042). Respondents with a bachelor's degree showed a similar pattern, although the association narrowly missed statistical significance (OR 0.270, 95% CI 0.072-1.016, P=.053). Other education-related estimates generally suggested lower cancer information burden among respondents with higher educational attainment, but the CIs included the null value.

Non-Hispanic respondents classified in the Other racial and ethnic category had significantly higher odds of reporting greater concern regarding cancer information quality than non-Hispanic White respondents (OR 2.798, 95% CI 1.222-6.408, P=.02). This estimate should be interpreted cautiously because of the heterogeneity and likely small sample size of this category.

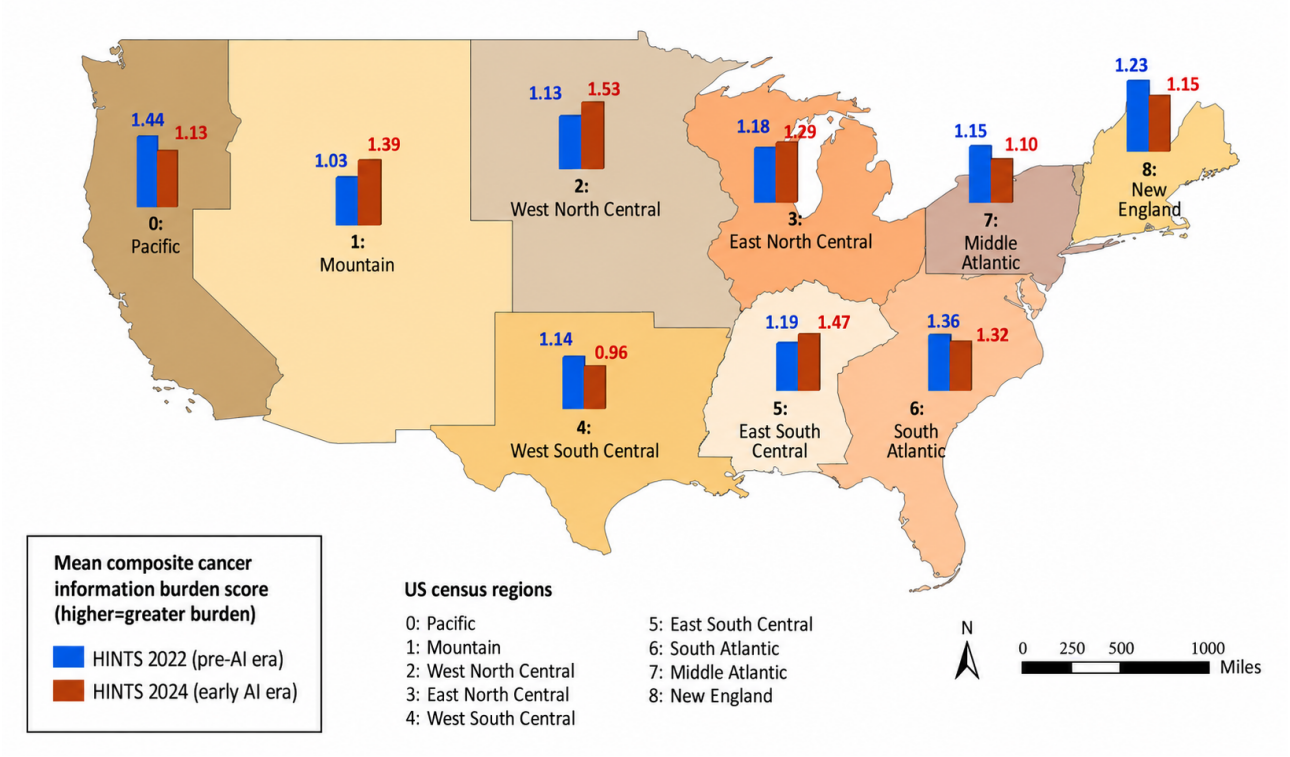
Age was not significantly associated with the cancer information burden outcomes after covariate adjustment. Household income and US census division also demonstrated limited associations across the models. Differences between the 2022 and 2024 HINTS cycles were not statistically significant for any outcome. These findings indicate that the measured associations did not differ significantly between the 2 survey cycles.

Figure 2 presents exploratory, survey-weighted mean cancer information burden scores across US census divisions. Certain divisions appeared to have higher or lower mean scores for selected burden measures, but the patterns were not uniform across outcomes or survey cycles. No formal pairwise or omnibus comparisons were conducted for this figure; therefore, these descriptive differences should not be interpreted as statistically significant geographic variation. Because the number and weighted distribution of cancer survivors differed across census divisions and survey cycles, some of the visual variation may reflect differences in regional sample composition. Multimedia Appendix 1 reports the unweighted sample sizes and weighted percentages of cancer survivors by census division and survey cycle.

The regional distribution of cancer survivors did not differ significantly between HINTS 2022 and HINTS 2024, based on a survey-weighted design-adjusted Pearson test ($F_{7.48,2594.52}$=0.943, P=.48). Thus, the descriptive differences shown in Figure 2 were not accompanied by evidence of a significant change in the geographic composition of the cancer-survivor sample across survey cycles.

XSL•FO
RenderX

Figure 2. Exploratory, survey-weighted mean cancer information burden scores across US census regions (Health Information National Trends Survey [HINTS] 2022 vs HINTS 2024). Values represent descriptive survey-weighted means for perceived effort in obtaining cancer information, difficulty understanding cancer information, frustration during information seeking, concern regarding information quality, and the composite burden score across the US census divisions. No formal pairwise or omnibus statistical comparisons were conducted. Regional patterns should therefore be interpreted as exploratory. Unweighted sample sizes and weighted distributions by census division and survey cycle are reported in Multimedia Appendix 1. Data source: HINTS 2022 and 2024 survey-weighted means estimated using linearized variance.



Secondary Analysis of the Interaction Between Age and Survey Cycle

Table 3 presents the secondary interaction analyses examining whether age-related associations with cancer information burden differed between HINTS 2022 and HINTS 2024. Across all 5 outcomes, the age × survey-cycle interaction terms were not statistically significant, indicating no evidence that the association between age and cancer information burden differed between HINTS 2022 and HINTS 2024. For effort in obtaining cancer information, the interaction between age and HINTS 2024 was not statistically significant (β=−0.004, 95% CI −0.034 to 0.025, *P*=.77). Similarly, no significant age-by-cycle interaction was observed for difficulty understanding cancer information (β=−0.019, 95% CI −0.058 to 0.020, *P*=.33), frustration during cancer information seeking (β=−0.012, 95% CI −0.043 to 0.020, *P*=.47), concerns regarding cancer information quality (β=−0.012, 95% CI −0.047 to 0.024, *P*=.52), or the composite cancer information burden score (β=−0.005, 95% CI −0.018 to 0.008, *P*=.45).

XSL•FO
RenderX

Table . Interaction between age and survey year in models of cancer information burden among US cancer survivors[a].

Outcome	Age, β (95% CI)	HINTS[b] 2024, β (95% CI)	Age × HINTS 2024, β (95% CI)	*P* value	Observations
Effort in obtaining cancer information	−0.003 (−0.025 to 0.018)	0.296 (−1.811 to 2.402)	−0.004 (−0.034 to 0.025)	.77	957
Difficulty understanding cancer information	0.003 (−0.030 to 0.035)	1.364 (−1.376 to 4.104)	−0.019 (−0.058 to 0.020)	.33	952
Frustration during information seeking	0.000 (−0.026 to 0.026)	0.644 (−1.596 to 2.885)	−0.012 (−0.043 to 0.020)	.47	953
Concerns regarding information quality	−0.006 (−0.036 to 0.025)	0.422 (−2.165 to 3.009)	−0.012 (−0.047 to 0.024)	.52	956
Composite cancer information burden score	−0.001 (−0.011 to 0.010)	0.288 (−0.613 to 1.188)	−0.005 (−0.018 to 0.008)	.45	969

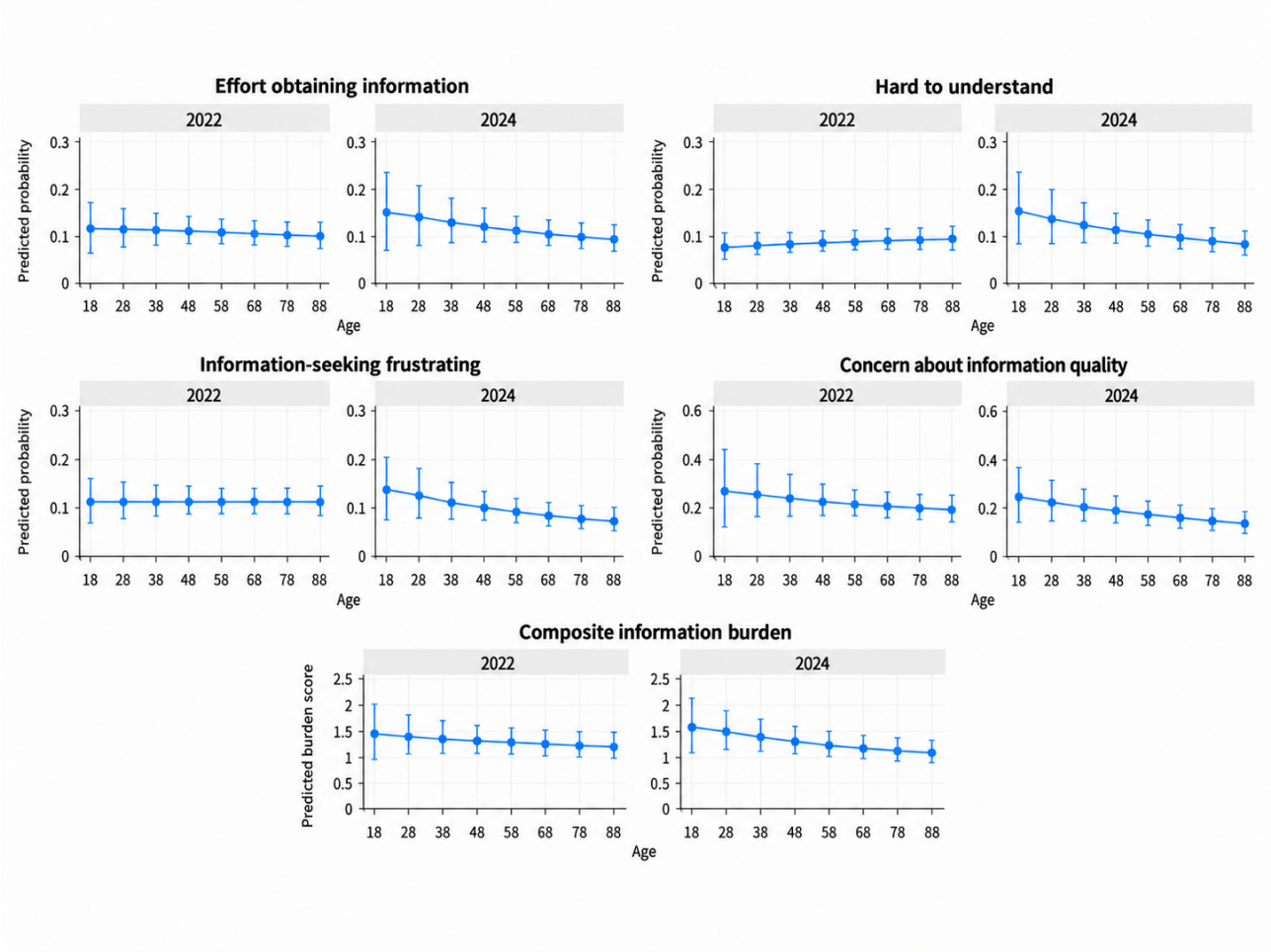
[a]Survey-weighted proportional-odds ordinal logistic regression models were estimated for the 4 ordinal cancer information burden outcomes. For these models, β estimates are unexponentiated coefficients expressed on the log-odds scale. Survey-weighted linear regression was used for the composite cancer information burden score, for which β estimates are unstandardized linear regression coefficients. Models were adjusted for sex, race and ethnicity, educational attainment, household income, and US census division. The age × survey-cycle interaction term indicates whether the association between age and each outcome differed between HINTS 2022 and HINTS 2024. HINTS 2022 was the reference survey cycle.

[b]HINTS: Health Information National Trends Survey.

Figure 3 presents predictive margins that illustrate age-specific trends in cancer information burden across the 2022 and 2024 survey cycles. Overall, predicted probabilities and burden scores were broadly comparable between survey years, with overlapping CIs across most age groups. Across several outcomes, younger cancer survivors demonstrated modestly higher predicted levels of cancer information burden compared with older survivors, particularly in 2024. This pattern was most apparent for difficulty understanding cancer information, concerns regarding cancer information quality, and the composite cancer information burden score. However, CIs were wider at younger ages, indicating greater uncertainty in these estimates. Collectively, these findings indicate that age-related associations with cancer information burden did not differ significantly between the 2022 and 2024 survey cycles. Because HINTS did not measure GenAI use or exposure, these findings cannot be attributed to changes in AI-enabled health information environments.

XSL•FO
RenderX

Figure 3. Predicted cancer information burden by age and survey year among US cancer survivors.



Discussion

Principal Findings and Interpretation

This study provides an updated, nationally representative assessment of cancer information-seeking burden among US cancer survivors using HINTS 2022 and 2024 data. The findings were broadly consistent with earlier HINTS-based and population-based studies showing that cancer information-seeking behaviors and experiences differ across demographic and socioeconomic groups [15-18]. The present analysis extends this literature by examining 4 burden dimensions separately, evaluating a 4-item composite measure, and assessing whether these associations were evident in recent HINTS cycles. Accordingly, the findings should be interpreted primarily as a contemporary and confirmatory update rather than as evidence of entirely new demographic patterns. In contrast, adjusted differences between the 2022 and 2024 survey cycles and age × survey-cycle interactions were not statistically significant. These findings should be interpreted as evidence regarding cancer information burden during 2022 and 2024 and not as evidence of GenAI effects because HINTS did not assess AI use, awareness, or exposure. Viewed through the Comprehensive Model of Information Seeking, the observed associations may reflect differences in antecedent characteristics that shape information-seeking experiences. However, the current study did not measure several constructs central to the framework, including the salience of current information needs, beliefs, prior cancer experience, information-channel characteristics, trust, digital access, and eHealth literacy. The findings therefore support a theory-informed interpretation but should not be considered a formal test of the model. The secondary age × survey-cycle analyses did not identify statistically significant interactions for any outcome. These results indicate that the association between age and cancer information burden did not differ significantly across the 2 survey cycles, although the analyses may have had limited power to detect modest interaction effects.

One of the most notable findings was the elevated cancer information burden reported among Hispanic cancer survivors. Hispanic respondents consistently demonstrated higher perceived effort associated with obtaining cancer information and greater frustration during cancer information seeking compared with non-Hispanic White survivors. Non-Hispanic Asian respondents also demonstrated significantly higher perceived effort in obtaining cancer information. These findings are consistent with prior research documenting differences in health literacy, access to culturally tailored information, and digital navigation experiences across population groups [20,26-28]. The associations remained after adjustment for education, income, and geographic region. However, the present study did not directly measure health literacy, structural barriers, digital navigation skills, or exposure to AI-generated health content. These factors therefore represent possible areas for future investigation rather than established explanations for the

XSL•FO
RenderX

observed associations. Overall, these findings are consistent with prior HINTS-based and population-based research. Davis et al [15] identified demographic correlates of cancer information-seeking experiences among US cancer survivors, while Elkefi and Matthews [16] documented differences in information-seeking behaviors and source preferences across a diverse survivor sample. Mahmood et al [17] further showed that patient-centered communication was associated with cancer information-seeking experiences, and Benzo et al [18] reported associations between digital information-seeking behaviors and sociodemographic characteristics, cancer history, and perceived health. The present study extends this literature by showing that selected demographic differences remain evident in recent HINTS data and vary across specific dimensions of cancer information-seeking burden.

The descriptive patterns shown in Figure 2 suggest possible geographic variation in cancer information burden, but these observations are exploratory. The figure was not accompanied by formal pairwise or omnibus testing, and the number and weighted distribution of cancer survivors varied across census divisions and survey cycles. Consequently, some visual differences may reflect regional sample composition rather than independent geographic effects. Census division was included as a covariate in the adjusted regression models, but most division-specific estimates were not statistically significant. Future studies with larger geographic samples should formally test regional differences and examine whether survey-cycle changes vary across areas.

Educational attainment also demonstrated a protective relationship with several dimensions of cancer information burden. Cancer survivors with bachelor's or postbaccalaureate degrees reported lower levels of frustration during cancer information seeking and generally lower composite burden scores. These associations may be related to differences in digital literacy, familiarity with health systems, or the ability to evaluate complex health information. However, these potential mechanisms were not directly measured in the present study. Prior research has consistently demonstrated that health literacy and digital literacy strongly influence the ability to effectively navigate online health information environments [29,30]. Earlier HINTS-based studies have also linked educational attainment and socioeconomic characteristics with cancer information-seeking experiences and source use [15,17]. The current findings therefore reinforce prior evidence that educational resources may shape how survivors experience the effort and frustration associated with obtaining cancer information. The present findings indicate that educational attainment was associated with selected dimensions of perceived cancer information burden. The study cannot determine whether this association was influenced by the increasing availability of AI-enabled information tools.

Interestingly, female cancer survivors reported significantly less difficulty understanding cancer information and less frustration during information seeking than male survivors. Prior HINTS-based and population-based studies have also identified sex differences in health information-seeking behaviors, preventive service use, and engagement with health care systems [31,32]. Female survivors may have greater prior experience seeking health information, communicating with health professionals, and managing health care for themselves or family members. These experiences may provide context for the observed associations. However, the present study did not directly measure prior information-seeking frequency, caregiving roles, digital literacy, or health care navigation experience, and these factors should not be interpreted as established explanatory mechanisms. Future studies should directly evaluate whether these characteristics account for sex differences in cancer information-seeking burden.

Age was examined as a potential modifier because digital access, online information-seeking experience, and familiarity with digital health tools may differ across age groups. However, the absence of statistically significant interactions suggests that age-related patterns were not detectably different between the 2022 and 2024 cycles. Because HINTS did not directly measure digital literacy, technology use, or AI exposure, the mechanisms underlying age-related information experiences could not be evaluated. Future studies with larger samples and direct measures of digital engagement should further assess whether age modifies the relationship between changing information environments and cancer information burden.

Although AI was not evaluated in this study, the findings provide a baseline for future research on cancer communication in changing digital environments. GenAI tools may offer opportunities to simplify medical terminology, tailor educational materials, and support patient engagement, but they may also introduce concerns related to information accuracy, bias, and unequal access. Because HINTS did not measure AI use, exposure, or reliance on AI-generated information, the present results cannot determine whether AI-enabled systems influenced cancer information burden. Future studies should directly assess the use of these tools and evaluate their associations with comprehension, trust, and information burden.

The incremental contribution of this study lies in its use of recent pooled HINTS cycles, its simultaneous evaluation of 4 specific burden dimensions, and its psychometric assessment of a composite cancer information burden measure. The high internal consistency and single-factor structure supported the use of the composite, while the item-specific models showed that associations were not uniform across all dimensions. This outcome-specific pattern provides more detail than treating cancer information seeking as a single behavior or experience.

Limitations

This study has several limitations. First, the cross-sectional design precludes causal inference. Second, the study relied on self-reported perceptions of cancer information burden rather than objective measures of information-seeking behavior or comprehension. Third, the analyses did not account for cancer-specific characteristics, including cancer type, time since diagnosis, treatment status, disease stage or severity, recurrence, or survivorship phase. These factors may influence information-seeking needs and perceived burden and could contribute to residual confounding. Future studies should examine whether the observed associations vary across cancer types and phases of care. The analyses also did not adjust for factors more directly related to digital information access and

XSL•FO
RenderX

navigation, such as internet use, frequency of online health information seeking, digital literacy, eHealth literacy, or access to digital technologies. These factors may influence perceived cancer information burden and could partly explain the observed associations. Future studies should incorporate these measures to better understand the pathways through which demographic and socioeconomic characteristics are associated with cancer information burden. Fourth, HINTS 2022 and 2024 did not measure GenAI awareness, use, chatbot engagement, AI-assisted search behavior, or exposure to AI-generated health information. Survey cycle therefore cannot be interpreted as a proxy for AI exposure, and neither differences nor the absence of differences between 2022 and 2024 can be attributed to GenAI. References to changes in the digital information environment are intended only to provide broader temporal context. Fifth, some subgroup analyses, particularly among non-Hispanic Asian respondents, were limited by relatively small sample sizes, resulting in wider CIs. Finally, although the study examined several dimensions of cancer information burden, additional factors such as trust in online information, patient portal engagement, and exposure to AI-generated misinformation warrant further investigation. Complete-case analysis reduced the adjusted analytic samples, primarily because of missing household income data. If missingness was systematically related to information burden, some selection bias may remain.

Despite these limitations, this study provides nationally representative evidence regarding cancer information burden in 2022 and 2024. The findings identify demographic and socioeconomic characteristics associated with selected dimensions of information burden during this period. Because the survey did not measure AI-related exposures, the study cannot determine whether technological developments influenced these associations or whether AI-enabled communication systems affected information burden. Future research should directly measure AI tool use and examine its longitudinal associations with patient understanding, trust, well-being, and cancer information burden.

Implications for Cancer and Digital Health Communication

Digital health information systems may create opportunities to improve access to understandable and personalized cancer information. However, their effects on cancer information burden were not evaluated in this study. The observed associations by race and ethnicity, sex, and educational attainment underscore the importance of accessibility, cultural responsiveness, readability, and health literacy–sensitive design in cancer communication. Future evaluations of AI-enabled communication tools should examine whether their benefits and limitations differ across population groups.

Conclusions

These findings reinforce and update prior HINTS-based evidence that cancer information-seeking experiences differ across survivor populations while showing that associations vary by the specific dimension of burden examined. The validated composite measure may support future studies seeking to monitor cancer information burden over time or across communication environments. Hispanic and non-Hispanic Asian survivors, and individuals with lower educational attainment, may experience greater difficulty obtaining, understanding, and evaluating cancer information. Cancer communication strategies should prioritize accessibility, readability, cultural responsiveness, and health literacy. Future studies should directly measure the use of GenAI and other emerging digital tools before evaluating their relationship with cancer information burden.

Acknowledgments

The authors would like to thank the UTHealth Houston School of Public Health for its support of this research.

Generative AI Use

During the preparation of this manuscript, the authors used OpenAI's ChatGPT (model: GPT-6) to assist with language editing, improving clarity, and refining the organization of the manuscript and responses to reviewers. The authors critically reviewed and edited all AI-assisted content and take full responsibility for the accuracy, integrity, and originality of the manuscript. Generative AI was not used to generate, analyze, or interpret study data or to make scientific conclusions.

Funding

This work was supported by startup funds provided by the University of Texas Health Science Center at Houston. The funder had no role in the study design; data collection, analysis, and interpretation; manuscript preparation; or the decision to submit the manuscript for publication.

Data Availability

The data used in this study are publicly available from the National Cancer Institute's Health Information National Trends Survey (HINTS) website.

Authors' Contributions

JG conceived and designed the study, conducted the statistical analyses, interpreted the findings, and wrote the original draft of the manuscript. SVM, YZ, and JHW contributed to the conceptual development of the study, reviewed the manuscript, and provided critical revisions and intellectual input. All authors have read and agreed to the published version of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Distribution of cancer survivors by US census division and Health Information National Trends Survey (HINTS) cycle.
[DOCX File, 15 KB - cancer_v12i1e101310_app1.docx]

References

1. Shaffer KM, Turner KL, Siwik C, et al. Digital health and telehealth in cancer care: a scoping review of reviews. Lancet Digit Health 2023 May;5(5):e316-e327. [doi: 10.1016/S2589-7500(23)00049-3] [Medline: 37100545]
2. Lazarou I, Krooupa AM, Nikolopoulos S, et al. Cancer patients' perspectives and requirements of digital health technologies: a scoping literature review. Cancers (Basel) 2024 Jun 21;16(13):2293. [doi: 10.3390/cancers16132293] [Medline: 39001356]
3. Briggs LG, Labban M, Alkhatib K, Nguyen DD, Cole AP, Trinh QD. Digital technologies in cancer care: a review from the clinician's perspective. J Comp Eff Res 2022 May;11(7):533-544. [doi: 10.2217/cer-2021-0263] [Medline: 35416050]
4. Yun HS, Bickmore T. Online health information–seeking in the era of large language models: cross-sectional web-based survey study. J Med Internet Res 2025:e68560. [doi: 10.2196/68560] [Medline: 40163112]
5. Wardle C, Urbani S, Wang E. Evolving health information–seeking behavior in the context of Google AI Overviews, ChatGPT, and Alexa: interview study using the think-aloud protocol. J Med Internet Res 2025 Oct 7;27:e79961. [doi: 10.2196/79961] [Medline: 41055948]
6. Budenz A, Sleight AG, Klein WMP. A qualitative study of online information-seeking preferences among cancer survivors. J Cancer Surviv 2022 Aug;16(4):892-903. [doi: 10.1007/s11764-021-01082-y] [Medline: 34505207]
7. Miller LE. Uncertainty management and information seeking in cancer survivorship. Health Commun 2014;29(3):233-243. [doi: 10.1080/10410236.2012.739949] [Medline: 23682727]
8. Ferraris G, Monzani D, Coppini V, et al. Barriers to and facilitators of online health information-seeking behaviours among cancer patients: a systematic review. Digit Health 2023;9:20552076231210663. [doi: 10.1177/20552076231210663] [Medline: 38107979]
9. Wu Y, Zhang L, Zhao X. Linking online health information seeking to cancer information overload among Chinese cancer patients' family members. Digit Health 2025;11:20552076251336308. [doi: 10.1177/20552076251336308] [Medline: 40297380]
10. Peng W, Carcioppolo N, Occa A, Ali K, Yang Q, Yang F. Feel worried, overloaded, or fatalistic? The determinants of cancer uncertainty management preferences. Health Commun 2021 Mar;36(3):347-360. [doi: 10.1080/10410236.2019.1692489] [Medline: 31760812]
11. Watson A, van Kessel K. Cancer survivors' experiences and explanations of postcancer fatigue: an analysis of online blogs. Cancer Nurs 2018;41(2):117-123. [doi: 10.1097/NCC.0000000000000460] [Medline: 27941353]
12. Sørensen K. Fostering digital health literacy to enhance trust and improve health outcomes. Comput Methods Programs Biomed Update 2024;5:100140. [doi: 10.1016/j.cmpbup.2024.100140]
13. Papp-Zipernovszky O, Horváth MD, Schulz PJ, Csabai M. Generation gaps in digital health literacy and their impact on health information seeking behavior and health empowerment in Hungary. Front Public Health 2021;9:635943. [doi: 10.3389/fpubh.2021.635943] [Medline: 34055714]
14. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health literacy and health information technology adoption: the potential for a new digital divide. J Med Internet Res 2016 Oct 4;18(10):e264. [doi: 10.2196/jmir.6349] [Medline: 27702738]
15. Davis SN, O'Malley DM, Bator A, Ohman-Strickland P, Hudson SV. Correlates of information seeking behaviors and experiences among adult cancer survivors in the USA. J Cancer Educ 2021 Dec;36(6):1253-1260. [doi: 10.1007/s13187-020-01758-6] [Medline: 32358717]
16. Elkefi S, Matthews AK. Exploring health information-seeking behavior and information source preferences among a diverse sample of cancer survivors: implications for patient education. J Cancer Educ 2024 Dec;39(6):650-662. [doi: 10.1007/s13187-024-02448-3] [Medline: 38807002]
17. Mahmood A, Kedia S, Ogunsanmi DO, Kabir U, Entwistle C. Patient-centered communication and cancer information-seeking experiences among cancer survivors: a population-based study in the United States. Patient Educ Couns 2025 Jun;135:108710. [doi: 10.1016/j.pec.2025.108710] [Medline: 40010060]
18. Benzo RM, Tetrick MK, Krok-Schoen JL, et al. Digital information-seeking behaviors among cancer survivors: associations with sociodemographic determinants, cancer history, and perceived health. J Cancer Surviv 2025 Nov 6:1-12. [doi: 10.1007/s11764-025-01927-w] [Medline: 41196508]
19. Gordon NP, Hornbrook MC. Older adults' readiness to engage with eHealth patient education and self-care resources: a cross-sectional survey. BMC Health Serv Res 2018 Mar 27;18(1):220. [doi: 10.1186/s12913-018-2986-0] [Medline: 29587721]

XSL•FO
RenderX

20. Valdez RS, Rogers CC. Consumer health informatics for racial and ethnic minoritized communities: minor progress, major opportunities. Yearb Med Inform 2022 Aug;31(1):167-172. [doi: 10.1055/s-0042-1742520] [Medline: 36463875]
21. Rousseau J, Gibbs L, Garcia-Cabrera C, et al. A pioneering EMR-embedded digital health literacy tool reveals healthcare disparities for diverse older adults. J Am Geriatr Soc 2024 Aug;72 Suppl 3:S97-S104. [doi: 10.1111/jgs.18935] [Medline: 38682826]
22. Johnson JD, Donohue WA, Atkin CK, Johnson S. A comprehensive model of information seeking: tests focusing on a technical organization. Sci Commun 1995 Mar;16(3):274-303. [doi: 10.1177/1075547095016003003]
23. Ruppel EK. Scanning health information sources: applying and extending the comprehensive model of information seeking. J Health Commun 2016;21(2):208-216. [doi: 10.1080/10810730.2015.1058438] [Medline: 26716985]
24. Yang Q, Van Stee SK, Rains SA. Comprehensive model of information seeking: a meta-analysis. J Health Commun 2023 Jun 3;28(6):360-374. [doi: 10.1080/10810730.2023.2214097] [Medline: 37293744]
25. Reifegerste D, Blech S, Dechant P. Understanding information seeking about the health of others: applying the comprehensive model of information seeking to proxy online health information seeking. J Health Commun 2020 Feb 1;25(2):126-135. [doi: 10.1080/10810730.2020.1716280] [Medline: 32009552]
26. Chesser A, Burke A, Reyes J, Rohrberg T. Navigating the digital divide: a systematic review of eHealth literacy in underserved populations in the United States. Inform Health Soc Care 2016;41(1):1-19. [doi: 10.3109/17538157.2014.948171] [Medline: 25710808]
27. López L, Green AR, Tan-McGrory A, King R, Betancourt JR. Bridging the digital divide in health care: the role of health information technology in addressing racial and ethnic disparities. Jt Comm J Qual Patient Saf 2011 Oct;37(10):437-445. [doi: 10.1016/s1553-7250(11)37055-9] [Medline: 22013816]
28. Whitehead L, Talevski J, Fatehi F, Beauchamp A. Barriers to and facilitators of digital health among culturally and linguistically diverse populations: qualitative systematic review. J Med Internet Res 2023 Feb 28;25:e42719. [doi: 10.2196/42719] [Medline: 36853742]
29. Gorski P. Education equity and the digital divide. AACE Rev 2005;13(1):3-45 [FREE Full text]
30. Afzal A, Khan S, Daud S, Ahmad Z, Butt A. Addressing the digital divide: access and use of technology in education. J Soc Sci Rev 2023 Jun 30;3(2):883-895. [doi: 10.54183/jssr.v3i2.326]
31. Vaidya V, Partha G, Karmakar M. Gender differences in utilization of preventive care services in the United States. J Womens Health (Larchmt) 2012 Feb;21(2):140-145. [doi: 10.1089/jwh.2011.2876] [Medline: 22081983]
32. Figueroa CA, Luo T, Aguilera A, Lyles CR. The need for feminist intersectionality in digital health. Lancet Digit Health 2021 Aug;3(8):e526-e533. [doi: 10.1016/S2589-7500(21)00118-7] [Medline: 34325855]

Abbreviations

GenAI: generative AI
HINTS: Health Information National Trends Survey
OR: odds ratio

Edited by M Balcarras; submitted 13.May.2026; peer-reviewed by A Chakraborty, J Stimpson, M Sawaya, S Kedia; revised version received 21.Aug.2026; accepted 21.Aug.2026; published 28.Oct.2026.

Please cite as:
Gao J, Munroe SV, Zhang Y, Windett JH
Cancer Information Burden Among US Cancer Survivors Using Health Information National Trends Survey 2022 and 2024: Cross-Sectional Study
JMIR Cancer 2026;12:e101310
URL: https://cancer.jmir.org/2026/1/e101310
doi:10.2196/101310

© Jingjing Gao, Sharon V Munroe, Yue Zhang, Jason Windett. Originally published in JMIR Cancer (https://cancer.jmir.org), 28.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predictors of Telehealth Use Among Cancer Survivors: Retrospective Study

Jasskiran Kaur, MD; Jeffrey Zhong, MD; Carley Mitchell, MD; Annie Zhang, MD; Qian Wang, MD; Melinda L Hsu, MD

University Hospitals Seidman Cancer Center, 11100 Euclid Avenue, Cleveland, OH, United States

Corresponding Author:
Melinda L Hsu, MD
University Hospitals Seidman Cancer Center, 11100 Euclid Avenue, Cleveland, OH, United States

Abstract

Background: As the number of cancer survivors continues to grow, optimizing long-term survivorship care models has become increasingly important. Telehealth has the potential to improve access to health care for survivors; however, studies evaluating telehealth in this population remain limited. Additionally, concerns persist regarding equity in technology access and digital literacy.

Objective: This study aimed to examine demographic factors and patient attitudes influencing telehealth use among cancer survivors compared to the general population.

Methods: Adult participants were identified from the nationally representative database Health Information National Trends Survey 6 (HINTS 6). Multivariable logistic regression was used to calculate the predictors of telehealth use among cancer survivors. χ^2 tests compared the prevalence of reported reasons of not using telehealth in the last 12 months between cancer survivors and the general population.

Results: A total of 5793 (weighted n=239,557,883) individuals were included in this study, 7.7% (weighted n=18,545,434) who are cancer survivors. 5092 individuals from the general population and 701 cancer survivors were included. Older age was associated with lower telehealth use (adjusted odds ratio [aOR] 0.11; 95% CI 0.02 - 0.59 for patients aged ≥65, compared to those under 40 y old). Higher education (aOR 2.55; 95% CI 1.24 - 5.27) and heart disease history (aOR 2.52; 95% CI 1.20 - 5.28) were associated with increased telehealth use. Employed (aOR 0.46; 95% CI 0.22 - 0.97) and retired (aOR=0.37; 95% CI: 0.18 - 0.77) cancer survivors were less likely to use telehealth than unemployed individuals. Of the nonusers, over 60% reported that telehealth options were not offered, and 80% preferred in-person visits. Technical issues and privacy concerns were not major factors in utilizing telehealth.

Conclusions: Despite greater telehealth use among cancer survivors, a negative association between older age and telemedicine utilization persists. Efforts should focus on improving access for older cancer survivors and addressing employment-related factors, patient attitudes, and telehealth availability. Future studies should explore personalized approaches to enhance cancer survivors' health care experiences. Our findings emphasize the need to address specific factors including age and employment related disparities, patient preferences and telehealth availability to optimize equitable access to telehealth and enhance the delivery of cancer survivorship care.

(JMIR Cancer 2026;12:e87403) doi:10.2196/87403

KEYWORDS

survivorship; telehealth; digital health; internet; disparities; healthcare delivery; cancer

Introduction

Cancer survivors represent approximately 5.4% of the US population, and the number of survivors is expected to increase by 7.9 million by 2040 [1]. As the number of cancer survivors grows, the need to optimize long-term survivorship care models is imperative. Cancer survivors face unique health challenges, as they require continued monitoring for risk of recurrence, assessment for long-term physical and psychosocial effects of both the disease and its treatments, management of comorbidities, screening for new malignancies, and health promotion. Moreover, cancer survivors may face considerable economic burden for myriad reasons, including the inability to work and the rising costs of cancer therapies [2-4]. Patients with cancer and cancer survivors represented a vulnerable population during the rise of the COVID-19 pandemic, and fear associated with contracting the virus significantly affected all aspects of medical care [5-7]. As a result, telehealth offered potential health care delivery solutions for all socially distanced patients [8].

XSL•FO
RenderX

Telehealth is defined as the delivery of health and health-related services via telecommunication technologies [9] and has been successfully implemented over the last two decades in noncancer contexts including diabetes care, hypertension, mental health, and chronic pain [10-13]. Telehealth for cancer patients was evaluated before the COVID-19 pandemic and was shown to be as effective at improving quality of life as in-person care [14]. Although prior studies have examined telehealth use among cancer survivors, these have largely been limited to specific cancer types [15,16]. Data on telehealth utilization among cancer survivors, particularly relative to the general population, remain limited. In our analyses, the general population serves as the reference group, enabling assessment of whether cancer survivorship is independently associated with the outcomes of interest.

While telehealth is a powerful tool with the potential to improve health care outcomes, inequities remain regarding access to the technologies required for successful telehealth interactions. This is exemplified by a decrease in telehealth utilization by adults who are uninsured, have lower incomes, and are without high school degrees [17]. Older adults may have less digital literacy as well, and a small qualitative study found older cancer survivors made use of online platforms for general health concerns but preferred in-person visits for major issues [18,19]. Further research regarding the use of and potential barriers to telehealth services is imperative to continued improvement of survivorship-care. Our study aims to analyze demographic factors and patient attitudes that influence telehealth uptake among cancer survivors compared to the general population.

Methods

This study was a retrospective cross-sectional analysis of publicly available data from the Health Information National Trends Survey 6 (HINTS 6) [20]. HINTS data were assessed from the National Cancer Institute HINTS website. which collects nationally representative data about how adults use different communication channels to obtain health information. Telehealth visits were defined in HINTS 6 as a telephone or video appointment. Telehealth data was extracted for adult patients from March 7, 2022, to November 8, 2022. A total of 5793 individuals were included in this study, 5092 individuals from the general population and 701 cancer survivors. Our outcome of interest was overall telehealth utilization rates, reasons for having telehealth visits, and patient perceptions of telehealth utilization.

Outcome Variables

Participants were first asked whether they had used telehealth (defined as a telephone or video visit with a health care provider) in the past 12 months and whether they had been offered a telehealth option for care. Telehealth was assessed using items from section D of HINTS 6 [20].

Among those who used telehealth, additional questions assessed reasons for choosing telehealth (eg, provider recommendation, convenience, infection avoidance, or ability to include caregivers). Participants were also asked about their experiences and perceptions of telehealth, including preferences for in-person versus telehealth visits, concerns about privacy, perceived difficulty using the technology, technical problems encountered, and whether the quality of care was comparable to in-person visits.

Among those who did not use telehealth, items assessed whether telehealth was offered and reasons for nonuse (eg, preferred in person, privacy concerns, perceived difficulty using the technology)

Statistical analysis included a χ^2 test to compare the prevalence of telehealth use in the last 12 months between cancer survivors and the general population. Additionally, multivariable logistic regression was used to calculate the adjusted odds ratio (aOR) of telehealth use in cancer survivors compared to the general population. Our model was adjusted for multiple variables — age, sex, race/ethnicity income, education, marital status, BMI, employment status, insurance status, physical activity, diabetes, hypertension, lung disease, heart disease and smoking status.

All analyses incorporated the HINTS complex survey weights, and variance estimates were calculated using the Taylor series linearization method to account for the survey's stratified sampling design. Listwise deletions were used to treat missing values across these variables. All statistical analyses were performed with SAS 9.4. A 2-sided *P* value of <.05 was considered statistically significant.

Ethical Considerations

HINTS is deidentified, publicly available data, and has been designated as exempt research and approved by the Westat IRB [21].

Results

Patient Characteristics

A total of 239,557,883 individuals were included in the weighted sample, of which 7.7% (weighted n=18,545,434) were cancer survivors. 5092 individuals from the general population and 701 cancer survivors were included. Cancer survivors differed from the general population across multiple sociodemographic and clinical characteristics (Table 1). As expected, cancer survivors were more likely to be older (*P*<.001) and retired (*P*<.001). Cancer survivors were also more likely than the general population to be female (*P*<.001), non-Hispanic White (*P*<.001), married (*P*=.01), and have insurance (*P*<.001). Clinically, cancer survivors were more likely to be smokers (*P*<.001) and have additional medical comorbidities including diabetes (*P*=.002), hypertension (*P*<.001), heart disease (*P*<.001), and lung disease (=.003).

XSL•FO
RenderX

Table . Table 1. Demographic features comparing general population vs noncancer survivors 2022 (Weighted N=239,557,883)[a].

Variables	General population n=5092 (weighted n=221,012,449); 92.3%		Cancer survivors n=701 (weighted n=18,545,434); 7.7%		*P* value
	Unweighted n	Weighted % (95% CI)	Unweighted n	Weighted % (95% CI)	
Age					<.001
18 - 39	1213	35.2 (32.6 - 37.7)	17	3.6 (1.3 - 5.8)	
40 - 64	2202	45.5 (43.1 - 48.0)	215	39.8 (33.6 - 46.0)	
65+	1677	19.3 (17.9 - 20.7)	469	56.7 (50.3 - 63.0)	
Sex					<.001
Female	3068	49.9 (47.2 - 52.6)	437	61.8 (56.6 - 66.9)	
Male	2024	50.1 (47.4 - 52.8)	264	38.2 (33.1 - 43.4)	
Race					<.001
NHW[b]	2731	60.1 (57.4 - 62.8)	450	77.1 (72.3 - 81.9)	
NHB[c]	791	11.3 (10.2 - 12.5)	90	7.4 (5.3 - 9.5)	
Hispanic	917	17.5 (15.4 - 19.5)	68	9.0 (5.7 - 12.2)	
Others	429	11.1 (8.6 - 13.6)	39	6.6 (3.7 - 9.4)	
Education					.53
≤ High school	1245	28.4 (26.4 - 30.5)	180	26.2 (22.0 - 30.5)	
Some college	1446	38.9 (36.3 - 41.5)	212	42.1 (37.1 - 47.0)	
≥ College	2385	32.7 (30.5 - 34.9)	304	31.7 (28.0 - 35.4)	
Income (USD)					.78
≤50,000	2210	37.4 (35.1 - 39.8)	322	36.7 (31.8 - 41.5)	
>50,000	2876	62.6 (60.2 - 64.9)	376	63.3 (58.5 - 68.2)	
Marital status					.01
Unmarried	2442	44.7 (42.4 - 47.1)	352	36.9 (31.5 - 42.4)	
Married	2626	55.3 (52.9 - 57.6)	344	63.1 (57.6 - 68.5)	
Employment status					<.001
Unemployed	1212	27.1 (24.0 - 30.1)	142	21.5 (16.3 - 26.7)	
Employed	2556	56.2 (53.4 - 59.0)	180	30.2 (23.7 - 36.8)	
Retired	1317	16.7 (15.4 - 18.0)	374	48.3 (41.9 - 54.7)	
Insurance status					<.001
No	446	11.0 (9.3 - 12.8)	13	2.8 (0.5 - 5.1)	
Yes	4612	89.0 (87.2 - 90.7)	687	97.2 (94.9 - 99.5)	
Smoking status					<.001
Never	3293	66.8 (64.7 - 68.8)	373	50.9 (45.1 - 56.8)	
Current	554	12.0 (10.4 - 13.6)	73	12.4 (8.4 - 16.3)	
Former	1187	21.2 (19.6 - 22.8)	245	36.7 (30.9 - 42.5)	
BMI					.62
Underweight/normal	1519	31.4 (29.2 - 33.5)	218	30.1 (24.9 - 35.2)	
Overweight	1664	33.6 (31.1 - 36.2)	241	36.5 (31.5 - 41.5)	
Obese	1810	35.0 (32.7 - 37.2)	229	33.4 (28.3 - 38.6)	
Moderate intensivity activity (min/wk)					.06
≤60	2210	43.0 (40.8 - 45.2)	326	51.6 (45.3 - 57.9)	

XSL•FO
RenderX

Variables	General population n=5092 (weighted n=221,012,449); 92.3%		Cancer survivors n=701 (weighted n=18,545,434); 7.7%		*P* value
	Unweighted n	Weighted % (95% CI)	Unweighted n	Weighted % (95% CI)	
60 - 150	1125	21.2 (19.2 - 23.1)	139	19.9 (15.2 - 24.6)	
151 - 300	1006	21.7 (18.9 - 24.5)	134	17.3 (13.0 - 21.7)	
>300	644	14.2 (12.3 - 16.1)	79	11.1 (8.0 - 14.3)	
Diabetes	1050	16.6 (15.0 - 18.3)	198	25.8 (20.8 - 30.8)	.001
Hypertension	2194	36.0 (33.7 - 38.2)	428	55.5 (51.0 - 59.9)	<.001
Heart disease	462	7.0 (6.0 - 7.9)	120	16.3 (12.4 - 20.1)	<.001
Lung disease	658	11.4 (10.1 - 12.7)	141	19.6 (14.7 - 24.4)	.003
Depression	1674	35.6 (33.4 - 37.9)	214	32.7 (27.6 - 37.8)	.32
Cancer information					
Type of cancer					—[d]
Liquid			38	4.0 (2.2 - 5.8)	
Solid			469	67.8 (62.8 - 72.8)	
Others			186	28.3 (23.2 - 33.3)	
Time since diagnosis					—[d]
≤5 y			272	34.6 (29.5 - 39.7)	
>5 y			587	65.4 (60.3 - 70.5)	

[a]Values are presented as unweighted n with weighted percentages and 95% confidence intervals. Percentages and confidence intervals were calculated using survey weights and therefore may not correspond directly to the unweighted counts.

[b]NHW: Non-Hispanic White.

[c]NHB: Non-Hispanic Black.

[d]P-values are not reported because these variables were collected exclusively in the cancer cohort and were not compared between groups.

Telehealth use was more common among cancer survivors than the general population (n=320, 48.2% vs n=2042, 38.4%; *P*=.01, aOR 1.48; 95% CI 1.07 - 2.04) (Table 2). Among cancer survivors who utilized telehealth, older age was associated with a lower likelihood of using telehealth compared to those under 40 years old (aOR 0.11; 95% CI 0.02 - 0.59 for age ≥65). Additionally, telehealth usage among cancer survivors was greater among those with higher education levels (aOR 2.55; 95% CI 1.24 - 5.27) and a history of heart disease (aOR 2.52; 95% CI 1.20 - 5.28). However, cancer survivors who were employed (aOR 0.46; 95% CI 0.22 - 0.97) or retired (aOR 0.37; 95% CI 0.18 - 0.77) were less likely to use telehealth than those who were unemployed (Table 3).

XSL•FO
RenderX

Table . Comparing telehealth use and perception of telehealth use, comparing cancer survivors to the general population.

Telehealth utilization	General population n=5092 (weighted n=221,012,449); 92.3%		Cancer survivors n=701 (weighted n=18,545,434); 7.7%		*P* value	aOR[a] 95% CI[b]
	Unweighted n	Weighted % (95% CI)	Unweighted n	Weighted % (95% CI)		
Overall						
Had telehealth	2042	38.4 (36.2 - 40.5)	320	48.2 (41.4 - 54.9)	.01	1.48 (1.07 - 2.04)
Among individuals who used telehealth						
Reason for having televisit						
Recommended by provider	1364	72.9 (70.0 - 75.7)	224	79.4 (73.3 - 85.4)	.07	1.51 (0.97 - 2.35)
Avoid infection	947	51.5 (47.5 - 55.5)	114	36.9 (27.6 - 46.2)	.008	0.63 (0.40 - 0.996)
Convenience	1242	66.6 (63.1 - 70.0)	158	56.6 (47.2 - 66.1)	.05	0.80 (0.50 - 1.30)
Including friends/families	378	22.6 (19.1 - 26.2)	68	23.5 (15.3 - 31.8)	.84	0.99 (0.55 - 1.76)
Having technical difficulty in telehealth visit	380	18.2 (14.7 - 21.7)	80	25.7 (18.2 - 33.1)	.08	1.36 (0.82 - 2.25)
Think telehealth is as good as in person	1439	75.8 (72.0 - 79.6)	214	75.3 (69.3 - 81.3)	.90	0.90 (0.60 - 1.33)
Has privacy concerns with telehealth	285	15.9 (11.7 - 20.1)	40	10.4 (6.1 - 14.7)	.07	0.43 (0.21 - 0.89)

[a]aOR: adjusted odds ratio.

[b]Adjusted for age, sex, race, income, education, marital status, BMI, employment status, insurance status, physical activity, diabetes, hypertension, lung disease, heart disease and smoking status.

XSL•FO
RenderX

Table . Predictors of receiving telehealth.

Variable	Using telehealth
	aOR[a] 95% CI
Cancer survivor	
No	Ref
Yes	1.48 (1.07 - 2.04)[b]
Among cancer survivors only - predicting factors[c]	
Age	
18 - 39	Ref
40 - 64	0.18 (0.03 - 1.03)
65+	0.11 (0.02 - 0.59)
Gender	
Female	Ref
Male	1.54 (0.87 - 2.71)
Race	
NHW[d]	Ref
NHB[e]	0.75 (0.33 - 1.70)
Hispanic	1.77 (0.70 - 4.50)
Others	0.88 (0.35 - 2.25)
Education	
≤ High school	Ref
Some college	1.93 (0.98 - 3.82)
≥ College	2.55 (1.24 - 5.27)
Income (USD)	
≤50,000	Ref
>50,000	0.96 (0.53 - 1.74)
Marital status	
Unmarried	Ref
Married	0.90 (0.48 - 1.70)
Employment status	
Unemployed	Ref
Employed	0.46 (0.22 - 0.97)
Retired	0.37 (0.18 - 0.77)
Insurance status	
No	Ref
Yes	1.52 (0.30 - 7.60)
Smoking status	
Never	Ref
Current	0.62 (0.24 - 1.61)
Former	1.40 (0.80 - 2.44)
BMI	
Underweight/normal	Ref
Overweight	0.72 (0.38 - 1.36)

XSL•FO
RenderX

Variable	Using telehealth
	aOR[a] 95% CI
Obese	1.06 (0.50 - 2.23)
Moderate intensity activity (min/wk)	
≤60	Ref
60 - 150	1.43 (0.68 - 3.00)
151 - 300	0.72 (0.35 - 1.49)
>300	1.58 (0.69 - 3.63)
Diabetes	
No	Ref
Yes	0.89 (0.51 - 1.55)
Hypertension	
No	Ref
Yes	1.35 (0.77 - 2.35)
Heart disease	
No	Ref
Yes	2.52 (1.20 - 5.28)
Lung disease	
No	Ref
Yes	1.57 (0.86 - 2.88)
Cancer information	
Type of cancer	
Solid	Ref
Liquid	1.07 (0.34 - 3.35)
Others	0.82 (0.44 - 1.54)
Time since diagnosis (years)	
≤5	Ref
>5	0.95 (0.52 - 1.75)

[a]aOR: adjusted odds ratio.

[b]Model adjusted for cancer survivor status, age, sex, race/ethnicity, income, education, marital status, BMI, employment status, insurance status, physical activity, diabetes, hypertension, lung disease, heart disease and smoking status.

[c]Model adjusted for age, sex, race/ethnicity, income, education, marital status, BMI, employment status, insurance status, physical activity, diabetes, hypertension, lung disease, heart disease, smoking status, type of cancer, and time since cancer diagnosis.

[d]NHW: Non-Hispanic White.

[e]NHB: Non-Hispanic Black.

Among those utilizing telehealth, avoiding infection was a more frequent motivator for the general population compared to cancer survivors (n=947, 51.5% vs n=114, 36.9%; *P*=.008, aOR 0.63; 95% CI 0.40 - 0.996). No significant differences were observed between cancer survivors and the general population in other reasons for telehealth use, including provider recommendation, convenience, or the ability to include family and friends. Similarly, there were no significant differences in perceptions of telehealth, including perceived difficulties, or equivalence to in-person visits (Table 2). Although privacy concerns were significantly less likely among cancer survivors after adjustment for covariates (aOR 0.43; 95% CI 0.21 - 0.89) (Table 2), the unadjusted comparison did not reach statistical significance (*P*=.07).

Among individuals who had not used telehealth in the past 12 months, over 80% (general population, n =411 and cancer survivors, n=68) did not use telehealth because they preferred in-person visits, and over 60% (general population, n=1875 and cancer survivors, n=231) of both groups reported that their health care providers did not offer telehealth services (Table 4). There were no significant differences in privacy concerns or technical difficulties as reasons for not utilizing telehealth (Table 2).

XSL•FO
RenderX

Table . Among those who did not have telehealth last year, exploring reasons—comparing cancer survivors versus the general population.

Variables	General population n=5092 (weighted n=221,012,449); 92.3%		Cancer survivors n=701 (weighted n=18,545,434); 7.7%		*P* value	aOR[a] 95% CI[b]
	Unweighted n	Weighted % (95% CI)	Unweighted n	Weighted % (95% CI)		
Offered by providers					.02	
Yes	474	16.6 (14.1 - 19.1)	76	20.6 (14.4 - 26.7)		1.60 (1.02 - 2.49)
No	1875	62.9 (59.6 - 66.1)	231	67.6 (60.9 - 74.3)		1.24 (0.87 - 1.78)
Did not schedule appointment	532	20.5 (17.5 - 23.5)	41	11.8 (6.7 - 16.9)		0.82 (0.45 - 1.48)
Prefer in person	411	83.7 (73.8 - 93.5)	68	88.9 (78.1 - 99.6)	.48	0.35 (0.06 - 2.02)
Privacy concerns	71	17.1 (7.4 - 26.7)	26	14.9 (3.8 - 26.0)	.76	1.41 (0.39 - 5.07)
Technical difficulty	95	17.8 (12.3 - 23.3)	16	22.1 (10.6 - 33.6)	.51	0.81 (0.30 - 2.17)

[a]aOR: adjusted odds ratio.

[b]Using the general population as the reference group, model was adjusted for age, sex, race, income, education, marital status, BMI, employment status, insurance status, physical activity, diabetes, hypertension, lung disease, heart disease and smoking status.

Discussion

Principal Results

Our study retrospectively analyzed telehealth utilization in both the general population and cancer survivors, using the HINTS 6 data. We identified several factors that may affect the likelihood of cancer survivors utilizing telehealth.

We found older age and lower education level in cancer survivors to be associated with decreased use of telehealth, which is consistent with results from prior research [17-19,22-24]. Chen et al [25] found that prostate cancer survivors with less than a high school education had lower use of telehealth than those who had received any college education. Moreover, there is evidence that cancer patients with higher education levels have more positive perceptions of telemedicine with regards to satisfaction and privacy [26]. Along these same lines, Kjeldsted et al [27] noted that cancer patients with lower education levels reported less positive experiences with telemedicine, specifically with regard to comfort and confidence levels in teleconsultations.

Cancer survivors in our study with a history of heart disease were more likely to use telehealth. Inglist et al [28] published a Cochrane systematic review identifying 41 randomized controlled trials and found telehealth interventions reduced the relative risk of all-cause mortality by 20%. Consequently, various telehealth interventions have been used in heart disease management including structured telephone or video support for monitoring risk factors monitoring such as hypertension, and remote monitoring or wearable implantable devices including arrhythmia detection. These types of interventions familiarize many patients to telehealth technologies, therefore they may be more willing to engage in them in other settings like oncology [29].

Unlike previous research, we did not find privacy concerns or technical difficulties as major factors associated with reduced telehealth use [30]. Lower odds of older cancer survivors using telehealth may support prior evidence regarding the digital divide; this indicates the need to investigate other potential barriers for older cancer patients and use of telehealth including access to the internet and technology. A study from a primary care setting found that patient portal use was 28%, and patients with the lowest broadband access were less likely to utilize their patient portal [31].

Among the patients who had not utilized telehealth services over the prior 12 months, our study found that 60% of both cancer survivors and the general population were not offered this option.

This may be a result of system-level factors such as medical licensure, payment models and lack of training for implementation of telehealth interventions [32]. However, telehealth regulatory changes during the pandemic led to payment parity and equal reimbursement for telehealth visits as well as elimination of previous licensure requirements [33,34]. Furthermore, patient privacy regulations, particularly certain HIPAA (Health Insurance Portability and Accountability ACT of 1996) requirements, including the use of various video conferencing platforms, were waived during the pandemic [34]. Sequentially, a US Department of Health and Human Services report found a 63-fold increase in telehealth utilization [35]. Hopefully this trend in increased uptake will continue.

Preceding research has established potential benefits of telehealth compared to in person care, including improved quality of life, decreased costs [36,37], and improved access to health care for patients in rural communities [38]. However, research also suggests that telehealth may exacerbate inequities by further increasing the digital divide [39]. To mitigate this risk, specific guidelines and personalized care will need to be implemented. The Journal of the National Cancer Institute has recently published the Framework for Integrating Telehealth Equitably (FITE), which is intended to guide equitable implementation of telehealth [40]. Potential interventions, such

XSL•FO
RenderX

as the Telehealth Task Force described by Worster et al [41], have demonstrated that providing tailored digital support can successfully engage older, racially diverse, and socioeconomically disadvantaged cancer patients, thereby helping to reduce the digital divide. A designated task force or telehealth navigators may similarly assist in addressing the needs of older cancer survivors, and those with low health literacy. The lower telehealth utilization observed among employed cancer survivors compared to their unemployed counterparts may reflect scheduling constraints inherent to the working population, such as limited availability during standard clinic hours or difficulty finding a private setting for a telehealth visit during the workday. To address such barriers, the Multinational Association for Supportive Care in Cancer (MASCC)–ASCO standards and practice recommendations for survivorship care emphasize flexible scheduling and the inclusion of both video and audio-only appointment options to increase accessibility across diverse patient populations [42].

Limitations

Our study does have limitations, as it focuses on data from a snapshot of time in 2022. It should also be noted that this data was collected during the COVID-19 pandemic, which had a large impact on the utilization of telehealth, and it is unclear how telehealth is being used in the post-pandemic setting. The strengths of our study include the inclusion of all cancer types, the large sample size, and the use of a model adjusted for multiple variables—age, sex, race/ethnicity income, education, marital status, BMI, employment status, insurance status, physical activity, diabetes, hypertension, lung disease, heart disease and smoking status—which allowed us to more directly compare the cancer survivor population to the general population.

Conclusion

As the health care delivery landscape continues to transform, understanding the factors that influence telehealth utilization among cancer survivors is imperative for ensuring equity and access to care, as well as tailoring interventions to meet the unique needs of this population.

Although cancer survivors in our study were significantly more likely than the general population to use telehealth, fewer than half did so. Digital literacy may not be a barrier to telehealth use in cancer survivors as they did not experience more difficulties in their telehealth visits.

The results of our study highlight the need for further efforts to mitigate barriers to telehealth uptake by cancer survivors. Health care providers also need to consider patients' specific concerns and preferences when promoting telehealth. Future studies are needed to explore and develop personalized approaches to health care delivery, with the goal of optimizing cancer survivors' outcomes and experiences.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data Availability

The datasets generated and analyzed during the current study are available in the nationally representative database, Health Information National Trends Survey 6 (HINTS 6) [43].

Authors' Contributions

Conceptualization: MH, QW
Data curation: AZ, CM
Formal analysis: AZ, CM
Methodology: MH, QW
Writing – original draft: JK, JZ
Writing – review & editing: AZ, CM, JK, JZ, MH, QW

Conflicts of Interest

None declared.

References

1. Tonorezos E, Devasia T, Mariotto AB, et al. Prevalence of cancer survivors in the United States. J Natl Cancer Inst 2024 Nov 1;116(11):1784-1790. [doi: 10.1093/jnci/djae135] [Medline: 39002121]
2. Nekhlyudov L, Mollica MA, Jacobsen PB, Mayer DK, Shulman LN, Geiger AM. Developing a quality of cancer survivorship care framework: implications for clinical care, research, and policy. J Natl Cancer Inst 2019 Nov 1;111(11):1120-1130. [doi: 10.1093/jnci/djz089] [Medline: 31095326]
3. Guy GP Jr, Ekwueme DU, Yabroff KR, et al. Economic burden of cancer survivorship among adults in the United States. J Clin Oncol 2013 Oct 20;31(30):3749-3757. [doi: 10.1200/JCO.2013.49.1241] [Medline: 24043731]
4. Zheng Z, Jemal A, Han X, et al. Medical financial hardship among cancer survivors in the United States. Cancer 2019 May 15;125(10):1737-1747. [doi: 10.1002/cncr.31913] [Medline: 30663039]

XSL•FO
RenderX

5. Paterson C, Bacon R, Dwyer R, et al. The role of telehealth during the COVID-19 pandemic across the interdisciplinary cancer team: implications for practice. Semin Oncol Nurs 2020 Dec;36(6):151090. [doi: 10.1016/j.soncn.2020.151090] [Medline: 33218886]
6. Riera R, Bagattini Â, Pacheco RL, Pachito DV, Roitberg F, Ilbawi A. Delays and disruptions in cancer health care due to COVID-19 pandemic: systematic review. JCO Glob Oncol 2021 Feb;7:311-323. [doi: 10.1200/GO.20.00639] [Medline: 33617304]
7. Desai A, Gupta R, Advani S, et al. Mortality in hospitalized patients with cancer and coronavirus disease 2019: A systematic review and meta-analysis of cohort studies. Cancer 2021 May 1;127(9):1459-1468. [doi: 10.1002/cncr.33386] [Medline: 33378122]
8. Gilroy KJ. Incorporating telemedicine as part of COVID-19 outbreak response systems. Am J Manag Care 2020 Apr 1;26(4):147-148. [doi: 10.37765/ajmc.2020.42784]
9. What is telehealth. NEJM Catal. 2018. URL: https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0268 [accessed 2025-07-25]
10. Pratt SI, Bartels SJ, Mueser KT, et al. Feasibility and effectiveness of an automated telehealth intervention to improve illness self-management in people with serious psychiatric and medical disorders. Psychiatr Rehabil J 2013 Dec;36(4):297-305. [doi: 10.1037/prj0000022] [Medline: 24320837]
11. Garg SK, Parkin CG. The emerging role of telemedicine and mobile health technologies in improving diabetes care. Diabetes Technol Ther 2019 Jun;21(2_suppl):2019. [doi: 10.1089/dia.2019.0090]
12. Wakefield BJ, Holman JE, Ray A, et al. Effectiveness of home telehealth in comorbid diabetes and hypertension: a randomized, controlled trial. Telemed J E Health 2011 May;17(4):254-261. [doi: 10.1089/tmj.2010.0176] [Medline: 21476945]
13. Kroenke K, Krebs EE, Wu J, Yu Z, Chumbler NR, Bair MJ. Telecare collaborative management of chronic pain in primary care: a randomized clinical trial. JAMA 2014 Jul 16;312(3):240-248. [doi: 10.1001/jama.2014.7689] [Medline: 25027139]
14. Larson JL, Rosen AB, Wilson FA. The effect of telehealth interventions on quality of life of cancer patients: a systematic review and meta-analysis. Telemed J E Health 2018 Jun;24(6):397-405. [doi: 10.1089/tmj.2017.0112] [Medline: 29112484]
15. Ajmera P, Miraj M, Kalra S, et al. Impact of telehealth interventions on physiological and psychological outcomes in breast cancer survivors: a meta-analysis of randomised controlled trials. Front Oncol 2022;12:1017343. [doi: 10.3389/fonc.2022.1017343] [Medline: 36686741]
16. Damholdt MF, Mehlsen M, O'Toole MS, Andreasen RK, Pedersen AD, Zachariae R. Web-based cognitive training for breast cancer survivors with cognitive complaints-a randomized controlled trial. Psychooncology 2016 Nov;25(11):1293-1300. [doi: 10.1002/pon.4058] [Medline: 26763774]
17. Karimi M, Lee E, Couture S, et al. National Survey Trends in Telehealth Use in 2021: Disparities in utilization and audio vs video services. Assistant Secretary for Planning and Evaluation Health and Human Services. 2022. URL: https://aspe.hhs.gov/sites/default/files/documents/fdd365e43a2bbb925e2df805a6ffda94/telehealth-hps-ib.pdf
18. Arthur EK, Pisegna J, Oliveri JM, Aker H, Krok-Schoen JL. Older cancer survivors' perspectives and use of telehealth in their cancer survivorship care in the United States: a ResearchMatch® sample. J Geriatr Oncol 2022 Nov;13(8):1223-1229. [doi: 10.1016/j.jgo.2022.08.004] [Medline: 35985929]
19. Knapp R. Telehealth and equity - AP-NORC. AP-NORC. 2022. URL: https://apnorc.org/projects/telehealth-and-equity/
20. National Cancer Institute. Health Information National Trends Survey 6 (HINTS 6): Annotated English Instrument.: National Cancer Institute; 2022. URL: https://hints.cancer.gov/docs/Instruments/HINTS6-AnnotatedEnglishInstrument.pdf [accessed 2026-04-01]
21. National Cancer Institute. Institutional Review Board (IRB) approvals for the Health Information National Trends Survey (HINTS). URL: https://hints.cancer.gov/about-hints/institutional-review-board.aspx [accessed 2026-04-09]
22. Jewett PI, Vogel RI, Ghebre R, et al. Telehealth in cancer care during COVID-19: disparities by age, race/ethnicity, and residential status. J Cancer Surviv 2022 Feb;16(1):44-51. [doi: 10.1007/s11764-021-01133-4] [Medline: 34800257]
23. Anderson M, Anderson M. 1 Technology use among seniors.: Pew Research Center; 2024. URL: https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/ [accessed 2025-01-25]
24. Qian AS, Schiaffino MK, Nalawade V, et al. Disparities in telemedicine during COVID-19. Cancer Med 2022 Feb;11(4):1192-1201. [doi: 10.1002/cam4.4518] [Medline: 34989148]
25. Chen LW, Usinger DS, Katz AJ. Telehealth use and perceptions among prostate cancer survivors. Cancer Med 2023 Aug;12(16):17308-17312. [doi: 10.1002/cam4.6328] [Medline: 37455582]
26. Picardo E, Baù MG, Anatrone C, et al. Oncophone20 study: Patients' perception of telemedicine in the COVID‐19 pandemic during follow‐up visits for gynecological and breast cancers. Intl J Gynecology & Obste 2021 Dec;155(3):398-403. [doi: 10.1002/ijgo.13825]
27. Kjeldsted E, Lindblad KV, Bødtcher H, et al. A population-based survey of patients' experiences with teleconsultations in cancer care in Denmark during the COVID-19 pandemic. Acta Oncol 2021 Oct;60(10):1352-1360. [doi: 10.1080/0284186X.2021.1956688] [Medline: 34338113]
28. Inglis SC, Clark RA, Dierckx R, Prieto-Merino D, Cleland JG. Structured telephone support or non-invasive telemonitoring for patients with heart failure. Cochrane Libr 2015;2015(10). [doi: 10.1002/14651858.CD007228.pub3]

XSL•FO
RenderX

29. Takahashi EA, Schwamm LH, Adeoye OM, et al. An overview of telehealth in the management of cardiovascular disease: a scientific statement from the American Heart Association. Circulation 2022 Dec 20;146(25):e558-e568. [doi: 10.1161/CIR.0000000000001107] [Medline: 36373541]
30. Kato-Lin YC, Thelen ST. Privacy concerns and continued use intention of telemedicine during COVID-19. Telemed J E Health 2022 Oct;28(10):1440-1448. [doi: 10.1089/tmj.2021.0603] [Medline: 35235434]
31. Rodriguez JA, Lipsitz SR, Lyles CR, Samal L. Association between patient portal use and broadband access: a national evaluation. J GEN INTERN MED 2020 Dec;35(12):3719-3720. [doi: 10.1007/s11606-020-05633-4] [Medline: 31925739]
32. Gajarawala SN, Pelkowski JN. Telehealth benefits and barriers. J Nurse Pract 2021 Feb;17(2):218-221. [doi: 10.1016/j.nurpra.2020.09.013] [Medline: 33106751]
33. Shachar C, Engel J, Elwyn G. Implications for telehealth in a postpandemic future. JAMA 2020 Jun 16;323(23):2375. [doi: 10.1001/jama.2020.7943]
34. Admin: COVID-19 telehealth coverage policies - CCHP. CCHP. 2021. URL: https://www.cchpca.org/resources/covid-19-telehealth-coverage-policies/ [accessed 2025-07-25]
35. New HHS study shows 63-fold increase in Medicare telehealth utilization during the pandemic. CMS. 2023. URL: https://www.cms.gov/newsroom/press-releases/new-hhs-study-shows-63-fold-increase-medicare-telehealth-utilization-during-pandemic [accessed 2025-07-25]
36. Abbott DE, Macke RA, Kurtz J, et al. Financial and temporal advantages of virtual consultation in veterans requiring specialty care. Mil Med 2018 Jan 1;183(1-2):e71-e76. [doi: 10.1093/milmed/usx006] [Medline: 29401334]
37. Beswick DM, Vashi A, Song Y, et al. Consultation via telemedicine and access to operative care for patients with head and neck cancer in a Veterans Health Administration population. Head Neck 2016 Jun;38(6):925-929. [doi: 10.1002/hed.24386]
38. Charlton M, Schlichting J, Chioreso C, et al. Challenges of rural cancer care in the United States. Cancer Network 2020 [FREE Full text]
39. Choxi H, VanDerSchaaf H, Li Y, Morgan E. Telehealth and the digital divide: identifying potential care gaps in video visit use. J Med Syst 2022 Jul 30;46(9):58. [doi: 10.1007/s10916-022-01843-x] [Medline: 35906432]
40. Rendle KA, Tan ASL, Spring B, et al. A framework for integrating telehealth equitably across the Cancer Care Continuum. JNCI Monographs 2024 Jun 26;2024(64):92-99. [doi: 10.1093/jncimonographs/lgae021]
41. Worster B, Waldman L, Garber G, et al. Increasing equitable access to telehealth oncology care in the COVID-19 National Emergency: Creation of a telehealth task force. Cancer Med 2023 Feb;12(3):2842-2849. [doi: 10.1002/cam4.5176] [Medline: 36210751]
42. Hart NH, Nekhlyudov L, Smith TJ, et al. Survivorship care for people affected by advanced or metastatic cancer: MASCC-ASCO Standards and Practice Recommendations. JCO Oncol Pract 2024 Sep;20(9):1160-1172. [doi: 10.1200/OP.23.00716] [Medline: 38684036]
43. Health Information National Trends Survey 6 (HINTS 6): HINTS 6 methodology report. National Cancer Institute. 2023. URL: https://hints.cancer.gov/docs/methodologyreports/HINTS_6_MethodologyReport.pdf [accessed 2025-07-25]

Abbreviations

FITE: Framework for Integrating Telehealth Equitably
HINTS 6: Health Information National Trends Survey 6
HIPAA: Health Insurance Portability and Accountability ACT
MASCC: Multinational Association for Supportive Care in Cancer

Edited by M Balcarras; submitted 08.Nov.2025; peer-reviewed by KI Alcaraz, WV Harten, X Ding; revised version received 25.Apr.2026; accepted 30.Apr.2026; published 03.Jul.2026.

Please cite as:
Kaur J, Zhong J, Mitchell C, Zhang A, Wang Q, Hsu ML
Predictors of Telehealth Use Among Cancer Survivors: Retrospective Study
JMIR Cancer 2026;12:e87403
URL: https://cancer.jmir.org/2026/1/e87403
doi:10.2196/87403

© Jasskiran Kaur, Jeffrey Zhong, Carley Mitchell, Annie Zhang, Qian Wang, Melinda L Hsu. Originally published in JMIR Cancer (https://cancer.jmir.org), 3.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cardiovascular Symptom Tracking Among Patients With Cancer in Cardio-Oncology Care: Qualitative Study Using the Capability, Opportunity, Motivation–Behavior (COM-B) Framework

Roberto M Benzo[1,2], PhD; Macy K Tetrick[1], BS; Rujul Singh[1,3], BS; Sara M St George[4,5], PhD; Vanina Pavia Aubry[5], MD, MS; Peter Washington[6], PhD; Anvitha Gogineni[7], BS; Sanam M Ghazi[8,9], MD; Olivia Preston[8], MS; Daniel Addison[8,9], MD

[1]Division of Cancer Prevention and Control, Department of Internal Medicine, College of Medicine, 3650 Olentangy River Rd, Columbus, OH, United States

[2]The Ohio State University Wexner Medical Center, The Ohio State University Comprehensive Cancer Center, Columbus, OH, United States

[3]Grossman School of Medicine, New York University, New York, NY, United States

[4]University of Miami Miller School of Medicine, Sylvester Comprehensive Cancer Center, Miami, FL, United States

[5]Department of Public Health Sciences, University of Miami Miller School of Medicine, Miami, FL, United States

[6]Division of Clinical Informatics and Digital Transformation (DoC-IT), Department of Medicine, University of California, San Francisco, San Francisco, CA, United States

[7]College of Medicine, The Ohio State University, Columbus, OH, United States

[8]Cardio-Oncology Program, Division of Cardiology, The Ohio State University Medical Center, Columbus, OH, United States

[9]Cardio-Oncology Program, Division of Cardiology, The University of Texas Southwestern Medical Center, Dallas, TX, United States

Corresponding Author:
Roberto M Benzo, PhD
Division of Cancer Prevention and Control, Department of Internal Medicine, College of Medicine, 3650 Olentangy River Rd, Columbus, OH, United States

Abstract

Background: Cardiovascular disease is a leading noncancer cause of morbidity and mortality among individuals diagnosed with cancer. Although modern cancer therapies have improved survival, many are associated with cardiotoxic effects that increase the risk of cardiovascular complications. Early identification of cardiovascular symptoms and signs may support timely clinical management. However, tracking cardiovascular health during cancer care can be complex for both patients and medical team participants, and important indicators, such as blood pressure changes, heart rate variability, fluid retention, chest pain, or new-onset fatigue, may be missed between clinic visits. Mobile health technologies offer potential tools to facilitate remote monitoring, yet behavioral factors influencing cardiovascular tracking in cardio-oncology remain insufficiently understood.

Objective: This study aimed to identify barriers and facilitators influencing cardiovascular symptom tracking among patients with cancer and medical team participants using the capability, opportunity, motivation–behavior (COM-B) framework.

Methods: This qualitative descriptive study included adult patients with cancer (n=12) receiving cardiotoxic therapies and members of their health care teams (n=12), including oncologists, nurses, advanced practice providers, and allied health professionals. Participants were recruited from a cardio-oncology clinic. Semistructured interviews explored perceptions of tracking cardiovascular symptoms and signs. Data were analyzed using rapid qualitative analysis and organized according to COM-B domains. The study received ethical approval, and participants provided informed consent.

Results: Patients with cancer and medical team participants described cardiovascular symptom tracking as a shared responsibility. Capability-related themes included uncertainty regarding which symptoms to monitor, appropriate thresholds for concern, and variability in communication practices. Opportunity-related themes included time pressures, competing clinical demands, and the influence of patient-medical team participant relationships and social support. Motivation-related themes reflected perceived benefits of tracking, including reassurance for patients and improved clinical insight for medical team participants, alongside concerns about information burden and unclear actionability. Both groups expressed interest in digital tools that provide clear guidance, defined response pathways, and integration within health care workflows.

Conclusions: Cardiovascular symptom tracking in cancer care is shaped by interrelated behavioral and contextual factors affecting both patients and medical team participants. Digital systems intended to support tracking should address knowledge clarity, workflow integration, and actionability of reported information. Understanding these determinants may inform the design and implementation of cardiovascular monitoring strategies in cardio-oncology settings.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e100279) doi:10.2196/100279

KEYWORDS

cardio-oncology; symptom monitoring; cancer survivorship; mobile health; digital health; qualitative research

Introduction

Cardiovascular disease (CVD) is a leading noncancer cause of mortality in populations with cancer and an important contributor to morbidity. A recent systematic review and meta-analysis of 136 studies across 16 countries reported that individuals with cancer experience a 55% higher risk of cardiovascular mortality compared with the general population (standardized mortality ratio=1.55, 95% CI 1.40 - 1.72), with increased risk observed for both heart disease and cerebrovascular conditions [1]. Similarly, a study in the United States shows that patients with cancer remain at increased risk of CVD-related death from the time of diagnosis forward into survivorship, where risk is particularly elevated during the first year following diagnosis [2]. Contemporary cardio-oncology reviews further emphasize that cardiovascular and oncologic conditions are biologically and clinically interconnected, underscoring the need for systematic attention to cardiovascular risk throughout the cancer continuum [3]. Collectively, these findings support the concern for CVD risk in survivorship rather than as a secondary complication in cancer care.

Beyond mortality differences, new cardiovascular events are common during cancer treatment. In a cohort of 17,389 Australian patients treated with chemotherapy, nearly half developed CVD requiring hospital admission during follow-up [4]. More than 50% of these events occurred within 12 months of starting treatment. Patients who developed CVD during follow-up had higher subsequent all-cause mortality [4]. In a population-based cohort study conducted in Alberta, Canada, more than 4.5 million adults were followed for a median of 11.8 (IQR 6.4-11.8) years [5]. Among 224,016 individuals with a new cancer diagnosis, adjusted analyses showed a higher risk of cardiovascular mortality, stroke, heart failure, and pulmonary embolism compared with participants without cancer, while myocardial infarction risk was not increased. Together, these findings suggest that cardiovascular complications frequently arise early in the cancer trajectory and are strongly linked to adverse outcomes, underscoring the need for proactive cardiovascular risk identification and monitoring during and after cancer therapy.

Despite the high incidence of cardiovascular events during and after cancer treatment, early cardiovascular symptoms, such as dyspnea, fatigue, palpitations, or peripheral edema, may emerge between scheduled oncology encounters and may be missed or underdetected [6]. Because these symptoms overlap with common cancer- or treatment-related effects, they may not always prompt cardiovascular evaluation. Systematic electronic monitoring of patient-reported outcomes (PROs) has been linked to positive outcomes in oncology settings. In the multicenter patient-reported outcomes to enhance cancer treatment (PRO-TECT) cluster-randomized trial conducted across 52 community oncology practices, weekly electronic symptom reporting among patients with metastatic cancer significantly reduced emergency department visits and delayed deterioration in physical function, symptom burden, and health-related quality of life (HRQoL) compared with usual care [7]. Similarly, implementation of electronic medical record (EMR)–integrated symptom and needs monitoring in a large ambulatory oncology population at an academic cancer center was associated with reduced emergency department usage and hospitalizations [8]. These findings highlight that systematic symptom tracking can enhance early detection and improve patient-centered and usage outcomes, yet cardiovascular-specific symptom monitoring has not been well-studied in or implemented in cardio-oncology practice.

Digital PRO monitoring and mobile health (mHealth) tools may help address gaps in cardiovascular risk and symptom tracking across the cancer continuum by extending monitoring beyond episodic clinic visits. In a systematic review of mHealth technologies used to study cardiovascular health in populations with cancer, the most commonly captured measures were physical activity and heart rate, primarily via mobile apps and commercial wearables, and integration with EMRs was limited, representing a significant implementation barrier to translating patient-generated data into clinical action [9]. Complementing this landscape, EMR-integrated symptom and needs monitoring in ambulatory oncology has demonstrated that workflow-embedded PRO systems are feasible and associated with reductions in emergency department visits and hospitalizations, while also revealing variability in patient engagement [8]. Qualitative work focused specifically on symptom-tracking for cardiotoxicity, demonstrating that both patients and medical team participants perceive value in a patient-facing mHealth application to facilitate earlier recognition of treatment-related cardiotoxicity. However, medical team participants emphasized the need for simple design, clear triage and escalation pathways for urgent symptoms, and integration with existing clinical workflows and EMR constraints to ensure a safe and feasible implementation [10]. Together, these findings suggest that although digital tools for symptom monitoring are increasingly available, cardiovascular-specific tracking in cardio-oncology remains underdeveloped and insufficiently integrated into routine practice.

Cardiovascular symptom tracking to support cancer care ultimately depends on the behaviors of both patients and medical team participants. In this context, relevant behaviors include patients actively monitoring and reporting cardiovascular signs and symptoms, and medical team participants promoting, reviewing, and responding to that information within routine care. The capability, opportunity, motivation–behavior (COM-B) model, which is central to the Behavior Change Wheel framework, conceptualizes behavior as arising from the interaction of capability, opportunity, and motivation [11]. Capability refers to an individual's ability to engage in a behavior and includes both physical skills and psychological

XSL•FO
RenderX

capacity (eg, knowledge and understanding). Opportunity refers to external factors that enable or constrain behavior, including physical opportunity (eg, time and resources) and social opportunity (eg, support and norms). Motivation encompasses processes that energize and direct behavior, including reflective processes (eg, intentions and beliefs) and automatic processes (eg, habits and emotional responses).

The framework was developed to identify determinants of behaviors and to inform intervention design and has been widely applied to organize barriers and facilitators in implementation research. Applying COM-B to cardiovascular symptom and sign tracking in the cardio-oncology setting provides a structured approach to understanding determinants of these interrelated patient and medical team participant behaviors. In this study, COM-B was used as an analytic framework to classify barriers and facilitators to cardiovascular symptom tracking.

Methods

Study Design

We used a descriptive qualitative design guided by the COM-B model. Qualitative data were collected through semistructured interviews focused on identifying barriers and facilitators to tracking symptoms and signs relevant to cardiovascular health (eg, fatigue, chest pain, palpitations, and blood pressure) in the context of cancer treatment and survivorship. Interviews also explored participants' perspectives on desirable features of digital health solutions to support cardiovascular symptom monitoring, with these features subsequently mapped to COM-B domains.

The target behavior in this study was tracking cardiovascular symptoms and signs between clinical appointments. This behavior involves complementary roles for patients with cancer and medical team participants. Patients engage in tracking using available tools and strategies and report this information during or between visits. The medical team often inquires about cardiovascular health during appointments and may ask patients to monitor specific signs or symptoms between visits to support clinical care discussions and decisions. Because tracking involves actions and expectations on both sides of the clinical relationship, both patients with cancer and medical team participants were interviewed.

The COM-B framework guided both the development of the interview guides and the analysis of qualitative data. The model conceptualizes behavior as the interaction of capability (physical and psychological), opportunity (physical and social), and motivation (reflective and automatic). In this study, the framework was used to classify participant-reported barriers and facilitators related to tracking behaviors.

Participants and Recruitment

Medical team participants (n=12) and patients with cancer (n=12) who were currently undergoing or had previously undergone cardiotoxic anticancer therapy as defined by established cardio-oncology guidelines [12] (eg, tyrosine kinase inhibitors, chemotherapy, immune checkpoint inhibitors, and radiation) participated in one-on-one semistructured interviews to discuss their experiences and perspectives related to tracking symptoms and signs relevant to cardiovascular health. Eligible medical team participants were (1) aged 18 years or older, (2) English-speaking, (3) currently practicing as a licensed health care professional or allied health professional involved in cancer or cardio-oncology care (eg, physician, advanced practice provider, nurse, or allied health discipline such as physical therapy, exercise physiology, or nutrition), (4) employed within an oncology or cardio-oncology care setting, and (5) involved in the care of patients who were receiving or had received cardiotoxic anticancer therapy.

Eligible patients were (1) aged 18 years or older; (2) English-speaking; (3) diagnosed with cancer; (4) currently undergoing or had previously undergone a cancer regimen that includes a targeted, biologic, radiation, or immune-based therapy (eg, tyrosine kinase inhibitors or immune checkpoint inhibitors); (5) able to engage in physical activity (self-report); and (6) interested in symptom or sign tracking or increasing physical activity levels to improve health (self-report). Although physical activity was discussed during interviews, this analysis focuses on tracking symptoms and signs relevant to cardiovascular health; physical activity–related findings are reported separately.

Medical team participants were recruited through word of mouth and professional referrals. Potential patients were recruited via medical team participant-based referrals within the cardio-oncology clinic. Medical team participants introduced the study to eligible patients during routine medical appointments. If patients expressed interest, members of the research team contacted them via telephone or email to provide additional study information and confirm eligibility. The telephone was the primary mode of contact unless the participant preferred email. If a potential participant could not be reached by phone, a voicemail was left. Research staff followed a 3-call limit protocol when attempting contact. Eligible participants who provided informed consent were enrolled and scheduled for a semistructured interview with a trained member or members of the research team.

Of the 32 patients with cancer screened through the EMR, 7 were ineligible. The remaining 25 preliminarily eligible patients were approached by the clinical study team during medical appointments in the cardio-oncology clinic and asked whether they were interested in speaking with the research study team. The research team attempted to contact all 25 patients by telephone or email. Of the total, 5 patients were unreachable despite repeated contact attempts, and 4 were not contacted after the recruitment target had been met. Moreover, 16 patients were successfully screened by the research team; of these, 2 were ineligible and 2 declined to participate. In total, 12 patients with cancer were enrolled and completed interviews.

Data Collection

Semistructured interviews were conducted remotely via videoconferencing (eg, Zoom [Zoom Communications] and Microsoft Teams) between December 12, 2023, and July 29, 2024. Interviews were facilitated by the lead author (RMB), one of the research team members (MKT), or both, who were trained in qualitative interviewing techniques and followed a semistructured guide designed to promote open discussion.

XSL•FO
RenderX

The interview guides were developed based on the COM-B model domains, including motivation, opportunity, and capability [11]. Open-ended questions were structured to elicit participants' perspectives related to tracking symptoms and signs relevant to cardiovascular health. Patients were asked about their experiences tracking their own symptoms and signs, whereas medical team participants were asked about their perspectives on patients' tracking behaviors. Topics included current tracking practices, contextual factors influencing tracking (eg, available tools and time constraints), motivational factors, and perceptions of potential digital tools to support tracking. Example prompts included: "How do you currently keep track of your own/your patients' signs and symptoms?" (capability), "Can you describe any opportunities or external factors that help you track signs and symptoms?" (opportunity), and "Can you describe your reasons for tracking signs and symptoms?" (motivation). The full interview guides are provided in Multimedia Appendix 1.

Interviews lasted approximately 60 - 90 minutes. All interviews were audio- and video-recorded with the participant's permission, transcribed using the auto-generated transcript from Zoom or Teams, and reviewed and corrected, when necessary, by a research team member (MKT). Participants were given the option to turn their camera on or off. Participants were assigned unique identifiers not associated with personal identifying information. Participants received an electronic US $25 gift card delivered via email upon completion of the interview.

Sociodemographic information from patients (age, sex, race, ethnicity, education, employment, income, height, and weight) was collected via REDCap surveys. Similar demographic information was collected verbally from medical team participants before interviews, including age, sex, race, ethnicity, job position, and years of experience in their current role.

Data Analysis

Demographic and medical characteristics were analyzed using descriptive statistics. Rapid qualitative analysis (RQA) was applied to analyze the one-on-one interviews [13-15]. RQA is a team-based qualitative approach in which research team members review audio recordings and transcripts, summarize key points, and identify patterns across interviews to generate themes.

A total of 5 study team members (RMB, MKT, SMSG, VPA, and AG) were involved in transcript review, data summarization, and were adequately trained in RQA. Research assistants first reviewed and corrected auto-generated transcripts for accuracy. Following this step, a researcher (RMB) reviewed each transcript and interview recording and identified segments corresponding to the COM-B domains and subdomains. Relevant excerpts were grouped according to these domains and further categorized as facilitators or barriers.

Each transcript was then independently reviewed by a second researcher (MKT), who repeated the process while having access to the initial coding for comparison. After double review, identified themes and supporting quotes were entered into a structured database to allow for comparison across participants. The research team met regularly to reconcile interpretations and ensure consistency.

Themes were finalized by examining patterns across participants and grouping similar concepts within COM-B subdomains. Because each subdomain is inherently aligned with a COM-B domain, themes were organized accordingly within capability, opportunity, or motivation and illustrated with representative quotes, with subcomponents (eg, physical or psychological capability; physical or social opportunity; and automatic or reflective motivation) used for classification.

Ethical Considerations

This study was approved by The Ohio State University Institutional Review Board (IRB #2021C0018). All participants provided informed consent before participation. Interviews were audio-recorded with participant permission. Identifying information was removed and replaced with unique identifiers to maintain confidentiality.

Results

Participant Characteristics

Participant demographic and clinical characteristics are summarized in Table 1.

XSL•FO
RenderX

Table . Demographic and clinical characteristics of sample population.

Variables	Patient with cancer (n=12), n (%)	Medical team participant (n=12), n (%)
Average age in years (SD)	65.3 (12.6)	37.8 (6.4)
Average time in position (years)	—[a]	4.5[b]
Sex		
Female	6 (50.0)	7 (58.3)
Male	6 (50.0)	5 (41.7)
Race or ethnicity		
White, non-Hispanic	10 (83.3)	6 (50.0)
White, Hispanic	—	1 (8.3)
Black or African American	1 (8.3)	1 (8.3)
Asian	—	3 (25.0)
Multiple races or ethnicities	1 (8.3)	1 (8.3)
BMI (kg/m^2)		
18.5 - 24.9	3 (25.0)	—
25 - 29.9	5 (41.7)	—
≥30	4 (33.3)	—
Average BMI	26.9[b]	—
Education completed		
High school diploma or GED	1 (8.3)	—
More than high school/some college	1 (8.3)	—
College degree	3 (25.0)	—
Graduate or professional degree	7 (58.3)	—
Employment		
Full-time	1 (8.3)	—
Part-time	0 (0)	—
Unemployed	1 (8.3)	—
Retired or disabled	10 (83.3)	—
Annual income (US $)		
50,000-74,999	4 (33.3)	—
75,000-99,999	1 (8.3)	—
100,000-199,999	5 (41.7)	—
Prefer not to answer	2 (16.7)	—
Job position		
Cardio-oncologists	—	4 (33.3)
Oncologists	—	3 (25.0)
Radiation oncologists	—	1 (8.3)
Nurse practitioner	—	1 (8.3)
Physical therapist	—	1 (8.3)
Exercise physiologist	—	1 (8.3)
Dietitian	—	1 (8.3)
Cancer site		
Acute myeloid leukemia	3 (25.0)	—

XSL•FO
RenderX

Variables	Patient with cancer (n=12), n (%)	Medical team participant (n=12), n (%)
Breast	2 (16.7)	—
Prostate	1 (8.3)	—
Lymphoma	1 (8.3)	—
Multiple myeloma	1 (8.3)	—
Uterine	1 (8.3)	—
Chronic lymphocytic leukemia	1 (8.3)	—
Lung	1 (8.3)	—
Neuroendocrine malignancy	1 (8.3)	—
Cancer treatment received		
Chemotherapy	6 (50.0)	—
Radiation	5 (41.7)	—
Stem cell transplant	4 (33.3)	—
Tyrosine kinase inhibitors	4 (33.3)	—
Hormonal therapy	3 (25.0)	—
Surgery	3 (25.0)	—
Immune checkpoint inhibitors	1 (8.3)	—
Other	4 (33.3)	—

[a]Not applicable.

[b]% value not applicable.

In total, 12 patients with cancer participated in interviews. The mean age was 65.3 (SD 12.6) years, and participants were evenly distributed by sex (6/12, 50% female). Most identified as White and non-Hispanic (10/12, 83.3%). The mean BMI was 26.9 kg/m² (SD 4.53), with 41.7% (5/12) classified as overweight and 33.3% (4/12) classified as obese. The majority had completed a graduate or professional degree (7/12, 58.3%) and were retired or disabled (10/12, 83.3%). The most frequently reported annual household income category was US $100,000-US $199,999, representing 41.7% (5/12) of participants. Cancer diagnoses included acute myeloid leukemia (3/12, 25%), breast cancer (2/12, 16.7%), and a range of other malignancies represented in smaller frequencies. Participants reported receiving multiple cancer-directed therapies, including chemotherapy (6/12, 50%), radiation (5/12, 41.7%), tyrosine kinase inhibitors (4/12, 33.3%), stem cell transplant (4/12, 33.3%), hormonal therapy (3/12, 25%), surgery (3/12, 25%), immune checkpoint inhibitors (1/12, 8.3%), and other therapies (4/12, 33.3%).

A total of 12 medical team members were interviewed. Most participants were female (7/12, 58.3%). The mean age was 37.8 (SD 6.4) years, and participants had been in their current position for a mean of 4.5 (SD 2.78) years. Participants identified as White and non-Hispanic (6/12, 50%), Asian (3/12, 25%), White and Hispanic (1/12, 8.3%), Black or African American (1/12, 8.3%), or multiple races or ethnicities (1/12, 8.3%). Professional roles included cardio-oncologists (4/12, 33.3%), oncologists (3/12, 25%), radiation oncologists (1/12, 8.3%), and allied health professionals, including a nurse practitioner, physical therapist, exercise physiologist, and dietitian (each 1/12, 8.3%).

Qualitative Themes

Themes were organized by COM-B domain and categorized as facilitators or barriers. Table 2 summarizes the themes identified among patients and medical team participants, along with the number of participants who endorsed each theme. Illustrative quotes can be found in Multimedia Appendix 2.

XSL•FO
RenderX

Table . Facilitators, barriers, and desired digital health features for cardiovascular symptom tracking, mapped to COM-B[a] domains.

Domain and subdomain	Theme	Patients, n	Medical team, n
Capability			
Facilitators			
Psychological	Awareness/knowledge of own symptoms/health	12	—[b]
Psychological	Was made aware of cardiotoxicity risks by medical team and asked to track	7	—
Physical	Journaling, using other recording tools to track	7	—
Psychological	Knowledge of cardiotoxicity	—	11
Physical	Medical team participant asks about symptoms	—	12
Barriers			
Psychological	Did not know about cardiotoxic risks or how, when to track	5	—
Psychological	Recall bias	5	—
Psychological	Lack of tech-savviness	4	—
Physical	Symptoms deterred them from tracking	4	—
Psychological	Hard to know which symptoms are cardiotoxic-related	—	8
Opportunity			
Facilitators			
Social	Patient–medical team participant encounters	11	7
Social	Social support	8	7
Physical	Physical and digital tools	11	10
Social	Staff support	—	5
Barriers			
Physical	Lack of time	6	7
Social	Medical team participant not requesting data	6	—
Physical	Lack of physical or digital tools	—	10
Social	Limited staff support	—	2
Motivation			
Facilitators			
Automatic	Habitual tracking	7	3
Reflective	Perceived value of monitoring for health management	11	—
Automatic	Emotional processes	6	—
Reflective	Planning their daily lives	2	—
Reflective	Clinical decision making	—	10
Reflective	Patient motivation and awareness	—	5
Reflective	Understanding health outside of the clinic	—	3
Barriers			

XSL•FO
RenderX

Domain and subdomain	Theme	Patients, n	Medical team, n
Reflective	Lack of emphasis to track from medical team participants	3	—
Reflective	Overwhelmed by demands of being a survivor	4	—
Automatic	Emotional processes	6	—
Reflective	Don’t want overload of information	—	4
Reflective	Patient motivation and burden	—	5
Digital health solution features			
Capability			
Simple design	—	6	5
Entering and tracking data	—	9	—
Integrate with other devices	—	3	—
Opportunity			
Reminders, prompts, widgets	—	7	—
Reminders, alerts, triaging features	—	—	7
Medical team participant communication	—	4	—
Educational insights	—	6	—
Motivation			
Tracking features	—	4	—
Tracking signs and visualizing trends	—	—	6
Reward system	—	1	4
Motivational messaging	—	3	—
Gamification	—	2	—
Sharing with medical team participant	—	2	—

[a]COM-B: capability, opportunity, motivation–behavior.

[b]Not applicable.

Within the capability domain, 3 facilitator themes and 4 barrier themes were identified among patients, compared with 2 facilitator themes and 1 barrier theme among medical team participants. Opportunity included 3 facilitator themes and 2 barrier themes among patients, and 4 facilitator themes and 3 barrier themes among medical team participants. Within the motivation domain, 4 facilitator themes and 3 barrier themes were identified among patients, while 4 facilitator themes and 2 barrier themes were identified among medical team participants.

Digital health solution features were also described and mapped to COM-B domains. Among patients, 5 motivation-related, 3 opportunity-related, and 3 capability-related features were identified. Among medical team participants, 2 motivation-related features and 1 feature related to opportunity and capability were identified.

Domain: Capability

Overview

In the COM-B framework, capability refers to an individual’s capacity to engage in a behavior. Capability includes psychological capability, such as knowledge, cognitive skills, and memory processes, and physical capability, such as the physical ability to perform the behavior. In the context of cardiovascular symptom tracking, capability reflects whether patients with cancer and medical team participants possess the knowledge, skills, and abilities necessary to monitor and interpret signs and symptoms. We identified 3 facilitator themes and 4 barrier themes among patients, and 2 facilitator themes and 1 barrier theme among medical team participants. Patients described varying levels of awareness of cardiotoxicity risks and challenges with real-time documentation, while medical team participants noted difficulty attributing symptoms specifically to cardiotoxicity.

XSL•FO
RenderX

Psychological Capability

The first theme related to knowledge of cardiovascular risk. Medical team participants described awareness that certain cancer treatments increase the risk of CVD. In contrast, although patients described awareness of their symptoms, just over half reported being aware that their cancer treatment could increase cardiovascular risk. Some patients reported that they had not been informed about the potential cardiac effects of cancer treatment. As some patients described:

> *They tested my heart before I was allowed to have the transplant, but I don't believe they said anything about the risk to my heart causing any heart problems.*
>
> *I never really thought to ask about well, how will this affect my heart, or will this affect my heart in a bad way, or whatever.*
>
> *No, we didn't discuss that [increased cardiovascular risk]. I don't remember. Well, I don't, I don't remember them discussing that, but I guess if I had looked at the papers that come with all the drugs that one drug that they took me off of definitely can impact the heart can be a side effect.*

Medical team participants described obtaining information about signs and symptoms during clinical encounters and noted difficulty determining whether symptoms were related to cardiotoxicity, given the variability in patient presentation. As one medical team participant stated:

> *They're not always related to cardiotoxicity, it could just be coincidental…but then, other times, it's clearly due to their cardiotoxicity of their cancer therapy.*

Most patients reported that their medical team had asked them to track signs and symptoms, and some described already monitoring select measures. However, several patients expressed uncertainty about which information to track and reported challenges documenting symptoms in real time. Some patients also reported relying on memory to recall symptoms between appointments rather than documenting them in real time. Only a subset of patients described using systematic tools, such as journals or applications, to record cardiovascular symptoms over time.

> *I don't know what else would be important to track of the, you know, what's being tracked. Blood pressure, all this other stuff.*
>
> *You know, [oncologist name] has been asking me, and once again this is my recall. I could not be remembering correctly. But he, you know, he mainly asked, "Do you have chest pain? Do you have this? Do you have that?" But I haven't been able to like, pull out a notebook and say, "On February third I had this and February second…"*

Finally, some patients noted that their lack of tech-savviness was a barrier to digital tracking. As one patient noted:

> *Oh, you gotta, you gotta keep in mind, I'm a bit technologically challenged.*

Physical Capability

Physical capability refers to the physical capacity to perform a behavior. In the context of tracking cardiovascular signs and symptoms, this includes whether patients feel physically able to consistently monitor and document symptoms in their daily lives.

Some patients described that severe symptoms, such as fatigue or feeling unwell, limited their ability to consistently track signs and symptoms. When symptoms were most burdensome, tracking was sometimes deprioritized or not completed at all.

Domain: Opportunity

Overview

In the COM-B framework, opportunity refers to external factors that make a behavior possible or prompt it. Opportunity includes both social (interpersonal influences, cultural norms, and professional roles) and physical (environmental resources, tools, time, and system-level structures) opportunities. In the context of cardiovascular symptom tracking, opportunity reflects the extent to which the clinical environment, social networks, and available tools support or constrain tracking behaviors for both patients with cancer and medical team participants. We identified 3 facilitator themes and 2 barrier themes among patients, and 4 facilitator themes and 3 barrier themes among medical team participants. Both groups highlighted patient–medical team participant encounters, social support, and tool access as facilitators, while time constraints and limited standardized resources emerged as key barriers.

Social Opportunity

The first theme identified by both patients and medical team participants was the role of patient–medical team encounters in shaping tracking behaviors. Most medical team participants described obtaining information about signs and symptoms during follow-up visits. Similarly, most patients reported that tracking was more likely to occur when their medical team explicitly asked them to monitor specific signs and symptoms between appointments. Conversely, some patients described that when tracking was not requested, they were less likely to engage in monitoring. For example, as 2 participants noted:

> *…during treatment for radiation for like esophagus cancer, for example, it's a five and a half week long treatment. So, we see them every week. They, you know, they kind of report what's going on every week.*
>
> *…I think my doctors definitely help, obviously. But they they've been awesome, and even in terms of explaining like, why, it's important. I think that's like a huge thing that sometimes providers miss, is explaining why specifically we're tracking these things and why it's important that we're kind of tracking them for ourselves, what they could in indicate in the future. And then they've also been very good at kind of explaining everything, and the process of how to do that.*

Another theme identified by both groups was the role of social support (ie, family, friends, and social support groups) in tracking signs and symptoms. Medical team participants

XSL•FO
RenderX

described that family members and friends could contribute additional information about patients' symptoms and help reinforce monitoring behaviors (eg, medications and blood pressure). For example, 1 medical team participant stated:

> *So, their family and friends are huge in getting the full picture of the patients a lot of times… So, getting all the information and, or, tracking like you said, tracking symptoms, making sure they're taking their medication, making sure they're taking their blood pressures, not downplaying their blood pressures if there's something that should be concerned.*

Patients also described family members and friends as providing support for tracking, including encouragement and concern when symptoms occurred. One patient stated:

> *My husband has been great cause he wants me to feel good. So, I think, having like a support system of people that want you to feel as good as you can and are concerned when you don't feel good and kind of want to mark that, that's important.*

A related theme among medical team participants involved staff roles in communication and follow-up. Participants described nurses as frequently serving as a point of contact between patients and oncologists. Some medical team participants also described workload and staffing constraints that affected their ability to follow up on patients' signs and symptoms outside of scheduled visits. For example, 1 medical team participant stated:

> *I don't have the bandwidth to call every patient and check on them and the patients I see here on Wednesday. [Nurse's name] helps with that.*

Physical Opportunity

Among both patients and medical team participants, most reported using physical and digital tools to track signs and symptoms. Most patients reported using paper-based tools (eg, notebooks, binders, or written logs). Patients also described using digital tools (eg, note-taking applications, meditation or glucose monitoring apps, blood pressure cuffs, and KardiaMobile) and device features (eg, alarms) to monitor signs and symptoms. A few patients reported not using any applications to track signs or symptoms. Some patients reported that having physical or digital tools organized in a single location helped with tracking. For example, 1 participant noted:

> *I found in life that it helps to be a little more organized. You know I mentioned. I try to always have the blood pressure cuffs in a one place, and it's taken me a long time in life to realize that being organized really helps.*

Most medical team participants described limited availability of paper-based or digital tools to provide to patients for tracking. Participants on the medical team reported lacking standardized resources to monitor patients between visits. For example, 1 medical team participant stated:

> *I think leveraging, like, 21st century technology to help patients, like, better identify and capture their symptoms, and then to also help them kind of engage in the preventative, or, you know, meaningful cardiac preservation or survivorship activities is really kind of a big part of the future.*

Another theme identified among both the medical team and patients was the time barrier. Most medical team participants consider time constraints to be a major barrier to understanding trends in patients' signs and symptoms. Similarly, many patients mentioned that finding time to track their signs and symptoms was a common obstacle.

> *So, I think cause sometimes, especially when I was working full time and also trying to like manage treatments and stuff, it just felt like there was such a limited amount of time and memory to do those things [tracking].*

Domain: Motivation

Overview

Motivation within the COM-B model refers to processes that direct and energize behavior. It includes reflective motivation, which involves conscious decision-making, planning, and evaluation of outcomes related to tracking behaviors, and automatic motivation, which involves emotional responses and habitual processes that influence engagement in tracking cardiovascular symptoms and signs. Themes were identified within both reflective and automatic motivation among patients and medical team participants. We identified 4 facilitator themes and 3 barrier themes among patients, and 4 facilitator themes and 2 barrier themes among medical team participants. Patients described reassurance and a sense of control as motivating tracking, while survivorship demands and emotional distress could diminish engagement; medical team participants valued tracking for clinical decision-making but expressed concern about information volume and actionability.

Reflective Motivation

Patients identified the perceived value of monitoring for health management and planning. Patients reported recognizing the importance of tracking their cardiovascular symptoms and described this recognition as motivating their engagement in tracking behaviors. Participants reported that monitoring contributed to a clearer understanding of their symptoms and allowed them to observe patterns or changes in symptoms over time. For example, 1 patient stated:

> *The one thing that I enjoy is being able to look at the data. I mean, that's just my thing. I like being able to look and see my trends where I've been, where I'm going.*

Another patient shared:

> *Honestly, I think that knowledge and data are kind of like a huge piece of mind for me personally like being to look able to look at trends especially like makes me think about chemo both with, you know, the signs of like my blood pressure, my heart rate, but also with the symptoms being able to kind of plan how I'm going to feel, I know on Day 5. I'm going to feel, you know, the most tired, and things like that. Allows me to kind of find the best way to live my life to the max, you know.*

XSL•FO
RenderX

However, some patients reported that when the medical team participants did not emphasize the importance of monitoring, they were less likely to prioritize tracking. Participants described reduced urgency or engagement when tracking was not explicitly discussed during clinical encounters. One patient noted:

> *I think that's like a huge thing that sometimes providers miss, is explaining why specifically we're tracking these things and why it's important that we're kind of tracking them for ourselves, what they could in indicate in the future.*

The second theme identified among patients was the use of tracking signs and symptoms to inform planning of daily activities and life events (eg, travel, holidays, and family events). Patients described using tracked information to guide their activities, adjusting plans based on how they felt or anticipated feeling. One patient described avoiding travel during treatment due to anticipated side effects and stated:

> *If I know when I'm going to be tired, I'm not going to make plans for those days. I'm going to try to give myself rest. But then, on the days where I know I start ticking up, then I know that I can start making plans with people, or we can choose to take a trip at that time.*

Some patients also described feeling overwhelmed by the demands of managing their health, which they reported affected their engagement in tracking. One participant shared:

> *I think it gets to be just a lot sometimes. Both logistically, in terms of like, there's so many appointments, there's so many medications. There are so many points, again, of like data, of like taking care of yourself, that sometimes you're just like, I want to sit on this couch all day. I don't want to think about any of this crap, like I need a day or 2 or 3, you know.*

The first theme identified among medical team participants was the use of signs and symptom data to inform clinical decision-making. Medical team participants described using PROs and other tracking data (eg, EMR data) to guide treatment adjustments and medication management. Participants also described tracking data as important for monitoring and identifying changes in cardiovascular health status. For example, 1 medical team participant stated:

> *But having better kind of updates on their symptom progression will allow us to better manage their cardiotoxicity or cardiovascular disease that may be independent from their cancer, cancer therapy.*

Medical team participants also described using symptom-tracking data to address specific patient complaints and inform clinical management decisions. Participants noted that tracking data provided additional context when evaluating symptoms. For example, 1 medical team participant stated:

> *You ask them whether they're doing okay, whether they're having any symptoms. If they're having any symptoms, then you try to make adjustments to the medications.*

In contrast, some medical team participants described information overload as a potential barrier to encouraging tracking. Participants noted concerns about receiving large volumes of patient-generated data and the time required to review it. For example, 1 medical team participant stated:

> *You like run the risk of collecting too much information, and, you know, kind of overburdening… You know, everybody's busy nowadays. And so, like sorting through you know, like reams of paperwork is not necessarily gonna be super helpful.*

The second theme identified among medical team participants was encouraging tracking to obtain information about patients' health outside the clinic, in their real-world environments. Participants described that clinical measurements may differ from those recorded at home and that tracking between visits provides additional context about patients' status. For example, 1 medical team participant stated:

> *Usually, I ask them, you know, "When you go home over the next, you know, month or so before I see you next, I want you to write down your blood pressure at different times of the day and make sure that you bring that back. So I know, you know, is your blood pressure really high at home, or is it just high in clinic because you're anxious?"*

Medical team participants described patient motivation as influencing engagement in tracking, noting variability in how consistently patients monitored and reported symptoms. Participants also described that asking patients to track signs and symptoms could add to patients' responsibilities, which, in some cases, limited engagement. At the same time, medical team participants reported encouraging patients to track their health between visits and emphasized the importance of patients being aware of changes in their health status. Some participants described prompting patients to take an active role in monitoring their health. For example, 1 medical team participant noted:

> *The big thing is for them to be self-aware of what's going on… So, being more self-sufficient and taking control of their health is the biggest goal, because, like so, we can't hold their hand forever.*

Automatic Motivation

Medical team participants described routinely asking about signs and symptoms during patient encounters, characterizing this practice as part of standard workflow. Similarly, several patients described tracking as part of their daily routine. Patients reported that engaging in routine tracking was reassuring and associated it with feeling more in control of their health. One patient shared:

> *It's just kind of a routine. I guess it just kind of reassures me that I know what's right. I just feel like it's probably, you know, a lot psychological, but it feels like I'm keeping track of something, my body, you know, and I've got control of something right?*

The second theme among patients revealed that emotional processes can influence the tracking of signs and symptoms. Patients mentioned concerns about their health and the potential

XSL•FO
RenderX

for worsening conditions, which may motivate them to monitor their symptoms closely, as they seek to anticipate changes and mitigate negative outcomes.

> *I probably subconsciously was more apt to worry about the cancer coming back. I probably worried more about any symptoms or anything and contributing to that.*

In contrast, some patients described that frequent or intense tracking became distressing or anxiety-producing. One participant shared:

> *Sometimes I see myself as going overboard with tracking so much that you become too somatically focused, and it starts becoming anxiety-producing. Because you're always tuned into every breath, and this and that, rather than living your life and having fun and doing things that are fun.*

Digital Health Solution Features

Both patients and medical team participants described specific features that could support cardiovascular symptom tracking.

Among medical team participants, the most frequently endorsed feature was reminders, alerts, or triaging capabilities (7/12). Medical team participants described the value of automated prompts to encourage patient logging, as well as alert systems that would highlight abnormal values and help prioritize review. One participant noted:

> *I will say reminders and things like keeping track of the blood pressure log, etc., through the phone, that works.*

Another medical team participant emphasized the importance of triaging information efficiently:

> *... like you only get the red flag if it's a lab that's abnormal or it's an imaging say, that's abnormal. So, in theory, when you're going through your inbox, you kind of triage, what's what I really need to like open up and look at versus it's probably normal.*

Of the 12 medical team participants, 6 highlighted the importance of tracking functionality that allows visualization of trends over time. Medical team participants described wanting both patients and medical team participants to see symptom progression and recovery patterns:

> *I've often thought like how nice it would be to, to, you know, kind of keep track of patients when they're not in my clinic to make sure that everything's going okay, and they're not in trouble, but then also to for the patient to kind of log things and kind of see their progress and see whether you know how they've progressed and how they're recovering. So, I think it'd be, you know, valuable in multiple aspects.*

Moreover, 4 medical team participants emphasized the potential benefit of incorporating reward systems or positive reinforcement to promote long-term tracking behaviors. One participant explained:

> *If they have some other commendation for whatever they're done in an app then they're more likely to do… and that's supposed to reinforce and something like that would definitely help people start doing or sticking onto the habit.*

Finally, 5 medical team participants stressed that any digital solution must be simple and user-friendly, with clear and actionable outputs:

> *I think if the app is simple, I think it could be helpful, that the app is simple and captures relative symptoms that are, that are useful, and that can be turned into actionable items that would be useful.*

Discussion

Principal Findings

This qualitative study examined barriers and facilitators to cardiovascular symptom and sign tracking from both patient and medical team perspectives, in a cardio-oncology setting, using the COM-B framework. We found that tracking behaviors related to cardiovascular health were shaped by interrelated factors across capability, opportunity, and motivation domains. Patients described variability in knowledge regarding cardiovascular health during and after cancer treatment, including uncertainty about which symptoms or signs to monitor and when to communicate changes. Medical team participants similarly identified gaps in standardized processes for cardiovascular symptom tracking and expressed concerns about time constraints and information burden. At the same time, both groups identified various motivational drivers, including perceived value of monitoring for maintaining cardiovascular health, reassurance, and improved clinical decision-making. Social opportunity, particularly patient–medical team participant communication and family involvement, also emerged as central to supporting tracking behaviors.

These findings suggest that cardiovascular symptom tracking in cancer care is not limited by awareness of cardiotoxicity alone but is influenced by clarity of expectations, workflow integration, and alignment between patient and medical team participant roles.

Comparison With Previous Literature

Randomized trials in oncology have demonstrated that structured electronic symptom monitoring can improve some clinical and PROs. For example, in the PRO-TECT cluster-randomized trial, weekly electronic symptom reporting reduced visits to the emergency department and delayed deterioration in physical function and HRQoL among patients with metastatic cancer [7]. Similarly, the use of an EMR-integrated symptom and needs monitoring in an ambulatory, racially, ethnically, and linguistically diverse oncology population (My Wellness Check [American Heart Association]) has been associated with reductions in emergency department usage and hospitalization [8]. These findings are consistent with broader evidence that PRO monitoring enhances symptom detection and clinical communication, supporting improved system-level outcomes in oncology care [16].

These studies establish the potential clinical impact of systematic symptom monitoring but do not examine the behavioral mechanisms influencing long-term engagement in symptom

XSL•FO
RenderX

tracking, particularly for cardiovascular risk in cardio-oncology contexts.

Our findings extend this literature by identifying capability-, opportunity-, and motivation-related determinants influencing the adoption and implementation of cardiovascular symptom tracking. Whereas previous trials primarily evaluate downstream clinical outcomes, our qualitative data illuminate upstream behavioral and contextual factors (including knowledge gaps, workflow constraints, uncertainty about symptom attribution, and perceived role boundaries) that may influence the feasibility of structured cardiovascular monitoring.

Our findings align with the qualitative work on cardiotoxicity-focused mHealth tools, where both patients and medical team participants expressed interest in digital cardiovascular symptom tracking but stressed the importance of clear reporting guidance for urgent symptoms, integration with existing clinical systems, and minimal workflow disruption [10]. Additionally, a recent systematic review of mHealth tools for cardiovascular health in patients with cancer found a fragmented array of technologies and measures, limited integration with electronic health records (EHRs), and few studies specifically addressing structured cardiovascular symptom collection [9]. More broadly, mHealth technologies enable continuous monitoring of both physiological and patient-reported data, but their implementation is often constrained by variability in engagement and limited integration into clinical care workflows [9], aligning with recent implementation-focused reviews demonstrating that the effectiveness of remote symptom monitoring depends on patient engagement, workflow integration, and actionable use of reported data [16].

Together, these studies suggest that although digital monitoring tools are increasingly available, implementation may depend less on technological capability alone and more on alignment with behavioral determinants and system-level structures. This aligns with findings from previous work on EHR-integrated interventions, which highlight that successful implementation depends on alignment with clinical workflows, defined roles, and manageable data flows rather than technology alone [17,18]. By applying the COM-B framework, our study provides a behavioral lens to inform the design and integration of cardiovascular symptom-tracking approaches in oncology settings. Notably, cardiovascular symptom tracking may differ from general oncology symptom monitoring in several respects: the overlap between cardiotoxicity-related and cancer- or treatment-related symptoms, the evolving nature of cardiovascular risk across the cancer care continuum, and the need for coordination across oncology and cardiovascular specialties.

Implications for Behavioral Determinants and Implementation

Cardiovascular symptom tracking in cancer care emerged as a shared behavior shaped by capability, opportunity, and motivation among both patients and medical team participants. Patients were often unsure which cardiovascular symptoms or signs should be tracked and when changes should prompt communication, suggesting that education alone may be insufficient without clear expectations for monitoring and reporting. Although patients generally recognized the relevance of cardiovascular health, many described ambiguity about expectations for monitoring and reporting. Medical team participants similarly described variability in how cardiovascular monitoring was addressed and incorporated into routine care. Such variability likely reflects differences in clinical context, patient complexity, and workflow demands, underscoring the importance of clarifying expectations and responsibilities when considering strategies to support consistent tracking practices. These challenges highlight the importance of clearly defining which cardiovascular symptoms should be tracked and when reporting is warranted, particularly in contexts where symptom attribution is uncertain. They also suggest that effective tracking approaches may need to be integrated into existing workflows and support coordination across oncology and cardiovascular care teams, agreeing with broader systematic reviews identifying knowledge gaps, workflow constraints, and variability in perceived value as central barriers to PRO implementation [19].

Opportunity constraints were prominent. Time limitations, competing priorities, and concerns about information burden shape whether tracking can be integrated into routine workflows. Previous implementation research in oncology has shown that monitoring systems are more likely to be adopted when embedded within existing clinical processes, supported by defined triage pathways, and aligned with established roles rather than added as an additional task [7-9]. EMR-integrated monitoring programs that include automated alerts and predefined response workflows have been described as feasible in ambulatory oncology settings when these structural supports are in place [20]. Observational analyses have also reported associations between participation in such programs and health care usage outcomes [8]. More broadly, implementation science literature emphasizes that long-term adoption of digital monitoring systems depends on minimizing disruption to clinical workflows and reducing cognitive and time burden on medical team participants [21], along with recent umbrella reviews highlighting workflow disruption, limited integration with EHRs, and time and resource constraints as key barriers to implementing patient-generated health data systems in oncology care [19]. The findings of this study add to this literature by identifying behavioral factors that may influence uptake of cardiovascular-specific tracking within cardio-oncology settings.

Motivational influences were equally apparent. Patients described reassurance and a sense of control as reasons for engaging in monitoring, whereas uncertainty about symptom meaning sometimes contributed to inconsistent reporting. Medical team participants were more inclined to engage with monitoring data when it was perceived as clinically meaningful but expressed concern when data volume was high or when response expectations were unclear. This aligns with previous real-world implementation studies demonstrating that long-term engagement with electronic PRO systems may vary over time and is influenced by reminder systems, medical team participant reinforcement, perceived relevance to care, and integration into clinical workflows [22]. These findings are also consistent with previous research showing that perceived usefulness and ease of use are central determinants of engagement with digital health

XSL•FO
RenderX

and self-monitoring tools in chronic disease management [23]. These perspectives indicate that perceived usefulness and manageable data flow may be important for long-term engagement over time.

Taken together, these findings indicate that knowledge clarity, workflow integration, and perceived value operate in combination. Introducing a digital tool without addressing expectations or integrating tracking into existing care processes may offer limited benefit.

Implications for Digital Health Design

Our findings inform the design of digital tools intended to support cardiovascular symptom tracking in cancer care. For patients, tools may benefit from providing clear guidance on which symptoms and signs to monitor and when changes should prompt communication with the medical team. Educational components embedded within the tracking process may help address gaps in psychological capability while avoiding additional burden during clinical encounters. Tools should also account for the overlap between cardiovascular, cancer-related, and treatment-related symptoms to avoid excessive data collection that increases patient burden or generates information that is not clinically actionable.

Medical team participants emphasized that digital tracking must be integrated into existing EMR systems and supported by clear triage processes. Previous implementation work in oncology indicates that feasibility is strengthened when monitoring systems include defined response workflows and designated roles for follow-up [20]. Broader implementation-oriented scholarship in digital oncology similarly emphasizes the importance of integrating patient-generated health data into existing clinical infrastructure and aligning digital tools with established workflows [21].

Broader mHealth research in populations with cancer has documented limited integration of patient-generated data into clinical systems and substantial variability in measures captured across platforms [9], echoing findings from recent reviews that emphasize the need for integrating patient-generated health data into clinical workflows and ensuring data are actionable at the point of care [19]. Participants emphasized that tracking is useful only when incoming information can be interpreted and acted upon efficiently. However, actionability may depend not only on data integration, but also on the availability of clinical resources, referral pathways, and support services that allow medical teams to respond to identified needs. For example, symptom or risk data may be more actionable in settings where patients can be referred to embedded supportive care, rehabilitation, behavioral health, nutrition, or cardiovascular services. In contrast, when follow-up services are unavailable, poorly coordinated, or associated with additional patient costs, tracking may generate information that is difficult to act on and may increase the burden for patients and medical team participants. This reflects broader challenges in translating patient-generated data into actionable clinical insights, particularly when data streams are not standardized or integrated with existing decision-support systems [9]. From an informatics perspective, this underscores the importance of systems that support interpretation within clinical workflows rather than treating data as isolated data entries [21].

In summary, digital cardiovascular symptom tracking systems should prioritize clarity, workflow alignment, and actionable outputs rather than simply increasing data volume. Participants' perspectives also indicate that minimizing cognitive burden and time constraints for both patients and medical team participants may be central to feasibility. Clear expectations regarding symptom relevance and defined response processes may support more consistent engagement over time.

Strengths and Limitations

This study has several strengths. First, it included perspectives from both patients and medical team participants, allowing examination of cardiovascular symptom tracking as an interdependent behavior rather than a unidirectional patient responsibility. Second, use of the COM-B framework provided a structured approach to organizing barriers and facilitators across psychological, social, and environmental domains while maintaining analytic transparency. Third, interviews focused specifically on cardiovascular symptoms and signs tracking in the cardio-oncology context, contributing to a growing body of research that has largely examined symptom monitoring more broadly in oncology.

Several limitations should be considered. Participants were recruited from a single, specialized cardio-oncology clinic and were predominantly highly educated and White, which may limit transferability to other institutions with different workflow structures or to more diverse and underserved populations. These findings may not reflect the experiences of individuals with lower health or digital literacy, limited access to technology, or those receiving care outside of specialized clinic settings. As with all qualitative studies, findings reflect self-reported perceptions and experiences rather than directly observed behaviors. We did not measure actual tracking adherence or clinical outcomes and therefore cannot draw conclusions regarding the effectiveness of specific tracking approaches. Additionally, because all interviews were conducted with participants who had access to videoconferencing capability, they may have been more interested in cardiovascular health or digital tools than nonparticipants, potentially introducing selection bias. Finally, while the COM-B framework provided a useful organizing structure, alternative theoretical approaches may have highlighted different aspects of tracking behavior. Future studies should include more diverse and community-based populations to better understand how cardiovascular symptom tracking can be implemented across different health care settings.

Conclusions

Cardiovascular symptom tracking in cancer care is a shared behavior requiring coordinated engagement between patients and medical team participants. The use of the COM-B framework allowed identification of capability, opportunity, and motivation factors that shape how cardiovascular symptoms and signs are monitored and communicated in cardio-oncology settings. Participants described ambiguity regarding expectations across both groups. Patients emphasized uncertainty about which

XSL•FO
RenderX

symptoms warranted tracking and about expectations for communicating symptom information, whereas medical team participants described time pressures, concerns about information burden, and limited standardized processes for reviewing and responding to symptom information. Medical team participants expressed concern about receiving large volumes of symptom information without clear mechanisms for prioritization or response, while patients described reassurance and a greater sense of control as benefits of tracking. Efforts to strengthen cardiovascular symptom tracking may benefit from addressing behavioral determinants alongside technological considerations. Digital systems that provide clear guidance, align with clinical workflows, and support actionable interpretation of symptom reports may facilitate engagement and perceived value. Future studies should examine how these behavioral insights inform implementation strategies and evaluation of cardiovascular monitoring approaches in routine cancer care.

Acknowledgments

This study was supported by discretionary research funds provided to Dr. Roberto M. Benzo through the Ohio State University Comprehensive Cancer Center (OSUCCC). We also thank the participants who shared their time and insights. The authors are grateful to the Lifestyles Lab team members for their contributions to data management and administrative coordination.

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) as a writing support tool to proofread for grammar. All content was written, critically reviewed, and edited by the authors, who take full responsibility for the final manuscript.

Funding

This study was supported by discretionary research funds provided to RMB through the Ohio State University Comprehensive Cancer Center (OSUCCC).

Data Availability

The statistical code used to analyze the data in this study is available from the corresponding author upon reasonable request. The datasets generated or analyzed during this study are not publicly available as it contains confidential patient information but can be made available from the corresponding author on reasonable request.

Authors' Contributions

RMB conceived and designed the study, oversaw data collection and data analysis, conducted the primary interpretation of findings, and led manuscript writing. RMB and MKT conducted interviews. MKT, RS, SMSG, and VPA coordinated data management, performed preliminary analyses, and assisted with manuscript writing. All authors contributed to study design, interpretation of results, and critical revision of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient and medical team interview guides.
[PDF File, 957 KB - cancer_v12i1e100279_app1.pdf]

Multimedia Appendix 2
Illustrative quotes for each theme from patients and medical team members.
[DOCX File, 41 KB - cancer_v12i1e100279_app2.docx]

References

1. Ng HS, Meng R, Marin TS, et al. Cardiovascular mortality in people with cancer compared to the general population: a systematic review and meta-analysis. Cancer Med 2024 Aug;13(15):e70057. [doi: 10.1002/cam4.70057] [Medline: 39096123]
2. Sturgeon KM, Deng L, Bluethmann SM, et al. A population-based study of cardiovascular disease mortality risk in US cancer patients. Eur Heart J 2019 Dec 21;40(48):3889-3897. [doi: 10.1093/eurheartj/ehz766] [Medline: 31761945]
3. Wilcox NS, Amit U, Reibel JB, Berlin E, Howell K, Ky B. Cardiovascular disease and cancer: shared risk factors and mechanisms. Nat Rev Cardiol 2024 Sep;21(9):617-631. [doi: 10.1038/s41569-024-01017-x] [Medline: 38600368]
4. Liu S, Horowitz JD, Koczwara B, Sverdlov AL, Packer N, Clark RA. Cardiac events among a cohort of 17,389 patients receiving cancer chemotherapy: short and long term implications. Cardiooncology 2024 Oct 16;10(1):72. [doi: 10.1186/s40959-024-00269-3] [Medline: 39415228]
5. Paterson DI, Wiebe N, Cheung WY, et al. Incident cardiovascular disease among adults with cancer: a population-based cohort study. JACC CardioOncol 2022 Mar;4(1):85-94. [doi: 10.1016/j.jaccao.2022.01.100] [Medline: 35492824]

XSL•FO
RenderX

6. Bonsu JM, Guha A, Charles L, et al. Reporting of cardiovascular events in clinical trials supporting FDA approval of contemporary cancer therapies. J Am Coll Cardiol 2020 Feb 18;75(6):620-628. [doi: 10.1016/j.jacc.2019.11.059] [Medline: 32057377]
7. Basch E, Schrag D, Jansen J, et al. Symptom monitoring with electronic patient-reported outcomes during cancer treatment: final results of the PRO-TECT cluster-randomized trial. Nat Med 2025 Apr;31(4):1225-1232. [doi: 10.1038/s41591-025-03507-y] [Medline: 39920394]
8. Natori A, Sookdeo VD, Koru-Sengul T, et al. Symptoms and Needs Monitoring in Diverse Ambulatory Oncology Patients: Usage Characteristics and Impact on Emergency Room Visits and Hospitalization: JCO Wolters Kluwer; 2023, Vol. 41:285-294. [doi: 10.1200/JCO.22.01038]
9. Benzo RM, Gogineni A, Tetrick MK, et al. mHealth technologies in research studying cardiovascular health in cancer: a systematic review. PLOS Digit Health 2025 Sep;4(9):e0001027. [doi: 10.1371/journal.pdig.0001027] [Medline: 40996972]
10. Gregory ME, Cao W, Rahurkar S, et al. Exploring the incorporation of a novel cardiotoxicity mobile health app into care of patients with cancer: qualitative study of patient and provider perspectives. JMIR Cancer 2023 Dec 12;9:e46481. [doi: 10.2196/46481] [Medline: 38085565]
11. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation Sci 2011 Dec;6(1). [doi: 10.1186/1748-5908-6-42]
12. Lyon AR, López-Fernández T, Couch LS, et al. 2022 ESC Guidelines on cardio-oncology developed in collaboration with the European Hematology Association (EHA), the European Society for Therapeutic Radiology and Oncology (ESTRO) and the International Cardio-Oncology Society (IC-OS). Eur Heart J 2022 Nov 1;43(41):4229-4361. [doi: 10.1093/eurheartj/ehac244] [Medline: 36017568]
13. Hamilton A. Qualitative methods in rapid turn-around health services research. Better evaluation knowledge. URL: https://www.betterevaluation.org/tools-resources/qualitative-methods-rapid-turn-around-health-services-research [accessed 2026-07-16]
14. Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. Psychiatry Res 2019 Oct;280:112516. [doi: 10.1016/j.psychres.2019.112516] [Medline: 31437661]
15. St George SM, Harkness AR, Rodriguez-Diaz CE, Weinstein ER, Pavia V, Hamilton AB. Applying rapid qualitative analysis for health equity: lessons learned using "EARS" with Latino communities. Int J Qual Methods 2023;22:16094069231164938. [doi: 10.1177/16094069231164938] [Medline: 38463016]
16. Aanes SG, Wiig S, Nieder C, Haukland EC. Implementing digital patient-reported outcomes in routine cancer care: barriers and facilitators. ESMO Real World Data Digit Oncol 2024 Dec;6:100088. [doi: 10.1016/j.esmorw.2024.100088] [Medline: 41646091]
17. Cresswell KM, Bates DW, Sheikh A. Ten key considerations for the successful implementation and adoption of large-scale health information technology. J Am Med Inform Assoc 2013 Jun;20(e1):e9-e13. [doi: 10.1136/amiajnl-2013-001684] [Medline: 23599226]
18. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
19. Rammant E, Ramsey I, M de Ligt K, et al. Barriers and facilitators for implementing patient-reported outcome measures in oncology practices: an umbrella review. Qual Life Res 2026 Apr 10;35(5):129. [doi: 10.1007/s11136-026-04205-z] [Medline: 41961151]
20. Penedo FJ, Medina HN, Moreno PI, et al. Implementation and feasibility of an electronic health record-integrated patient-reported outcomes symptom and needs monitoring pilot in ambulatory oncology. JCO Oncol Pract 2022 Jul;18(7):e1100-e1113. [doi: 10.1200/OP.21.00706] [Medline: 35290096]
21. Benzo RM, Noriega Esquives BS, Penedo FJ. APA handbook of health psychology, volume 2: clinical interventions and disease management in health psychology (vol. 2). In: Schneiderman N, Antoni MH, Penedo FJ, editors. APA Handbook of Health Psychology, Volume 2: Clinical Interventions and Disease Management in Health Psychology: American Psychological Association. [doi: 10.1037/0000395-027]
22. Patt D, Wilfong L, Hudson KE, et al. Implementation of electronic patient-reported outcomes for symptom monitoring in a large multisite community oncology practice: dancing the Texas two-step through a pandemic. JCO Clin Cancer Inform 2021 May;5(5):615-621. [doi: 10.1200/CCI.21.00063] [Medline: 34085537]
23. Tao D, Wang T, Wang T, Zhang T, Zhang X, Qu X. A systematic review and meta-analysis of user acceptance of consumer-oriented health information technologies. Comput Human Behav 2020 Mar;104:106147. [doi: 10.1016/j.chb.2019.09.023]

Abbreviations

COM-B: capability, opportunity, motivation–behavior
CVD: cardiovascular disease
EHR: electronic health record

XSL•FO
RenderX

EMR: electronic medical record
HRQoL: health-related quality of life
mHealth: mobile health
PRO: patient-reported outcome
PRO-TECT: patient-reported outcomes to enhance cancer treatment
RQA: rapid qualitative analysis

Edited by M Balcarras; submitted 04.May.2026; peer-reviewed by Y Zhu; revised version received 22.Jun.2026; accepted 29.Jun.2026; published 23.Jul.2026.

Please cite as:
Benzo RM, Tetrick MK, Singh R, St George SM, Pavia Aubry V, Washington P, Gogineni A, Ghazi SM, Preston O, Addison D Cardiovascular Symptom Tracking Among Patients With Cancer in Cardio-Oncology Care: Qualitative Study Using the Capability, Opportunity, Motivation–Behavior (COM-B) Framework
JMIR Cancer 2026;12:e100279
URL: https://cancer.jmir.org/2026/1/e100279
doi:10.2196/100279

© Roberto M Benzo, Macy K Tetrick, Rujul Singh, Sara M St George, Vanina Pavia Aubry, Peter Washington, Anvitha Gogineni, Sanam M Ghazi, Olivia Preston, Daniel Addison. Originally published in JMIR Cancer (https://cancer.jmir.org), 23.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Associated With the Use of the Lifestyle and Empowerment Techniques in Survivorship of Gynecologic Oncology (LETSGO) mHealth App in Routine Follow-Up After Gynecologic Cancer Treatment: Observational Study

Zaklina Tarabar[1,2], MSc; Elin Børøsund[3,4], PhD; Sindre H Fosstveit[5,6], PhD; Sveinung Berntsen[5,6], PhD; Milada Hagen[7], PhD; Ingvild Vistad[2,8], MD, PhD

[1]Hospital of Southern Norway, Egsveien 100, Kristiansand, Agder, Norway

[2]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

[3]Department of Digital Health Research, Division of Medicine, Oslo University Hospital, Oslo, Norway

[4]Department of Nursing and Health Sciences, Faculty of Health and Social Sciences, University of South-Eastern Norway, Drammen, Norway

[5]Department of Sport Science and Physical Education, University of Agder, Kristiansand, Agder, Norway

[6]Research Unit, Sorlandet Hospital, Kristiansand, Norway

[7]Department of Nursing and Health Promotion, Oslo Metropolitan University, Oslo, Norway

[8]Department of Obstetrics and Gynecology, Sorlandet Hospital, Kristiansand, Agder, Norway

Corresponding Author:
Zaklina Tarabar, MSc
Hospital of Southern Norway, Egsveien 100, Kristiansand, Agder, Norway

Abstract

Background: Mobile health (mHealth) apps offer new opportunities to support cancer survivors in managing their health, adopting healthier lifestyles, and increasing awareness of recurrence symptoms. However, despite their potential, little is known about survivors' use of such tools in routine follow-up care, or which factors are associated with their engagement.

Objective: The aim of the present study was to identify survivor characteristics associated with the use of the Lifestyle and Empowerment Techniques in Survivorship of Gynecologic Oncology (LETSGO) app. In addition, we aimed to describe engagement patterns among gynecologic cancer survivors during the first year of follow-up after treatment.

Methods: App data from 378 participants in the intervention group of the LETSGO multicenter clinical trial were included in the analysis. Using multiple logistic regression, we analyzed app data from the first year of LETSGO app use to identify participant characteristics associated with use. In addition, we described the frequency of user interactions with app features and identified the most frequently used features during the first year of the intervention in a subset of 49 participants with complete app log data.

Results: Of the 378 participants in the intervention group, 267 (70.6%) used the LETSGO app at least twice and were defined as app users. The median age of app users was 63 (53-70) years vs 70 (60-78) years among app nonusers (P<.001). App use was associated with younger age (odds ratio [OR] 0.96 per year, 95% CI 0.93 - 0.99; P=.003), higher education (OR 2.2, 95% CI 1.1 - 4.6; P=.03), and having ovarian cancer (OR 2.9; 95% CI 1.1 - 7.5; P=.03). The LETSGO app included a monthly reminder for symptom self-registration, and monthly peaks in the symptom monitoring feature corresponded with the timing of this reminder. App log data from the full first year of the intervention were available for 49 participants. Among these, user engagement was highest for the physical activity and activity goal-setting features, whereas the disease information feature showed steady but less frequent use.

Conclusions: Our findings indicate that LETSGO app use was associated with survivor characteristics such as age, education level, and cancer type. Considering user characteristics when tailoring mHealth apps may support user engagement in digital follow-up care and inform the development of mHealth tools that better align with survivors' needs and preferences. Engagement with the LETSGO app during the first year of follow-up among a subset of participants with complete app log data was highest for the physical activity and activity goal-setting features. The pattern of symptom registration following reminders indicates that scheduled prompts may support regular app engagement, while steady use of the disease information feature suggests an ongoing need for accessible health information during follow-up.

Trial Registration: ClinicalTrials.gov NCT04122235; https://clinicaltrials.gov/study/NCT04122235

XSL•FO
RenderX

(JMIR Cancer 2026;12:e89918) doi:10.2196/89918

KEYWORDS

app use; cancer survivors; digital follow-up care; digital health tools; mobile health apps; user engagement.

Introduction

Gynecologic cancer and its treatment can have a profound impact on survivors' well-being, with many women experiencing physical and psychosocial late effects [1-3]. As life expectancy following cancer treatment continues to improve, the focus of care has shifted from survival alone to promoting long-term well-being and quality of life [4-7]. Consequently, follow-up care has become an important component of survivorship, not only for monitoring recurrence but also for managing late effects and providing psychological support [8-10]. At the same time, the growing population of cancer survivors places increasing pressure on the outpatient clinics, where most patients are monitored for years after treatment. This underscores the need for more sustainable follow-up strategies that meet survivors' needs and are resource-efficient [10-12].

In this context, empowering cancer survivors to take an active role in managing their own health is increasingly recognized as a key component of future follow-up care. Such empowerment may help reduce the burden on outpatient clinics by enhancing survivors' ability to cope with cancer-related challenges and regain control over everyday life [13]. A central aspect of empowerment is strengthening the self-management skills of cancer survivors to promote self-care, address ongoing needs after treatment, and contribute to more efficient use of health care services [13-16].

Digital interventions, particularly mobile health (mHealth) apps, have shown promise in promoting empowerment and improving quality of life among cancer survivors [17-20]. The mHealth apps are internet-based apps on mobile devices that support health-related activities. Over recent years, an increasing number of mHealth apps have been developed to meet the specific needs of patients with cancer [21-23]. These mHealth technologies can complement health care professionals by providing tailored information, enabling self-monitoring, simplifying access to health care services, supporting behavior and lifestyle changes, and thereby reducing long-term health care costs [24,25]. They can also serve as platforms for collecting patient-reported outcomes, offering valuable insights into symptoms, functional status, and overall well-being without requiring in-person visits [26,27].

The effectiveness of mHealth apps in cancer survivorship for supporting behavior and lifestyle changes, as well as symptom monitoring, depends not only on their design and content but also on regular and sustained user engagement. Research indicates that digital health interventions, including mHealth apps, are effective only when users interact with them as intended [28,29]. Nonadherence, when users fail to use the digital intervention as intended, and attrition, when they discontinue use over time, remain common challenges [30,31]. Therefore, understanding user engagement and user perspectives is essential to improve the effectiveness of mHealth apps [30,32]. Although mHealth apps are increasingly used in cancer survivorship, knowledge regarding how users interact with these technologies over time and which factors influence sustained use remains limited [29]. As mHealth apps usually include more than one feature, it is important to identify which features are most used by cancer survivors because their needs may vary depending on diagnosis and individual characteristics [33].

In the Lifestyle and Empowerment Techniques in Survivorship of Gynecologic Oncology (LETSGO) multicenter clinical trial, we evaluated a new follow-up model involving nurse-led patient education, reinforced with the LETSGO app [16]. The comprehensive LETSGO app was designed to educate patients on late effects after cancer treatment, motivate physical activity, encourage healthier lifestyle changes, facilitate symptom monitoring for recurrence detection, and provide reminders to support physical activity goals in gynecologic cancer survivorship [16].

To improve our understanding of user engagement and to support the design and development of future digital follow-up interventions that better meet survivors' needs, this study's primary aim was to examine associations between selected demographic and clinical characteristics and the use of the LETSGO app among gynecologic cancer survivors in follow-up care. A secondary aim was to describe 1-year use patterns, including frequency of use and the most frequently used features.

Methods

Study Setting and Participants

The LETSGO multicenter clinical trial recruited women (N=741) from 12 Norwegian hospitals between December 2019 and February 2023 to test a new follow-up model for routine survivorship care after treatment for gynecological cancer, aimed at enhancing health-related empowerment through education and self-management strategies, as described in the study protocol [34]. Eligible women were aged 18 years or older, had a histologically confirmed diagnosis of ovarian, endometrial, cervical, or vulvar or vaginal cancer, were able to understand Norwegian, and had no cognitive impairments. Patients were assigned to their respective group (intervention group or standard follow‐up group) based on their hospital of residence. At the intervention hospitals, women received alternating nurse-led and physician-led follow-up consultations. The nurses focused on coaching participants in empowerment techniques, reinforced by the multifunctional LETSGO app, which was introduced to the women at the initial nurse-led consultation [34]. This substudy was conducted among 378 (51%) participants in the intervention group and is based on app registration data and app log data from the first year of LETSGO app use.

XSL•FO
RenderX

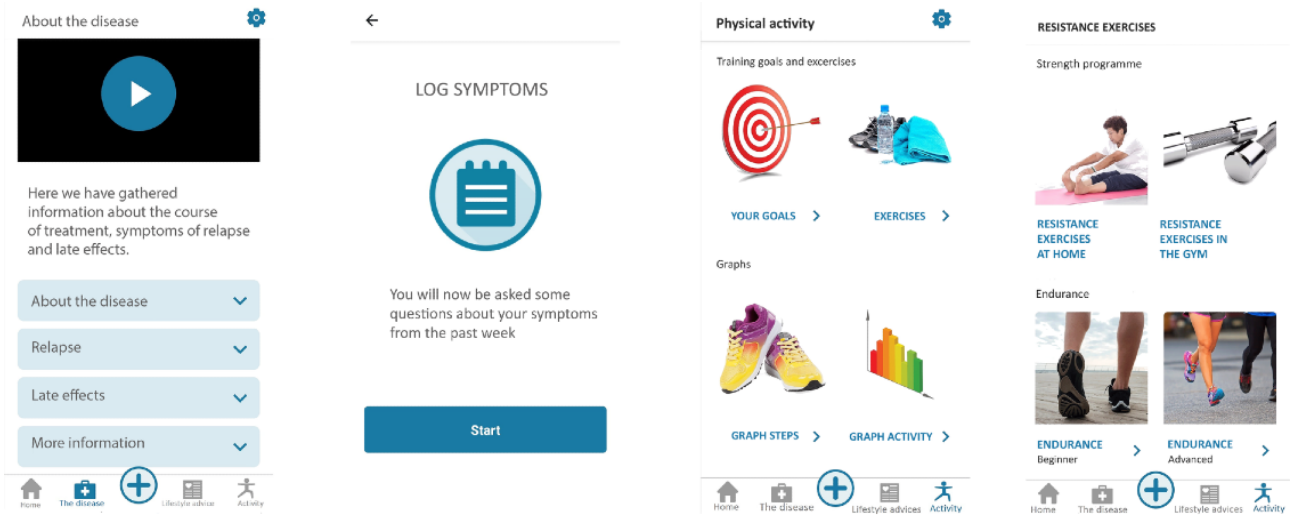
The LETSGO App

The LETSGO app was developed by a multidisciplinary team of experts in the field of gynecological cancer, behavior change, exercise oncology, and IT, alongside user representatives [34]. Its primary goal was to empower women to take a more active role in their own health by providing education and motivation. Pilot testing of the LETSGO intervention and study procedures provided valuable feedback on the app's functionality [35]. In response, technical improvements were made prior to its use in the intervention study. At the first nurse-led consultation, participants received a personal login code for the LETSGO app along with use instructions. Participants could also discuss app-related questions during follow-up consultations with nurses, including receiving help with technical or practical issues related to app use. Each participant registered their cancer type in the app to access diagnosis-specific information and saved the phone number of their outpatient clinic with direct access to the responsible nurse. The LETSGO app, accessible on smartphones and tablets, featured multiple modules, written content, links to external resources, and videos for additional guidance, as shown in Table 1 and Figure 1. Furthermore, the app offered users the option to log their own physical activities and set activity goals. Additionally, they received a wearable activity tracker to monitor their physical activity level (walking steps) [34].

Table . Content of the Lifestyle and Empowerment Techniques in Survivorship of Gynecologic Oncology (LETSGO) app.(Figure 2)

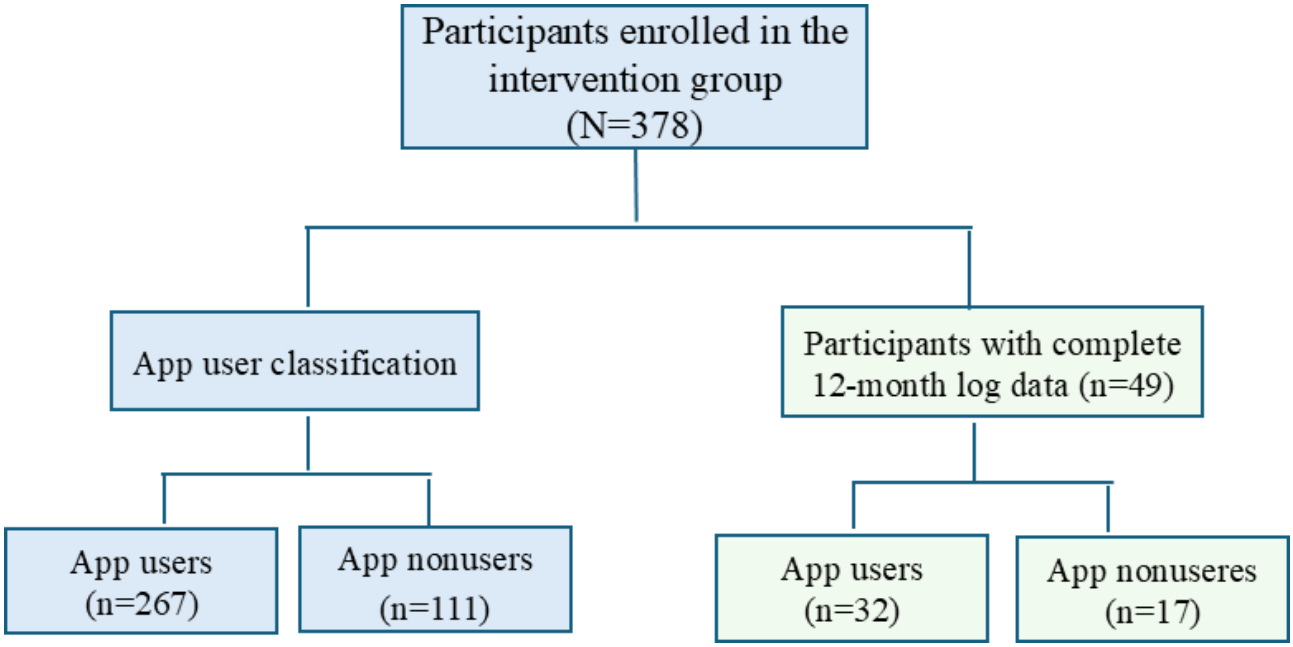
Features	Written content and videos
Contact information	• Telephone number to the responsible nurse
Disease information (ovarian, endometrial, cervical, or vulvar or vaginal cancer)	• Treatment • Signs of recurrence • Late effects • Links to external information
Lifestyle information and coaching	• Dietary guidance • Smoking cessation • Stress management • Yoga and mindfulness instructions
Physical activity	• Physical activity instructions • Resistance exercises • Endurance • Activity goal setting
Electronic reminder	• Electronic reminder to register physical activity • Electronic reminder to register symptoms
Symptom monitoring for recurrence	• 10 items related to cancer recurrence

Figure 1. Screenshots illustrating the core features of the Lifestyle and Empowerment Techniques in Survivorship of Gynecologic Oncology (LETSGO) app (Multimedia Appendices 1-4).



XSL•FO
RenderX

Figure 2. App user classification and availability of complete 12-month app log data.



The monthly monitoring of self-evaluated symptoms potentially indicating a recurrence included 10 items. These items were selected from the European Organisation for Research and Treatment of Cancer (EORTC) item library based on the most frequently reported recurrence symptoms for each type of gynecologic cancer [36]. The app sent users monthly reminders to complete the symptom monitoring questionnaire while also allowing them to complete the same questionnaire whenever needed. If their responses exceeded a predetermined threshold, an alert was triggered in the app, indicating a potential recurrence and advising them to contact the outpatient clinic. The phone number of the nurse at their outpatient clinic was automatically displayed on the screen for easy access.

Data Collection

At baseline in the LETSGO trial, participants completed patient-reported outcome measures, including sociodemographic characteristics (ie, age, BMI, education level, and employment status). Clinical information (ie, diagnosis and treatment modality) was obtained from medical records. These data were stored in the Viedoc software (Viedoc Technologies AB), a secure electronic data capture system, and extracted for inclusion in the analysis dataset. In this study, participants' education level was categorized as primary (≤10 years), secondary (11 - 13 years), or medium or high (≥14 years), and employment status was categorized as 0=no or 1=yes. Cancer diagnosis was categorized as endometrial, ovarian, cervical, or vulvar cancer. Treatment modality was categorized into three groups: (1) surgery only, (2) surgery plus chemotherapy, and (3) chemotherapy plus radiotherapy (with or without surgery) and other treatment combinations. Because of the low number of participants receiving chemotherapy and radiotherapy and other treatment combinations, these were combined into a single category.

All app data included in this study (app registration data and app log data) from the first year of the intervention were retrieved from the Services for Sensitive Data (TSD) platform at the University of Oslo, a secure system for handling sensitive research data. The app data included participant registrations in the physical activity, activity goal setting, and symptom monitoring features.

In addition, app log data capturing use frequency, measured as the number of clicks, were available for the physical activity, activity goal setting, disease information, and lifestyle information and coaching features. However, due to a miscommunication between the research team and the app developers regarding data storage, app log data were not stored from the beginning of the intervention period. Once this was identified, app log data collection was implemented. Consequently, app log data covering the full 12-month intervention period were available for only 49 participants. The analysis of app-use frequency over 12 months was limited to this subset of 49 participants.

App users were defined as those participants who, during the first year of the intervention, registered symptoms (symptom monitoring) at least twice and reported weekly physical activity (physical activity) and/or activity goals (activity goal setting) at least twice. There is currently no consensus on how engagement thresholds in digital health interventions should be operationalized, and our definition was pragmatically chosen to distinguish participants with repeated use of both features from those who did not meet this threshold, while also avoiding unnecessary exclusion of participants with low-frequency engagement [28,30].

Statistical Analyses

All statistical analyses were performed using SPSS Statistics (version 30; IBM Corp).

Participant characteristics were described using descriptive statistics. Continuous variables were presented as medians and ranges, while categorical variables were presented as counts and percentages. Group differences were assessed using the

XSL•FO
RenderX

Mann-Whitney *U* test for continuous variables and the Pearson chi-square test for pairs of categorical variables.

Multiple logistic regression analyses were conducted to identify participant characteristics associated with app use and to examine which groups were more likely to use the app. App use (coded as 0=no and 1=yes) was the dependent variable. Variables considered clinically relevant and potentially associated with app use, including demographic (eg, age, education level, and employment status) as well as clinical characteristics (ie, diagnosis and treatment modality), were entered as possible predictive factors. These independent variables were entered into the analysis simultaneously and treated as covariates, allowing the association between each variable and app use to be estimated while controlling for the others. Multicollinearity between independent variables was assessed, and no evidence of problematic multicollinearity was identified. Odds ratios (ORs) with 95% CIs were calculated to assess the strength of association.

Weekly symptom registrations over the first year of the intervention for all participants (N=378) were summarized to describe overall engagement with the symptom monitoring feature.

A subanalysis of the 49 participants with complete app log data covering the full 12-month intervention period was conducted to assess the frequency of user interactions with the key app features (physical activity, activity goal setting, disease information, and lifestyle information and coaching), and to identify which app features were most frequently used.

The app log data captured the number of user clicks within each feature. Results are presented as bar and line graphs. Differences in demographic and clinical characteristics between participants with complete app log data covering the full 12-month intervention period and those without were assessed using the Mann-Whitney *U* test for continuous variables and the Pearson chi-square test for categorical variables. All tests were 2-sided, and $P<.05$ was considered statistically significant. All results are considered exploratory; therefore, no correction for multiple testing was applied.

Ethical Considerations

The LETSGO study (ClinicalTrials identifier NCT04122235) was approved by the Regional Committee for Medical and Health Research Ethics (REK 2019/11093) and the data protection officers at participating hospitals. Written informed consent covering all aspects of the LETSGO study was obtained from the participants prior to inclusion. The data from the LETSGO app were stored on the TSD platform at the University of Oslo, and sociodemographic and clinical data were stored in Viedoc software.

Results

Characteristics of the Participants

Of the 378 participants included in the analysis, 267 (70.6%) were defined as app users. An overview of participant classification and availability of complete 12-month app log data is presented in Figure 2. Some baseline variables were missing for 39 participants. There were no statistically significant differences between those with and without missing data regarding age ($P=.23$), treatment modality ($P=.38$), and type of diagnosis ($P=.26$). Of those with missing data, 33.3% (13/39) were app users. The median age of all participants was 65 years, and app users were significantly younger than app nonusers (median age 63 vs 70 years). App users had significantly higher education and were more often employed compared with app nonusers (Table 2). Endometrial cancer was the most common diagnosis (221/378, 58.5%).

XSL•FO
RenderX

Table . Baseline characteristics of the participants (N=378).

	Total group (n=378)	App users[a] (n=267)	App nonusers[b] (n=111)	*P* value
Age (years), median (IQR)	65 (54-72)	63 (53-70)	70 (60-78)	<.001
BMI (kg/m^2), median (IQR)	27 (24-32)	27 (24-32)	28 (24-32)	.76
Missing, n (%)	41 (10.8)	14 (5.2)	27 (24.3)	—[c]
Diagnosis, n (%)				.18
Endometrial cancer	221 (58.5)	151 (56.6)	70 (63.1)	
Ovarian cancer	92 (24.3)	73 (27.3)	19 (17.1)	
Cervical cancer	58 (15.3)	39 (14.6)	19 (17.1)	
Vulvar cancer	7 (1.9)	4 (1.5)	3 (2.7)	
Treatment modality, n (%)				.29
Surgery only	213 (56.3)	145 (54.3)	68 (61.3)	
Surgery plus chemotherapy	121 (32)	92 (34.5)	29 (26.1)	
Chemotherapy plus radiotherapy and other combinations[d]	44 (11.6)	30 (11.2)	14 (12.6)	
Education, n (%)				<.001
Primary	50 (13.2)	28 (10.5)	22 (19.8)	
Secondary	156 (41.3)	124 (46.4)	32 (28.8)	
Medium or high	133 (35.2)	102 (38.2)	31 (27.9)	
Missing	39 (10.3)	13 (4.9)	26 (23.4)	
Employment, n (%)				<.001
Yes	120 (31.7)	103 (38.6)	17 (15.3)	
No	219 (57.9)	151 (56.6)	68 (61.3)	
Missing	39 (10.3)	13 (4.9)	26 (23.4)	

[a]App users: ≥2 registrations of symptoms and ≥2 registrations of weekly physical activity or activity goals during the first year of follow-up.

[b]App nonusers: Participants who did not meet the definition of app users.

[c]Not applicable.

[d]Participants receiving chemotherapy plus radiotherapy and other treatment combinations were grouped due to low numbers.

Factors Associated With App Use

Participants with missing data on variables included in the logistic regression model were excluded from the analysis, resulting in 337 complete cases included in the adjusted model.

Women with vulvar cancer were excluded from the regression analysis due to the low number of participants in this diagnosis group.

In the multiple logistic regression analysis, adjusted for age, education level, employment status, type of diagnosis, and treatment modality, increasing age was associated with lower odds of app use. For each additional year of age, the odds for app use were approximately 4% lower (OR 0.96 per year, 95% CI 0.93 - 0.99; *P*=.003). Younger participants had higher odds of app use, with a 10-year decrease in age being associated with approximately 50% higher odds of using the app. Participants with secondary education had more than 2-fold higher odds of app use compared with those with primary education (OR 2.23, 95% CI 1.09 - 4.56; *P*=.03). Participants diagnosed with ovarian cancer had higher odds of app use compared with those diagnosed with cervical cancer (OR 2.93, 95% CI 1.14 - 7.53; *P*=.03). Similarly, participants diagnosed with endometrial cancer had higher odds of app use compared with those diagnosed with cervical cancer (OR 3.31, 95% CI 1.16 - 9.49; *P*=.03). Other variables, including treatment modality and employment status, were not associated with app use (Table 3).

XSL•FO
RenderX

Table . Factors associated with app use based on multiple logistic regression analysis.

Variables	OR (95% CI)	*P* value
Age	0.96 (0.93 - 0.99)	.003
Education		
Primary	Reference	—[a]
Secondary	2.23 (1.09 - 4.56)	.03
Medium or high	1.48 (0.69 - 3.18)	.31
Employment status		
Yes	1.70 (0.81 - 3.56)	.16
Cancer type		
Cervical	Reference	—
Ovarian	2.93 (1.14 - 7.53)	.03
Endometrial	3.31 (1.16 - 9.49)	.03
Treatment modality		
Surgery only	Reference	—
Surgery plus chemotherapy	1.56 (0.79 - 3.12)	.20
Chemotherapy plus radiotherapy and other combinations	1.43 (0.53 - 3.83)	.48

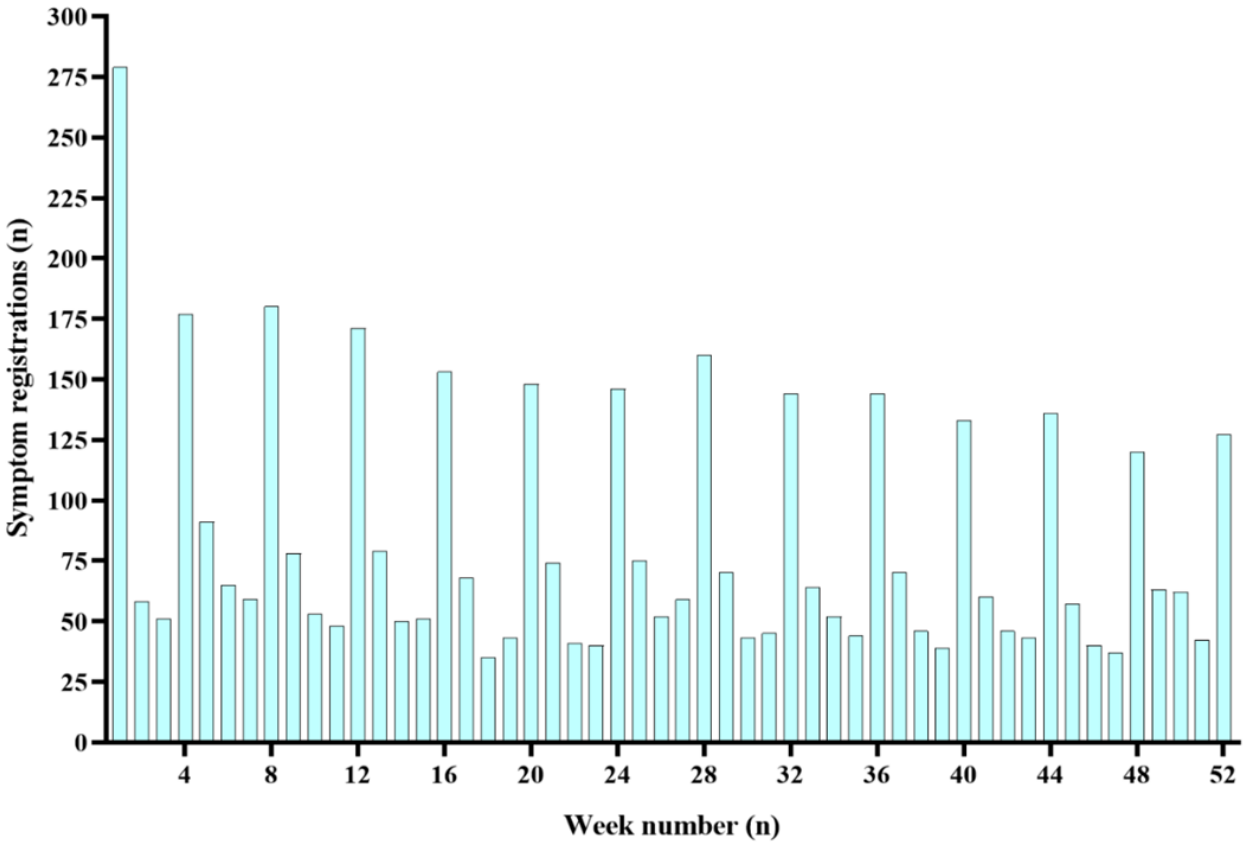
[a]Not applicable.

Registration in the Symptom Monitoring Feature

Of the 378 participants, 310 (82%) registered symptoms using the symptom monitoring feature at least once, including participants who did not meet the predefined criteria for app use. Symptom registration was most frequent in the first weeks after inclusion, followed by a decline and stabilization over the 1-year period (Figure 3). We observed monthly peaks in registrations that corresponded to the app reminder being sent once per month. However, participants also registered symptoms between reminders, indicating engagement with the feature outside the scheduled prompts.

XSL•FO
RenderX

Figure 3. Weekly number of symptom registrations over the 52-week follow-up period.



App Feature Use Over 12 Months for 49 Participants With Complete Log Data

A subanalysis of frequencies of app use among 49 participants over the first year of intervention showed that the proportion of users who interacted with each feature at least once per month was the highest for physical activity (n=39, 79.6%), followed by activity goal setting (n=29, 59.2%), while disease information and lifestyle information and coaching showed lower but relatively consistent engagement (n=10-20, 20.4% - 40.8%) over time (Figure 4). The median number of uses per month, defined as the number of individual participant registrations within app features, also indicated that physical activity and activity goal setting features were used most frequently (Figure 5). Among the 49 participants with complete log data, the median of app-use frequency was 104 (IQR 54 - 301; range 16 - 10,767). There were no statistically significant differences between participants with app log data covering the full intervention period (n=49) and those without (n=329) regarding age, BMI, education level, employment status, diagnosis, and treatment modality (data not shown).

XSL•FO
RenderX

Figure 4. Proportion of users engaging with each app feature at least once per month during the 12-month follow-up period (n=49).

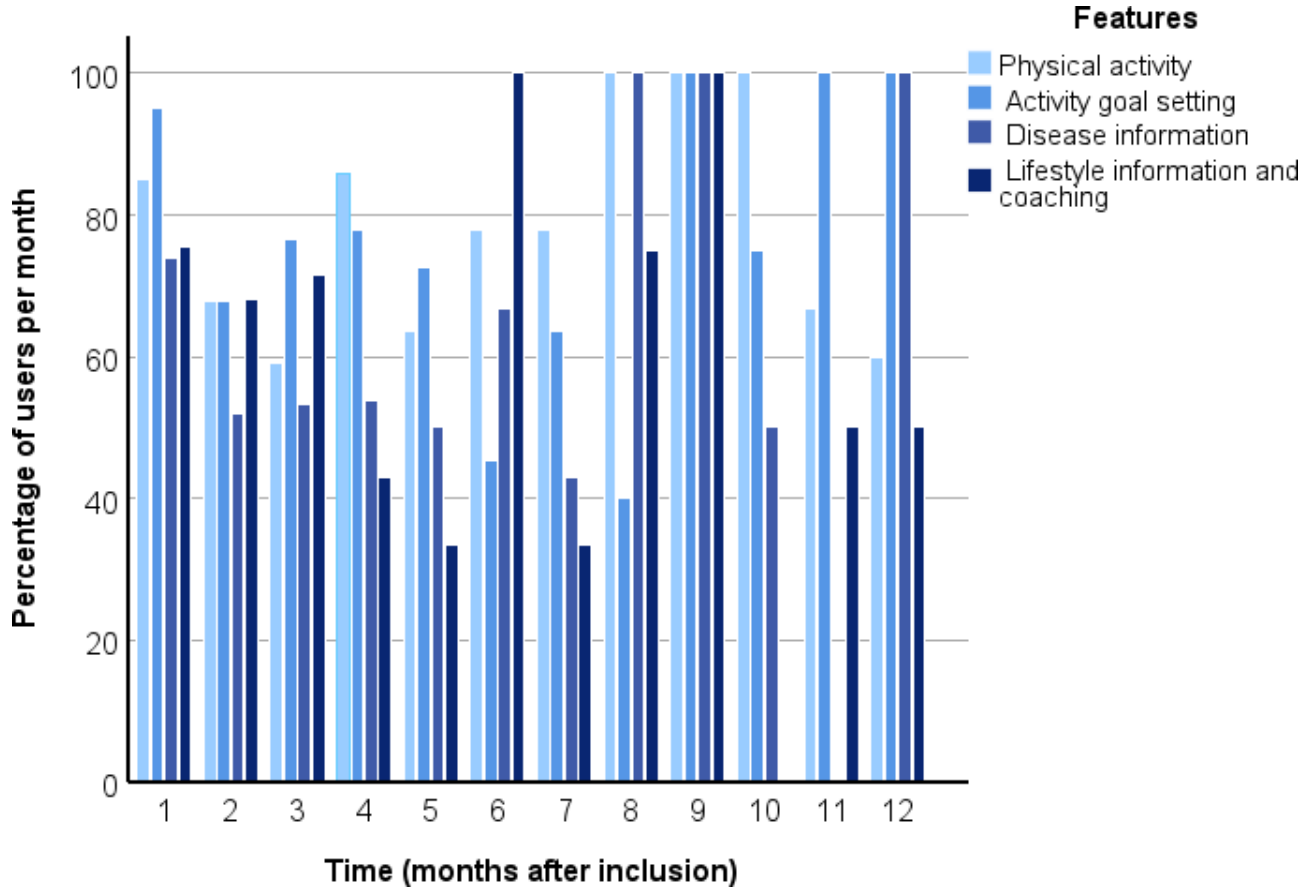
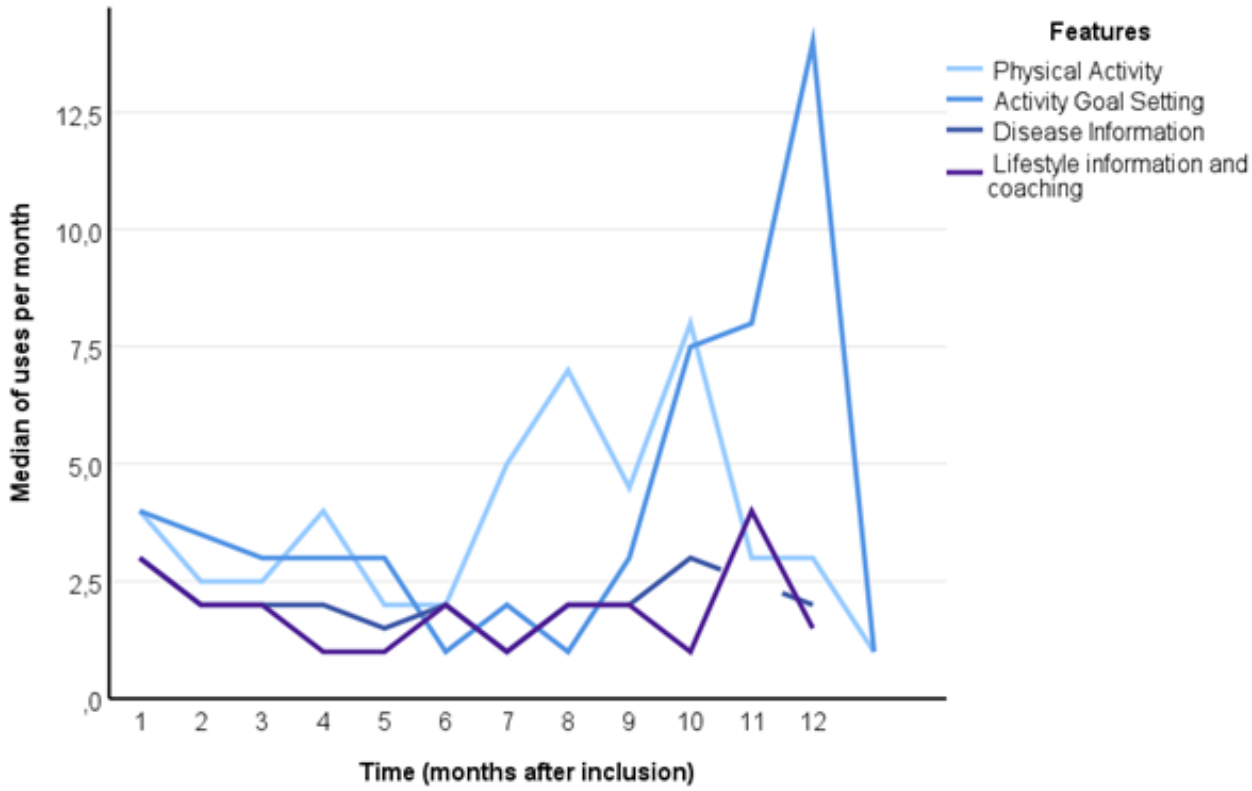


Figure 5. Median monthly use frequency of each app feature during the 12-month follow-up period (n=49).



XSL•FO
RenderX

Together, Figures 2 to 4 illustrate different aspects of app engagement during the first year of the intervention, including symptom registrations over time, the proportion of users engaging with different app features, and monthly use frequency.

Discussion

Principal Findings

In this observational study of 378 gynecologic cancer survivors participating in the LETSGO trial who had access to the LETSGO app as part of the intervention, app use was more common among younger women, those with secondary education, and women diagnosed with ovarian cancer. Symptom registration was highest shortly after inclusion and stabilized over time with monthly peaks corresponding to the app's reminder schedule. In a subanalysis of 49 app users with complete log data, engagement was highest for physical activity and activity goal setting features, while use of the disease information feature remained steady but less frequent.

The inverse association between age and app use aligns with previous studies showing that age may influence engagement with digital technologies, including mHealth technologies [37-40]. An umbrella review of 46 systematic and scoping reviews found that most digital interventions were effective regardless of age, whereas results on actual use and engagement showed greater variation [41]. Our findings align with this, indicating that although digital interventions may be effective across age groups, their uptake and engagement still differ by age. This may reflect differences in digital skills, familiarity with technology, or digital health literacy [42]. Older individuals may experience barriers related to digital health interventions, including lower digital competence, reduced confidence in using technology, lower digital health literacy, and less encouragement from health care providers [43,44].

Previous research [41] has highlighted that education has often been overlooked in existing research, even though it may contribute to the digital gap. Educational level may reflect both digital skills and general health awareness, which are important factors for engaging with digital health interventions. Earlier studies have reported higher app use among individuals with higher (tertiary) education [45,46]. In our study, however, app use was higher among women with secondary education compared with those with primary education. Among women with higher education, there was also a tendency toward higher app use compared with those with primary education, although this difference was not statistically significant. A possible explanation for the lack of a stronger association could be that women with higher education already have greater access to digital health technologies in general, which could reduce their need for this specific app [47]. However, this hypothesis was not directly examined in our study. This disparity underlines the importance of examining how education levels influence both use and sustained engagement with digital interventions in future research.

We found that women diagnosed with ovarian and endometrial cancer had higher odds of app use compared with women with cervical cancer. This finding may reflect differences in the disease trajectory, treatment intensity, and follow-up needs across diagnostic groups. One possible explanation for the observed association among women with ovarian cancer, suggested by previous research, is that ovarian cancer is often associated with greater symptom burden, longer treatment trajectories, and a perceived need for closer follow-up [48,49], which could make digital tools for symptom tracking and support more relevant and useful in their care routines [50].

Regular symptom monitoring may indicate that participants found the feature useful, which is supported by findings from a study reporting higher engagement when patients perceived symptom monitoring as relevant and meaningful for managing their symptoms [51]. In our study, the monthly peaks following reminders suggest that scheduled prompts support sustained engagement among participants. Some symptom registrations occurred outside the immediate monthly reminder pattern. However, the available data do not allow us to determine whether these registrations were self-initiated or delayed responses to reminders.

This finding aligns with previous research showing that structured symptom monitoring can encourage sustained app use over time [52]. Nevertheless, the distinction between reminder-driven and self-initiated use may be important to consider when designing future digital follow-up apps. Balancing structured guidance, such as reminders, with the opportunity for users to engage when they find it relevant could help maintain user engagement over time while ensuring that users feel supported without becoming dependent on reminders.

Although the impact of patient-directed alerts in the LETSGO app was not examined in our study, such alerts may still play a role in shaping how patients engage with and respond to the symptom monitoring feature. In the LETSGO app, patient-directed alerts are generated when symptom reports exceed predefined threshold values and prompt patients to take action, such as contacting a health care provider. Unlike many mHealth apps that notify clinicians, alerts in the LETSGO app are one-way and directed only to patients. Previous studies have shown that patient-directed symptom alerts can support earlier action and enhance patient self-management [26]. Sending alerts to patients rather than clinicians may also reduce the burden on health care providers [26]. This design supports patient empowerment approaches that encourage patients to take a more active role in their own health, which is one of the goals of mHealth interventions [32].

The high engagement with physical activity–related features in the subanalysis of 49 app users may indicate that physical activity and goal setting are particularly important for some participants in this subgroup. However, because this analysis was based on a small sample, these use patterns should be interpreted with caution and not generalized to the entire intervention group. Within these limitations, our findings are consistent with previous work highlighting the central role of physical activity features in digital interventions for cancer survivors [53,54], and they may reflect that such features foster feelings of progress, help users regain a sense of control, and support self-management during recovery [55]. This is further supported by a study [56] that found that physical activity

XSL•FO
RenderX

tracking and goal-oriented features were important drivers of motivation and engagement in physical activity in a mHealth intervention for breast cancer survivors.

The steady use of the disease information feature indicates an ongoing need for accessible and reliable information that supports participants' understanding of their condition. This observation should be interpreted in light of previous research showing that digital health literacy is often limited and influenced by factors such as age, gender, and education level [57]. This underscores the need to develop cancer-related information that is easily accessible and user-friendly across digital health platforms, including mHealth apps.

Study Strengths and Limitations

Strengths of this study include a large sample size (N=378), which allowed for robust analysis and increased the precision of the estimated associations between app use and participant characteristics. Because Norway ensures equal access to health care and consistent cancer treatment nationwide, our unselected, multihospital sample strengthens the external validity of the findings. Finally, app use was measured based on app log data rather than self-reports, thereby reducing possible recall bias.

However, the study has several limitations. The criteria for app use were defined by the minimum user threshold and therefore reflected minimal active use rather than sustained long-term engagement. Definitions of engagement in digital health research vary and are measured in different ways, with no clear standard for how engagement thresholds should be operationalized [28,30]. The results of this study reflect only 1 year of access to the LETSGO app, and engagement patterns over a longer period may differ. Although analyses were adjusted for demographic and clinical variables, unmeasured factors such as digital literacy or attitudes toward technology may have influenced app use. Missing data in some baseline variables may have affected the findings and introduced potential bias, as the logistic regression analyses were based on complete cases only. In addition, women with vulvar cancer were too few to be included in the regression analysis, precluding the possibility of drawing any conclusions for this diagnosis group. Analyses of app-use frequency over 12 months were limited to the 49 participants with complete log data. Although participants with and without complete log data did not differ in measured baseline characteristics, unmeasured technical or implementation-related factors may still have influenced log data availability. The small number of participants with complete log data also limited the precision of the estimates. Therefore, findings related to app feature use and use frequency should be interpreted with caution.

Implications for Practice

The observed variation in app use across demographic and clinical groups indicates that a one-size-fits-all digital intervention may not be sufficient. Addressing these factors when developing interventions can help promote equitable and sustained use across diverse user groups and enhance engagement among groups less likely to use the app. In clinical practice, this highlights the need for targeted onboarding and user support, such as motivational guidance and technical assistance, tailored to the needs of different patient groups to ensure equal access to and benefit from app-based support in follow-up care.

Although app log data for the full first year were available only for a subset of the participants, the highest engagement with the physical activity and activity goal setting features among these participants suggests that encouraging regular physical activity and goal setting may be perceived as supportive by survivors, helping them monitor their activity levels and maintain a sense of control during recovery. Symptom alerts delivered directly to survivors may support earlier self-initiated contact with health care providers, thereby potentially facilitating timely clinical assessment when needed.

Conclusions

App-use patterns indicate differences in engagement across app features and participant characteristics. Findings from this study underscore the need to consider demographic and clinical characteristics when implementing digital tools in follow-up settings to ensure equal participation and avoid disparities in follow-up care. Overall, this study provides insight into the potential relevance of the LETSGO app in routine follow-up after cancer treatment, suggesting that such mHealth approaches may be applicable across different gynecologic cancer diagnoses. These findings may also inform the development and implementation of future mHealth tools designed to support cancer survivors in routine follow-up care.

Acknowledgments

The authors would like to express their sincere appreciation to all participants for their contributions to this study. They also thank Anette Strømsbø Gjørv for assistance in improving the visual content of the LETSGO (Lifestyle and Empowerment Techniques in Survivorship of Gynecologic Oncology) app used in this manuscript.

During the preparation of this work, the authors used ChatGPT (GPT-4o; OpenAI) and Microsoft Copilot, which were available to students at the University of Oslo, to improve the readability and language of the Introduction and Discussion sections, without altering the content. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the final publication.

Funding

This study received financial support from Sorlandet Hospital, Kristiansand, Norway.

XSL•FO
RenderX

Data Availability

Demographic data used in this study were stored in Viedoc software (Viedoc Technologies AB), a secure electronic data capture system. App data were stored on the Services for Sensitive Data (TSD) platform at the University of Oslo, a secure system for handling sensitive research data, in accordance with Norwegian regulations for handling sensitive research data. Because of ethical restrictions, the data cannot be publicly available.

Authors' Contributions

Conceptualization: EB, IV
Data curation: SHF
Formal analysis: MH, ZT
Funding acquisition: IV
Investigation: IV
Methodology: IV, EB, MH
Project administration: IV
Supervision: IV, EB, MH
Visualization: MH, ZT, SHF
Writing—original draft: ZT
Writing—review and editing: EB, IV, MH, SB, SHF, ZT
All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Information about the disease, treatment, and late effects.
[PNG File, 207 KB - cancer_v12i1e89918_app1.png]

Multimedia Appendix 2
Symptom registration feature.
[PNG File, 966 KB - cancer_v12i1e89918_app2.png]

Multimedia Appendix 3
Physical activity feature.
[PNG File, 658 KB - cancer_v12i1e89918_app3.png]

Multimedia Appendix 4
Exercise program feature.
[PNG File, 957 KB - cancer_v12i1e89918_app4.png]

References

1. Pham EN, van den Berg CB, Boere I, et al. Quality of life and survival in patients with uterine carcinosarcoma: a tertiary center observational study. Gynecol Oncol Rep 2025;57:101679. [doi: 10.1016/j.gore.2025.101679] [Medline: 39902376]
2. Seland M, Skrede K, Lindemann K, et al. Distress, problems and unmet rehabilitation needs after treatment for gynecological cancer. Acta Obstet Gynecol Scand 2022 Mar;101(3):313-322. [doi: 10.1111/aogs.14310] [Medline: 34964982]
3. La Spina S, Scollo P, Pecorino B, et al. Life experience of survivors of gynecologic cancers: a survey conducted in Italy. Oncology (Williston Park) 2024 Jan 9;38(1):15-19. [doi: 10.46883/2024.25921012] [Medline: 38300531]
4. Gil-Ibanez B, Davies-Oliveira J, Lopez G, Díaz-Feijoo B, Tejerizo-Garcia A, Sehouli J. Impact of gynecological cancers on health-related quality of life: historical context, measurement instruments, and current knowledge. Int J Gynecol Cancer 2023 Nov 6;33(11):1800-1806. [doi: 10.1136/ijgc-2023-004804] [Medline: 37696648]
5. van Leeuwen M, Husson O, Alberti P, et al. Understanding the quality of life (QOL) issues in survivors of cancer: towards the development of an EORTC QOL cancer survivorship questionnaire. Health Qual Life Outcomes 2018 Jun 4;16(1):114. [doi: 10.1186/s12955-018-0920-0] [Medline: 29866185]
6. Froeding LP, Greimel E, Lanceley A, et al. Assessing patient-reported quality of life outcomes in vulva cancer patients: a systematic literature review. Int J Gynecol Cancer 2018 May;28(4):808-817. [doi: 10.1097/IGC.0000000000001211] [Medline: 29420364]
7. Sekse RJ, Dunberger G, Olesen ML, Østerbye M, Seibaek L. Lived experiences and quality of life after gynaecological cancer-an integrative review. J Clin Nurs 2019 May;28(9-10):1393-1421. [doi: 10.1111/jocn.14721] [Medline: 30461101]

XSL•FO
RenderX

8. Breistig S, Thorkildsen KM, Tveit Sekse RJ. Gynecological cancer survivors' experiences and desire for follow-up after recent treatment: a phenomenological hermeneutic study. Cancer Nurs 2024;47(5):E327-E335. [doi: 10.1097/NCC.0000000000001252] [Medline: 37272739]
9. Coping – follow-up medical care. National Cancer Institute National Institutes of Health. URL: https://www.cancer.gov/about-cancer/coping/survivorship/follow-up-care [accessed 2026-07-09]
10. Shapiro CL. Cancer survivorship. N Engl J Med 2018 Dec 20;379(25):2438-2450. [doi: 10.1056/NEJMra1712502] [Medline: 30575480]
11. Jacobs LA, Shulman LN. Follow-up care of cancer survivors: challenges and solutions. Lancet Oncol 2017 Jan;18(1):e19-e29. [doi: 10.1016/S1470-2045(16)30386-2] [Medline: 28049574]
12. Miller KD, Nogueira L, Devasia T, et al. Cancer treatment and survivorship statistics, 2022. CA Cancer J Clin 2022 Sep;72(5):409-436. [doi: 10.3322/caac.21731] [Medline: 35736631]
13. Avery J, Thomas R, Howell D, Dubouloz Wilner CJ. Empowering cancer survivors in managing their own health: a paradoxical dynamic process of taking and letting go of control. Qual Health Res 2023 Apr;33(5):412-425. [doi: 10.1177/10497323231158629] [Medline: 36825869]
14. Girgis A. The role of self-management in cancer survivorship care. Lancet Oncol 2020 Jan;21(1):8-9. [doi: 10.1016/S1470-2045(19)30715-6] [Medline: 31838013]
15. Howell D, Harth T, Brown J, Bennett C, Boyko S. Self-management education interventions for patients with cancer: a systematic review. Support Care Cancer 2017 Apr;25(4):1323-1355. [doi: 10.1007/s00520-016-3500-z] [Medline: 28058570]
16. Vistad I, Skorstad M, Berntsen S, et al. Empowerment and quality of life in gynecological cancer survivors: outcomes from a multicenter quasi-experimental cohort study from Norway (the LETSGO trial). Cancer 2025 Aug 15;131(16):e70040. [doi: 10.1002/cncr.70040] [Medline: 40781852]
17. Penedo FJ, Oswald LB, Kronenfeld JP, Garcia SF, Cella D, Yanez B. The increasing value of eHealth in the delivery of patient-centred cancer care. Lancet Oncol 2020 May;21(5):e240-e251. [doi: 10.1016/S1470-2045(20)30021-8] [Medline: 32359500]
18. Lee K, Kim S, Kim SH, et al. Digital health interventions for adult patients with cancer evaluated in randomized controlled trials: scoping review. J Med Internet Res 2023 Jan 6;25:e38333. [doi: 10.2196/38333] [Medline: 36607712]
19. Qin M, Chen B, Sun S, Liu X. Effect of mobile phone app-based interventions on quality of life and psychological symptoms among adult cancer survivors: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022 Dec 19;24(12):e39799. [doi: 10.2196/39799] [Medline: 36534460]
20. Buneviciene I, Mekary RA, Smith TR, Onnela JP, Bunevicius A. Can mHealth interventions improve quality of life of cancer patients? A systematic review and meta-analysis. Crit Rev Oncol Hematol 2021 Jan;157:103123. [doi: 10.1016/j.critrevonc.2020.103123] [Medline: 33190065]
21. Adam R, McMichael D, Powell D, Murchie P. Publicly available apps for cancer survivors: a scoping review. BMJ Open 2019 Sep 30;9(9):e032510. [doi: 10.1136/bmjopen-2019-032510] [Medline: 31575584]
22. Levine DM, Co Z, Newmark LP, et al. Design and testing of a mobile health application rating tool. NPJ Digit Med 2020;3:74. [doi: 10.1038/s41746-020-0268-9] [Medline: 32509971]
23. Number of mHealth apps available in the Apple App Store from 1st quarter 2015 to 2nd quarter 2024. Statista. URL: https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide [accessed 2025-09-23]
24. Jat AS, Grønli TM. Harnessing the digital revolution: a comprehensive review of mHealth applications for remote monitoring in transforming healthcare delivery. In: Younas M, Awan I, Grønli TM, editors. Mobile Web and Intelligent Information Systems: Springer; 2023. [doi: 10.1007/978-3-031-39764-6_4]
25. Briggs LG, Labban M, Alkhatib K, Nguyen DD, Cole AP, Trinh QD. Digital technologies in cancer care: a review from the clinician's perspective. J Comp Eff Res 2022 May;11(7):533-544. [doi: 10.2217/cer-2021-0263] [Medline: 35416050]
26. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565. [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
27. Denis F, Basch E, Septans AL, et al. Two-year survival comparing web-based symptom monitoring vs routine surveillance following treatment for lung cancer. JAMA 2019 Jan 22;321(3):306-307. [doi: 10.1001/jama.2018.18085] [Medline: 30667494]
28. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021 Jul 29;11(7):1473-1480. [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
29. Kemp E, Trigg J, Beatty L, et al. Health literacy, digital health literacy and the implementation of digital health technologies in cancer care: the need for a strategic approach. Health Promot J Austr 2021 Feb;32 Suppl 1:104-114. [doi: 10.1002/hpja.387] [Medline: 32681656]
30. Kelders SM, van Zyl LE, Ludden GD. The concept and components of engagement in different domains applied to eHealth: a systematic scoping review. Front Psychol 2020;11:926. [doi: 10.3389/fpsyg.2020.00926] [Medline: 32536888]
31. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. JMIR Res Protoc 2017 Aug 7;6(8):e156. [doi: 10.2196/resprot.6452] [Medline: 28784592]

XSL•FO
RenderX

32. van Gemert-Pijnen JE, Nijland N, van Limburg M, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 5;13(4):e111. [doi: 10.2196/jmir.1672] [Medline: 22155738]
33. Vaffis S, Whaley S, Axon DR, et al. Features of cancer mHealth apps and evidence for patient preferences: scoping literature review. JMIR Cancer 2023 Apr 28;9:e37330. [doi: 10.2196/37330] [Medline: 37115587]
34. Vistad I, Skorstad M, Demmelmaier I, et al. Lifestyle and Empowerment Techniques in Survivorship of Gynaecologic Oncology (LETSGO study): a study protocol for a multicentre longitudinal interventional study using mobile health technology and biobanking. BMJ Open 2021 Jul 12;11(7):e050930. [doi: 10.1136/bmjopen-2021-050930] [Medline: 34253678]
35. Skorstad M, Vistad I, Fegran L, Berntsen S, Johannessen B. Nurse-led consultations reinforced with eHealth technology: a qualitative study of the experiences of patients with gynecological cancer. BMC Nurs 2022 Nov 25;21(1):326. [doi: 10.1186/s12912-022-01104-9] [Medline: 36434602]
36. Vistad I, Bjørge L, Solheim O, et al. A national, prospective observational study of first recurrence after primary treatment for gynecological cancer in Norway. Acta Obstet Gynecol Scand 2017 Oct;96(10):1162-1169. [doi: 10.1111/aogs.13199] [Medline: 28795770]
37. Gordon NP, Crouch E. Digital information technology use and patient preferences for internet-based health education modalities: cross-sectional survey study of middle-aged and older adults with chronic health conditions. JMIR Aging 2019 Apr 4;2(1):e12243. [doi: 10.2196/12243] [Medline: 31518291]
38. Li C, Neugroschl J, Zhu CW, et al. Design considerations for mobile health applications targeting older adults. J Alzheimers Dis 2021;79(1):1-8. [doi: 10.3233/JAD-200485] [Medline: 33216024]
39. Chowdhury SZ, Stevens S, Wu C, et al. An age-old problem or an old-age problem? A UK survey of attitudes, historical use and recommendations by healthcare professionals to use healthcare apps. BMC Geriatr 2023 Feb 24;23(1):110. [doi: 10.1186/s12877-023-03772-x] [Medline: 36823564]
40. Győrffy Z, Boros J, Döbrössy B, Girasek E. Older adults in the digital health era: insights on the digital health related knowledge, habits and attitudes of the 65 year and older population. BMC Geriatr 2023 Nov 27;23(1):779. [doi: 10.1186/s12877-023-04437-5] [Medline: 38012565]
41. König LM, Western MJ, Denton AH, Krukowski RA. Umbrella review of social inequality in digital interventions targeting dietary and physical activity behaviors. NPJ Digit Med 2025 Jan 6;8(1):11. [doi: 10.1038/s41746-024-01405-0] [Medline: 39762352]
42. Yuen E, Winter N, Savira F, et al. Digital health literacy and its association with sociodemographic characteristics, health resource use, and health outcomes: rapid review. Interact J Med Res 2024 Jul 26;13:e46888. [doi: 10.2196/46888] [Medline: 39059006]
43. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
44. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
45. Causio FA, Beccia F, Tona DM, et al. Public perceptions and engagement in mHealth: a European survey on attitudes toward health apps use and data sharing. Eur J Public Health 2025 Jun 1;35(3):401-406. [doi: 10.1093/eurpub/ckaf036] [Medline: 40249875]
46. Wang C, Qi H. Influencing factors of acceptance and use behavior of mobile health application users: systematic review. Healthcare (Basel) 2021 Mar 22;9(3):357. [doi: 10.3390/healthcare9030357] [Medline: 33809828]
47. Zhao BY, Huang L, Cheng X, et al. Digital health literacy and associated factors among internet users from China: a cross-sectional study. BMC Public Health 2024 Mar 27;24(1):908. [doi: 10.1186/s12889-024-18324-0] [Medline: 38539176]
48. Vistad I, Cvancarova M, Lindviksmoen Astrup G, Rustøen T, Liavaag AH. Symptom experience and self-rated physical functioning in patients with ovarian cancer receiving chemotherapy: a longitudinal study. Int J Gynecol Cancer 2018 Jul;28(6):1167-1175. [doi: 10.1097/IGC.0000000000001287] [Medline: 29781824]
49. Lee YC, King MT, O'Connell RL, et al. Symptom burden and quality of life with chemotherapy for recurrent ovarian cancer: the Gynecologic Cancer InterGroup-Symptom Benefit Study. Int J Gynecol Cancer 2022 Jun 6;32(6):761-768. [doi: 10.1136/ijgc-2021-003142] [Medline: 35086926]
50. Mooney K, Gullatte M, Iacob E, et al. Essential components of an electronic patient-reported symptom monitoring and management system: a randomized clinical trial. JAMA Netw Open 2024 Sep 3;7(9):e2433153. [doi: 10.1001/jamanetworkopen.2024.33153] [Medline: 39269704]
51. Crafoord MT, Fjell M, Sundberg K, Nilsson M, Langius-Eklöf A. Engagement in an interactive app for symptom self-management during treatment in patients with breast or prostate cancer: mixed methods study. J Med Internet Res 2020 Aug 10;22(8):e17058. [doi: 10.2196/17058] [Medline: 32663140]
52. Cella D, Lancki N, Kuharic M, et al. Web-based cancer symptom self-management system: a randomized clinical trial. JAMA Netw Open 2025 May 1;8(5):e258353. [doi: 10.1001/jamanetworkopen.2025.8353] [Medline: 40323601]
53. Xiao K, Tang L, Chen Y, Zhou J, Yang Q, Wang R. The effectiveness of E-health interventions promoting physical activity in cancer survivors: a systematic review and meta-analysis of randomized controlled trials. J Cancer Res Clin Oncol 2024 Feb 2;150(2):72. [doi: 10.1007/s00432-023-05546-9] [Medline: 38305910]

XSL•FO
RenderX

54. Khoo S, Mohbin N, Ansari P, Al-Kitani M, Müller AM. mHealth interventions to address physical activity and sedentary behavior in cancer survivors: a systematic review. Int J Environ Res Public Health 2021 May 28;18(11):5798. [doi: 10.3390/ijerph18115798] [Medline: 34071342]
55. Wu HS, Gal R, van Sleeuwen NC, et al. Breast cancer survivors' experiences with an activity tracker integrated into a supervised exercise program: qualitative study. JMIR Mhealth Uhealth 2019 Feb 21;7(2):e10820. [doi: 10.2196/10820] [Medline: 30789349]
56. Signorelli GR, Monteiro-Guerra F, Rivera-Romero O, Núñez-Benjumea FJ, Fernández-Luque L. Breast cancer physical activity mobile intervention: early findings from a user experience and acceptability mixed methods study. JMIR Form Res 2022 Jun 22;6(6):e32354. [doi: 10.2196/32354] [Medline: 35731554]
57. Zhang Y, Xu P, Sun Q, Baral S, Xi L, Wang D. Factors influencing the e-health literacy in cancer patients: a systematic review. J Cancer Surviv 2023 Apr;17(2):425-440. [doi: 10.1007/s11764-022-01260-6] [Medline: 36190672]

Abbreviations

EORTC: European Organisation for Research and Treatment of Cancer
LETSGO: Lifestyle and Empowerment Techniques in Survivorship of Gynecologic Oncology
mHealth: mobile health
OR: odds ratio
TSD: Services for Sensitive Data

Edited by M Balcarras; submitted 26.Dec.2025; peer-reviewed by D Wong, R Desiree; revised version received 17.Jun.2026; accepted 18.Jun.2026; published 31.Jul.2026.

Please cite as:
Tarabar Z, Børøsund E, H Fosstveit S, Berntsen S, Hagen M, Vistad I
Factors Associated With the Use of the Lifestyle and Empowerment Techniques in Survivorship of Gynecologic Oncology (LETSGO) mHealth App in Routine Follow-Up After Gynecologic Cancer Treatment: Observational Study
JMIR Cancer 2026;12:e89918
URL: https://cancer.jmir.org/2026/1/e89918
doi:10.2196/89918

© Zaklina Tarabar, Elin Børøsund, Sindre Herskedal Fosstveit, Sveinung Berntsen, Milada Hagen, Ingvild Vistad. Originally published in JMIR Cancer (https://cancer.jmir.org), 31.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessment of Physical Activity Engagement and Promotion Among Cancer Survivors in Cardio-Oncology Care: Qualitative Study

Macy K Tetrick[1], BS; Daniel Addison[1,2,3,4], MD; Rujul Singh[1,5], BS; Sara M St. George[6,7], PhD; Vanina Pavia[7], MD, MS; Anvitha Gogineni[8], BS; Peter Washington[9], PhD; Sanam M Ghazi[2,3], MD; Olivia Preston[2], BS; James L Fisher[10], PhD; Roberto M Benzo[1,11], PhD

[1]Division of Cancer Prevention and Control, Department of Internal Medicine, College of Medicine, The Ohio State University, 3650 Olentangy River Rd, Columbus, OH, United States

[10]Arthur G. James Cancer Hospital and Richard J. Solove Research Institute, Columbus, OH, United States

[11]The Ohio State University Comprehensive Cancer Center, The Ohio State University Wexner Medical Center, Columbus, OH, United States

[2]Cardio-Oncology Program, Division of Cardiology, The Ohio State University Medical Center, Columbus, OH, United States

[3]Cardio-Oncology Program, Division of Cardiology, University of Texas Southwestern Medical Center, Dallas, TX, United States

[4]Division of Epidemiology, College of Public Health, Ohio State University, Columbus, OH, United States

[5]Grossman School of Medicine, New York University, New York, NY, United States

[6]Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL, United States

[7]Department of Public Health Sciences, University of Miami Miller School of Medicine, Miami, FL, United States

[8]College of Medicine, The Ohio State University, Columbus, OH, United States

[9]Division of Clinical Informatics and Digital Transformation (DoC-IT), Department of Medicine, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Macy K Tetrick, BS
Division of Cancer Prevention and Control, Department of Internal Medicine, College of Medicine, The Ohio State University, 3650 Olentangy River Rd, Columbus, OH, United States

Abstract

Background: Cardiovascular disease and cancer are leading causes of morbidity and mortality worldwide, with cancer survivors facing elevated cardiovascular risk due to shared risk factors and cardiotoxic treatments. Physical activity (PA) is a well-established strategy to improve cardiovascular health and survivorship outcomes; however, most cancer survivors do not meet recommended PA guidelines. Understanding multilevel influences on PA engagement and promotion is essential to improving implementation in cardio-oncology care.

Objective: This study aimed to examine barriers and facilitators to PA engagement and promotion in cancer survivorship from both survivor and medical team perspectives using the Capability, Opportunity, Motivation-Behavior (COM-B) framework.

Methods: A descriptive qualitative study was conducted with 12 cancer survivors and 12 medical team members involved in oncology care. Semistructured interviews explored experiences, perceptions, and practices related to PA across the cancer care continuum. Rapid qualitative analysis was used to identify themes, which were mapped to COM-B domains. Participants also completed surveys assessing PA, cardiovascular health (Life's Essential 8 [LE8]), and sociodemographic factors.

Results: Survivors had a mean age of 65.3 (SD 12.6) years and were evenly distributed by sex. PA engagement varied, with 42% (5/12) of survivors meeting recommended guidelines. The mean modified LE8-derived cardiovascular health score was 71.3 (SD 11.9), reflecting a standardized 6-domain composite (diet, PA, nicotine exposure, sleep, BMI, and blood pressure) calculated identically for all 12 participants. Across COM-B domains, physical limitation was the most commonly reported capability barrier (12/12 survivors), while current engagement in PA was the most prevalent capability facilitator (12/12 survivors). Opportunity was shaped by access to PA resources and social support, with social environment being the most frequently cited facilitator (12/12 survivors; 9/12 medical team members) and lack of access to PA resources being the most common barrier among medical team members (9/12). Motivation reflected perceived health outcomes as the most prevalent facilitator across both groups (12/12 survivors; 7/12 medical team members), alongside competing concerns such as safety and emotional responses. Medical team members' ability to promote PA was influenced by survivor readiness, clinical context, and resource availability.

Conclusions: PA engagement in cardio-oncology is shaped by dynamic interactions among capability, opportunity, and motivation across survivor and medical team member contexts. Although PA is widely valued, sustained engagement remains limited due

XSL•FO
RenderX

to fluctuating symptoms, environmental constraints, and clinical uncertainties. Interventions should prioritize flexible, survivor-centered approaches and incorporate supportive clinical communication and resource alignment to enhance PA across the survivorship trajectory.

(JMIR Cancer 2026;12:e104332) doi:10.2196/104332

KEYWORDS

cardio-oncology; physical activity; cancer survivorship; COM-B model; cardiovascular health; qualitative research; health behavior; exercise promotion; survivorship care; implementation science

Introduction

Background

Cardiovascular disease (CVD) and cancer are the 2 leading causes of death worldwide, highlighting the need for strategies to improve health among individuals affected by both conditions [1]. Compared with individuals without cancer, people with a cancer diagnosis experience a significantly higher risk of CVD, attributable to shared cardiometabolic and behavioral risk factors, cancer-related physiological changes, and exposure to cardiotoxic therapies [2,3]. As cancer survival has improved over recent decades, CVD has emerged as a leading cause of morbidity and mortality among cancer survivors, with noncancer deaths, particularly cardiovascular-related deaths, surpassing cancer-related deaths between 1973 and 2012 [4,5].

Cardiotoxic cancer therapies (eg, anthracyclines, chest radiation, and immune-based agents) significantly contribute to both short- and long-term cardiovascular morbidity and mortality [6]. As a result, CVD has become a major determinant of survivorship outcomes, with cancer survivors who develop CVD experiencing significantly worse survival and greater morbidity over time [7]. Consequently, the 8-year overall survival rate was 60% among those who developed CVD compared with 81% among those who did not, highlighting the critical role of cardiovascular health in long-term cancer survivorship [7].

Given the central role of cardiovascular health in long-term cancer survivorship, identifying effective and scalable strategies to mitigate cardiovascular risk is a clinical priority. Physical activity (PA) represents one such strategy, with well-established benefits for cardiovascular health, functional capacity, and quality of life among individuals with cancer [8]. In a meta-analysis of controlled PA interventions in cancer survivors, PA was associated with significant improvements in cardiorespiratory fitness both during and after cancer treatment, with weighted mean effect sizes (Cohen *d*) of 0.51 and 0.65, respectively ($P<.01$) [9]. Regular engagement in PA among cancer survivors improves cardiorespiratory fitness and reduces cardiovascular risk, helping to counteract treatment-related deconditioning and loss of aerobic capacity. This is clinically important given that cardiorespiratory fitness is a strong and independent predictor of cardiovascular and all-cause mortality [8,10]. In addition to cardiometabolic benefits, the American College of Sports Medicine has published PA recommendations that have been linked to improvement in common cancer-related symptoms, including fatigue, anxiety, depression, poor health-related quality of life, lymphedema, and poor physical function [11].

Clinical guidelines from the American Cancer Society (ACS) and the American Institute for Cancer Research (AICR) recommend regular PA before, during, and after cancer treatment, with adults advised to engage in a minimum of 150 minutes of moderate-to-vigorous physical activity (MVPA) per week, alongside muscle-strengthening activities [8,12]. Despite these recommendations, population-based data demonstrate that the majority of survivors do not meet recommended PA guidelines [13]. Despite ample evidence supporting the benefits of PA, engagement and promotion remain inconsistent, shaped by factors operating at the survivor, medical team member, and health care system levels [14,15]. Together, these findings suggest that the role of PA in cancer care is shaped by a range of factors that can act as both barriers to and facilitators of participation and promotion.

Compared with general oncology populations, patients receiving cardio-oncology care face actual or potential cardiovascular complications of cancer treatment in addition to cancer-related morbidity. PA may therefore be relevant to both cancer survivorship and cardiovascular health. Survivors may also have difficulty determining whether symptoms reflect cardiotoxicity, cancer treatment effects, or normal exertion, creating uncertainty about safe PA participation [16]. Further, promoting PA in cardio-oncology may require coordination between oncology and cardiology teams for counseling, safety guidance, and referral to exercise services [17]. These features make cardio-oncology a clinically distinct setting for examining the multilevel influences on PA engagement and promotion.

These multilevel influences highlight the need for a behavioral framework capable of integrating survivor, medical team member, and health care system determinants of PA engagement in cancer care. The Capability, Opportunity, Motivation-Behavior (COM-B) model provides a comprehensive framework for understanding and influencing health-related behaviors by explicitly identifying modifiable conditions required for a behavior to occur. Within this model, behavior is conceptualized as the result of interactions among individuals' physical and psychological capabilities, the social and environmental opportunities available to them, and both reflective and automatic motivational processes [18]. Applied to PA in cancer care, the COM-B framework offers a structured way to examine how survivor-level factors (eg, symptoms and beliefs), medical team member factors (eg, knowledge and time), and health care system factors (eg, access and resources) shape whether PA is initiated, sustained, and promoted.

Few studies have examined the influences of both survivor and medical team member factors in an integrated, theory-informed manner or explored how they are experienced and navigated

XSL•FO
RenderX

within clinical interactions. Limited qualitative work has simultaneously incorporated survivor and medical team perspectives to understand how PA engagement and promotion unfold across survivorship. As a result, there remains an incomplete understanding of the specific barriers, facilitators, and contextual conditions that shape PA as both a survivor behavior and a clinical practice.

Prior research suggests that PA engagement in cancer survivorship is shaped by not only individual factors but also clinical and system-level influences. Although clinicians widely endorse PA, gaps remain in how it is assessed, discussed, and supported in practice, with uncertainty regarding roles, time constraints, and available referral pathways limiting consistent promotion. Survivors also report interest in receiving guidance from their care teams, yet counseling and support vary across settings [19]. Existing qualitative work has largely focused on survivor experiences alone, with less attention to how survivor and medical team member perspectives intersect within clinical care. As a result, there remains limited understanding of how PA is both experienced by survivors and operationalized in clinical interactions.

Objectives

To address this gap, this qualitative study uses the COM-B framework to examine PA engagement and promotion in cancer survivorship from both survivor and medical team perspectives. By exploring how capability, opportunity, and motivation interact across roles and contexts, this study aims to generate actionable insights to inform more responsive clinical practice and intervention design for promoting and supporting PA throughout the survivorship trajectory.

Methods

Study Design

We conducted a descriptive qualitative study guided by the COM-B framework to examine PA engagement and promotion in the context of cancer treatment and survivorship. Semistructured interviews were used to explore perceived barriers and facilitators influencing PA behaviors among cancer survivors and medical team members involved in cancer care. The COM-B framework was used as an organizing lens to structure the examination of barriers and facilitators of PA behavior across medical team personnel (in terms of promotion) and cancer survivors (in terms of engagement).

In addition, cancer survivor participants completed a brief survey to capture self-reported sociodemographic characteristics, PA, and overall health measures. PA was assessed using the Rapid Assessment of Physical Activity (RAPA), a validated self-report measure of PA levels in adults [20]. The RAPA includes a series of yes/no items reflecting increasing levels of aerobic PA, with scores ranging from 1 (sedentary) to 7 (active). The final RAPA score was assigned based on the highest affirmed item. Consistent with the validated scoring criteria, participants with scores of 6 or higher were classified as "active." Item 6 of the RAPA indicates engaging in at least 30 minutes of moderate PA on 5 or more days per week, whereas item 7 indicates engaging in at least 20 minutes of vigorous PA on 3 or more days per week. The RAPA also includes items assessing strength and flexibility activities, asking whether participants engage in muscle-strengthening activities (eg, lifting weights and calisthenics) and flexibility activities (eg, stretching and yoga) at least once per week. These items are scored separately from the RAPA aerobic score and were not incorporated into the aerobic score or the study's classification of "active" status, which was based solely on the aerobic component. They are reported here descriptively. In addition to the RAPA, survivors reported weekly minutes of moderate or greater intensity activity as part of the PA domain of the Life's Essential 8 (LE8) cardiovascular health assessment (described below). This continuous self-report item was classified against the ACS/AICR guideline of 150 or more minutes per week, providing a second, independently derived measure of guideline-consistent PA. Because the RAPA and this continuous minutes-based measure use different item formats and scoring approaches, the resulting classifications may differ for some participants.

Select social determinants of health, including income, education, and employment status, were assessed using the Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences (PRAPARE) [21]. The full PRAPARE domain set was not analyzed for this manuscript.

Cardiovascular health was assessed using the American Heart Association's LE8 framework, which results in a composite score based on 8 health behaviors and factors: diet, PA, nicotine exposure, sleep health, BMI, blood lipids, blood glucose, and blood pressure [22]. The diet component was operationalized using a Mediterranean diet–based scoring approach, consistent with the Modified Mediterranean Eating Pattern for Americans (MEPA) recommended within the LE8 framework [22,23]. When individual LE8 components were missing, composite scores were calculated using the 6 domains available for all participants (diet, PA, nicotine exposure, sleep, blood pressure, and BMI). Blood glucose was not included in the survey instrument and therefore was not assessed. Blood lipid data were available for only 4 of 12 participants (33%) due to missing responses. Both domains were therefore excluded from the primary composite and are reported descriptively where available. Given the small sample size, validated imputation approaches were not considered appropriate to address missing domain data. Because some participants were unable to report clinical measures, survey-derived health measures were analyzed descriptively to characterize overall health status within the sample.

All 12 survivor participants completed the full survey battery. Blood lipid values were collected via self-report. Some participants left this item blank or did not respond, resulting in missing data for this domain. Blood glucose was not included in the survey instrument and was therefore not assessed for any participant.

The target behavior for this analysis was engagement in and promotion of PA throughout cancer survivorship. Because PA behaviors are shaped by complementary roles within the clinical relationship (survivors engaging in PA and medical team members providing encouragement and guidance), both survivor and medical team perspectives were included.

XSL•FO
RenderX

Participants and Recruitment

Survivor participants (n=12) were adults aged 18 years or older who had received a cancer diagnosis and were currently undergoing or had previously undergone cancer treatment involving cardiotoxic therapies, including targeted, biologic, radiation, or immune-based treatments. Medical team members (n=12) were health care professionals involved in the care of oncological survivors. Eligible survivor participants were required to have access to a device capable of supporting remote interviews (eg, smartphone, computer, or tablet) and to self-report interest in tracking cardiovascular symptoms and engaging in PA. Individuals who were pregnant, incarcerated, or unable to complete the interview in English were excluded.

Survivors were recruited through clinician referrals within a cardio-oncology clinic. Clinicians introduced the study to eligible survivors during routine appointments, and interested individuals were contacted by the research team to provide additional information and confirm eligibility. Medical team participants were recruited via professional referrals and word of mouth. All participants provided informed consent prior to enrollment.

Data Collection

Semistructured interviews were conducted with research staff members remotely via videoconferencing platforms (eg, Zoom and Microsoft Teams) between December 2023 and July 2024. Interviews were conducted by the lead author (MKT) and/or the study's principal investigator (RMB). Both were trained in rapid qualitative analysis (RQA) and qualitative interviewing techniques. Neither interviewer had a prior relationship with survivor or medical team participants. All participants were referred to the research team by the cardio-oncology clinic rather than recruited directly by the interviewers. Reflexivity was supported through regular team meetings in which interpretations were discussed and reconciled collaboratively rather than by a single analyst. While each interview followed the semistructured interview guide, interviewers remained flexible and allowed participants to elaborate on relevant topics as they arose to ensure complete data capture.

The semistructured interview guides used to conduct both survivor and medical team member interviews are available in Multimedia Appendix 1. Survivor interview questions explored experiences engaging in PA before, during, and after cancer treatment; reasons for being physically active or not being physically active; and external factors that may help or hinder PA engagement. Medical team participants were asked about their roles, confidence, and approaches to encouraging PA among survivors. Prompts were aligned with COM-B domains, including capability (eg, symptoms and knowledge), opportunity (eg, time, access, and social support), and motivation (eg, beliefs and priorities). Survivor participants received a US $25 gift card upon interview completion.

Interviews lasted approximately 60 - 90 minutes and were audio-recorded with participant consent. Most interviews were also video-recorded, although participants were given the option to keep their cameras off. Interviews were transcribed verbatim using a Zoom- or Teams-generated transcript and edited by a research team member when necessary. Survey data were analyzed descriptively to summarize participant characteristics, PA levels, cardiovascular health, and self-rated general health status.

Data Analysis

Analysis followed the RQA approach by Hamilton and Finley [24], with attention to best practices for rigor and validity [25,26]. RQA, used to analyze interview data, is a team-based approach that emphasizes systematic review and synthesis of qualitative data to identify actionable themes [24,25]. This study used a hybrid, deductive-dominant analytic approach: the COM-B framework was applied a priori to structure both the interview guides and the RQA summarization process. Research team members were trained in the RQA process. One researcher completed an initial review of each transcript, summarizing key points directly into structured templates organized by the COM-B domain and subdomain. A second researcher subsequently reviewed the same transcript, with access to the initial summary, to compare interpretations. Because reviewers were not blinded to one another's summaries, a formal interrater agreement statistic was not calculated. Discrepancies were resolved through team discussion and consensus. Within this deductively structured framework, the specific content of facilitator and barrier themes was generated inductively from participant data. Summaries were then organized into analytic matrices to enable comparison across participants and facilitate theme development. A template of the analytic matrix structure used for theme development is provided in Multimedia Appendix 2. Data collection and analysis occurred concurrently over approximately 1 year, with transcripts reviewed and summarized in a rolling manner as interviews were completed. As data collection and analysis progressed, the team observed that later interviews increasingly reinforced previously identified themes rather than introducing substantially new ones, consistent with thematic redundancy. Formal saturation metrics were not applied as a stopping criterion. Sample size was determined a priori based on the scope of the study. Final themes were mapped to the COM-B framework and categorized as barriers or facilitators within the overarching domains of capability, opportunity, and motivation. Consistent with the COM-B model, these domains were further specified as physical and psychological capability, physical and social opportunity, and reflective and automatic motivation. Representative quotations were selected to illustrate each theme. Sociodemographic characteristics of survivor participants were collected via a survey and summarized descriptively.

Ethical Considerations

This study was approved by The Ohio State University Institutional Review Board (IRB #2021C0018). All participants provided informed consent prior to participation. Interview recordings and transcripts were deidentified, assigned a unique study ID, and stored securely in accordance with institutional policies.

XSL•FO
RenderX

Results

Participant Characteristics

Participant demographic and clinical characteristics derived from both survey and interview data are presented in Table 1. A total of 12 cancer survivors participated in interviews and completed surveys. Survivor participants had a mean age of 65.3 (SD 12.6) years, with an equal representation of women and men. Most survivors identified as White and non-Hispanic (10/12, 83%). BMI values indicated that the majority of survivors were either overweight (5/12, 42%) or obese (4/12, 33%), with a mean BMI of 26.9 (SD 4.53) kg/m². Educational attainment was generally high, with 58% (7/12) of participants reporting completion of a graduate or professional degree. Most survivors (10/12, 83%) were retired or receiving disability benefits, and annual household income most frequently fell between US $100,000 and $199,999 (5/12, 42%). Cancer diagnoses were heterogeneous. The most commonly represented diagnosis was acute myeloid leukemia (3/12, 25%), followed by breast cancer (2/12, 17%), with several other malignancies represented in smaller numbers. Most survivors reported exposure to multiple cancer-directed therapies, reflecting complex treatment histories. These included chemotherapy (6/12, 50%), radiation therapy (5/12, 42%), tyrosine kinase inhibitors (4/12, 33%), stem cell transplantation (4/12, 33%), hormonal therapy (3/12, 25%), and surgery (3/12, 25%), with fewer participants reporting exposure to immune checkpoint inhibitors (1/12, 8%) or other treatment modalities (4/12, 33%).

XSL•FO
RenderX

Table . Demographic and clinical characteristics of the sample population.

Variable	Survivors (n=12)	Medical team members (n=12)
Age (years), mean (SD)	65.3 (12.6)	37.8 (6.4)
Time in position (years), mean	__[a]	4.5
Sex, n (%)		
Female	6 (50)	7 (58)
Male	6 (50)	5 (42)
Race or ethnicity, n (%)		
White, non-Hispanic	10 (83)	6 (50)
White, Hispanic	0 (0)	1 (8)
Black or African American	1 (8)	1 (8)
Asian	0 (0)	3 (25)
Multiple races or ethnicities	1 (8)	1 (8)
BMI (kg/m^2), n (%)		
18.5 - 24.9	3 (25)	—
25 - 29.9	5 (42)	—
≥30	4 (33)	—
BMI, mean (SD)	26.9 (4.53)	—
Employment[b] (most common category), n (%)		
Retired or disabled	10 (83)	—
Education[b] (most common category), n (%)		
Graduate or professional degree	7 (58)	—
Annual income[b] (most common category), n (%)		
US $100,000-199,999	5 (42)	—
Job position, n (%)		
Cardio-oncologist	—	4 (33)
Oncologist	—	3 (25)
Radiation oncologist	—	1 (8)
Nurse practitioner	—	1 (8)
Physical therapist	—	1 (8)
Exercise physiologist	—	1 (8)
Dietitian	—	1 (8)
Cancer type, n (%)		
Acute myeloid leukemia	3 (25)	—
Breast	2 (17)	—
Prostate	1 (8)	—
Lymphoma	1 (8)	—
Multiple myeloma	1 (8)	—
Uterine	1 (8)	—
Chronic lymphocytic leukemia	1 (8)	—
Lung	1 (8)	—
Neuroendocrine malignancy	1 (8)	—
Cancer treatment received, n (%)		

XSL•FO
RenderX

Variable	Survivors (n=12)	Medical team members (n=12)
Chemotherapy	6 (50)	—
Radiation	5 (42)	—
Stem cell transplantation	4 (33)	—
Tyrosine kinase inhibitors	4 (33)	—
Hormonal therapy	3 (25)	—
Surgery	3 (25)	—
Immune checkpoint inhibitors	1 (8)	—
Others	4 (33)	—

[a]Not collected or not applicable.

[b]Complete education, employment, and annual income data are provided in Table S2 in Multimedia Appendix 3.

In addition, 12 medical team members involved in the care of survivors with cancer completed interviews. Medical team members had a mean age of 37.8 (SD 6.4) years and reported an average of 4.5 years in their current professional roles. Of the 12 medical team members, 7 (58%) identified as female, and the team represented a range of racial and ethnic backgrounds, including White and non-Hispanic (6/12, 50%), Asian (3/12, 25%), White and Hispanic (1/12, 8%), Black or African American (1/12, 8%), and multiple racial or ethnic identities (1/12, 8%). Professional disciplines included cardio-oncology physician (4/12, 33%), medical oncologist (3/12, 25%), and radiation oncologist (1/12, 8%), as well as allied health professionals such as nurse practitioner (1/12, 8%), physical therapist (1/12, 8%), exercise physiologist (1/12, 8%), and dietitian (1/12, 8%).

Self-Reported PA

Based on self-reported weekly PA minutes collected as part of the LE8 PA domain, 42% (5/12) of survivor participants met recommended PA guidelines (≥150 min per week of MVPA), while 58% (7/12) did not. Using the separate, categorical RAPA instrument, 50% (6/12) of participants were classified as active, while 50% (6/12) were classified as underactive. Of the 12 participants, 9 (75%) were classified concordantly across the 2 measures: 4 met both criteria and 5 met neither; the remaining 3 were classified differently. These 2 measures rely on different item structures and scoring algorithms and are not expected to yield identical classifications for all participants.

For descriptive purposes only, and independent of the RAPA aerobic score reported above, 67% (8/12) of survivor participants reported engaging in strength training activities at least once per week, and 50% (6/12) reported engaging in flexibility activities at least once per week. Categorized jointly, 42% (5/12) reported both, 33% (4/12) reported one only, and 25% (3/12) reported neither.

The mean cardiovascular health score (modified LE8-derived score) was 71.3 (SD 11.9). Descriptive details, including LE8 domain-level scores, are provided in Table S1 in Multimedia Appendix 3.

COM-B Themes

The results are organized according to the COM-B framework, capturing themes related to capability, opportunity, and motivation across both survivor and medical team member perspectives. Themes and subthemes identified within each domain are summarized in Table 2. In the text below, we describe these themes in greater detail, using representative quotations to illustrate how participants experienced and interpreted PA engagement and promotion in cancer survivorship. A table of additional quotes representing each theme can be found in Multimedia Appendix 4.

XSL•FO
RenderX

Table . Summary of the domain system (interview data).

COM-B[a] domain, facilitator/barrier, and subdomain	Survivor interview theme (n)	Medical team member interview theme (n)
Capability		
Facilitator		
Physical	• Current engagement in PA[b] (n=12)	• Survivor self-efficacy (n=3)
Psychological	• Knowledge and awareness of PA levels and ability (n=9)	• Knowledge and awareness of survivor PA levels and ability (n=9)
Barrier		
Physical	• Physical limitations (n=12)	• Survivors' symptom- or treatment-related burden (n=2)
Psychological	—[c]	• Lack of knowledge on survivor PA levels (n=3)
Opportunity		
Facilitator		
Physical	• PA resources (n=11)	• Online PA resources (n=4)
Social	• Supportive social environments (classes, family, and friends) (n=12)	• Supportive social environments for survivors (n=9)
Barrier		
Physical	• Weather and time (n=7)	• Lack of access to PA resources (n=9)
Motivation		
Facilitator		
Reflective	• Perceived health outcomes (n=12) • PA-related resources (n=7) • Intentions to be active (n=7)	• Perceived health outcomes (n=7) • Survivors' inclination to be active (n=3)
Automatic	• Emotions surrounding PA (n=5) • Routine or habit (n=7)	—
Barrier		
Reflective	• Physical limitations (n=8) • Safety concerns, resource preferences (n=3)	• Treatment-related limitations (n=4)
Automatic	• Emotions surrounding PA (n=7)	• Survivors' emotions surrounding PA (n=3)

[a]COM-B: Capability, Opportunity, Motivation-Behavior.

[b]PA: physical activity.

[c]Not applicable or not available.

Domain: Capability

Overview

In the COM-B framework, capability refers to an individual's capacity to engage in a behavior and encompasses (1) physical capability (eg, physical strength, stamina, and symptom burden) and (2) psychological capability (eg, knowledge, understanding, and cognitive processes required to perform the behavior) [18]. In this study, capability captured factors influencing whether survivors were physically able and cognitively prepared to engage in PA, as well as whether medical team members felt equipped to assess and support PA participation.

Physical Capability

We identified 1 theme under physical capability: the *ability to engage in PA*. Every survivor mentioned their current engagement in activities and their ability to perform activities of daily living, such as household chores, which demonstrated their physical capability to stay active. One survivor mentioned:

XSL•FO
RenderX

> *I'm more of a person that takes on the chores of life. And that's where I get a lot of my physical activity in.*

For some, this baseline level of PA served as a foundation for further engagement in PA. At the same time, all survivor participants noted physical limitations, such as symptom burden and decreased fitness levels, that presented significant barriers to their ability to regularly engage in PA. One survivor noted:

> *There's been incidences during the whole time where I couldn't be very active… I've seen a lot of muscle atrophy set in. I've got very weak and things like that.*

If survivors were capable and self-driven to engage in PA, it facilitated medical team members' ability to promote PA. However, symptom-driven limitations also shaped medical team members' approaches, as they could only prescribe PA that aligned with what survivors' physical conditions allowed. One medical team member explained:

> *Every cancer patient is so different, and they have their own limitations and physical limitations.*

Psychological Capability

We identified 1 theme under psychological capability: the role that *knowledge and awareness* play in PA from survivors' and medical team members' perspectives. Many survivors mentioned their previous experience with PA, facilitating them with the cognitive tools to plan and engage in PA. Similarly, most medical team members emphasized that their knowledge of the benefits of PA and tips and tricks to help promote PA were key facilitators. On the other hand, some medical team members noted that a lack of knowledge regarding individual survivors' current PA levels and activity history hindered their ability to tailor and promote PA effectively. One medical team member stated:

> *People are super inexperienced on exercise, so the more information they have, the more successful they'll be with sticking with it.*

Domain: Opportunity

Overview

Within the COM-B framework, opportunity refers to external factors that enable or prompt behavior and includes (1) physical opportunity, such as environmental resources, access, time, and contextual constraints, and (2) social opportunity, including interpersonal influences, social norms, and support systems [18]. In this study, opportunity captured how features of survivors' physical environments and social contexts shaped their ability to engage in PA, as well as how these same conditions influenced medical team members' ability to promote PA.

Physical Opportunity

The first theme identified was the importance of *PA resources* for both medical team members and survivors regarding PA. Most survivors made note of access to public environments (eg, parks and gyms) along with tools, such as home PA equipment, online resources (eg, apps and YouTube), workout plans, and fitness trackers, as facilitators to engage in PA. Despite the availability of PA resources described by some participants, self-reported findings indicated that a majority of participants (7/12, 58%) did not meet recommended PA levels. Medical team members also identified access to online tools as a key facilitator for promoting PA among cancer survivors, noting that these resources expand accessible and flexible options for engaging in PA. One survivor mentioned:

> *I do think that like having things that you can do from the comfort of your own home is very important, because sometimes you don't feel like…looking like a normal person…on the flip side…gym accessibility would be awesome.*

However, several medical team participants also discussed that a *lack of access* to these resources is a barrier to promoting PA among survivors. Medical team members recognized that a lack of access to public parks or gyms, or tools and digital resources, reduces the opportunities for survivors to engage in regular PA. One medical team member noted:

> *…you want appropriate people doing the videos, not just someone,…that's a star on YouTube…Someone that's qualified is really important because our survivors are…receiving chemotherapy. They may not be in the best shape, and it's important to have the right…movements and things like that.*

Another medical team member stated:

> *You can only do so much. I mean, I can't get them a pair of running shoes. I can't get them a membership at the gym, or make it warmer outside to run, or whatever it is that would actually work for them.*

Another theme identified was *weather and time* regularly engaging in PA. Lack of time and poor weather conditions can be limiting factors in engaging in PA. One survivor stated:

> *I mean, it's harder in the winter time to walk and stuff like that…That's the hard time. It's hard to walk.*

Social Opportunity

Social support emerged as an integral facilitator for PA from both survivor and medical team member perspectives. For patients, access to local *classes, events, and support groups* created a sense of community and accountability, ultimately making engaging in PA more enjoyable. Many survivors also emphasized the supportive role that *their family, friends, pets, and medical team* provide in engaging in regular PA. Medical team members similarly highlighted the importance of supportive social environments to facilitate promoting PA among survivors. One medical team member mentioned:

> *Their families are the biggest cheerleaders…I have a lot of survivors…who are part of exercise groups or walking groups…and their goal is to get back to being with those people.*

Domain: Motivation

Overview

Within the COM-B framework, motivation refers to the internal processes that direct and sustain behavior and includes (1)

XSL•FO
RenderX

reflective motivation, such as conscious intentions, beliefs about outcomes, and evaluations of risk, and (2) automatic motivation, including emotional responses, habits, and routines [18]. In this study, motivation captured how survivors' beliefs, intentions, emotions, and habitual patterns influenced their engagement in PA, as well as how medical team members' perceptions of survivor motivation shaped their approaches to promoting PA.

Reflective Motivation

The first theme identified was the relationship between PA and *perceived health*. Both survivors and medical team members noted that the benefits of PA (eg, symptom relief and appearance) served as a motivating factor for engaging in and promoting PA. However, both groups also identified barriers to PA. Survivors reported that physical limitations, such as fatigue and discomfort, reduced their motivation to engage in regular PA. Medical team members echoed these concerns, noting that survivor-specific limitations related to treatment and disease often constrained their motivation to promote PA. One survivor noted:

> *Today, coughing and breathing is tough…I was thinking about getting on the bike…maybe trying for… 10 minutes or so. But usually, I'd like to go 30 minutes on the bike, if I can, but we'll see…if I get things moving and agitated then I start coughing a lot.*

Another survivor explained:

> *I feel like even in terms of physical appearance, again, there's so many things that change, but being able to still feel good about my body and like how I present myself is important to me.*

One medical team member stated:

> *It's extremely important. It's directly linked to their cardiovascular outcomes and their oncology outcomes…the physical activity is critical. I try to encourage all my survivors to be as physical as they can be.*

A second theme identified was the factors associated with PA-related resources. Several survivor participants reported that having access to at-home workouts and digital health tools was motivating. Many participants appreciated the flexibility that home workout options provided. This approach reduced logistical barriers, allowing them to engage in PA. At the same time, PA tracking tools provided motivating metrics, such as step count, that allowed survivors to understand and manage their PA. One survivor highlighted:

> *I have a lot of apps on my iPad about exercise and different things, so those kind of help a little bit. I will open up one and say, "Okay, let me see if I can get a 10-minute workout in or a 5-minute workout in."*

However, the positive influence of PA-related resources was tempered by *resource preferences and safety concerns*, which limited motivation to engage in PA. Few survivors were reluctant to use digital resources, such as online PA programs or tracking apps, due to discomfort with technology or a preference for more traditional methods. Moreover, some survivors expressed caution about the risks associated with PA, such as potential injury or the possibility of exacerbating symptoms, increasing their apprehension about engaging in PA. This focus on safety reflects survivors' conscious assessments of their physical limitations and the potential health consequences of PA. One survivor noted:

> *…during treatment, you're really afraid that you might hurt something or break something, that you're fragile and frail. It's pretty easy to say, "I'll wait till after treatment…"*

Several survivor participants also highlighted their *intentions* to engage in PA, which included a positive outlook toward PA and creating strategies to maintain accountability. Survivors noted that their motivation to engage in PA was influenced by the setting of goals and plans. Medical team members also observed this inclination in some survivors, stating that a survivor's motivation (eg, intentions, perceived beliefs, and consequences) to engage in PA serves as a key facilitator in promoting PA in survivors. One survivor explained:

> *Before that, I could bench press the 25 pounds. It's like, I wanna get back to that. I can do 20 now, but I can't do 25 yet.*

Automatic Motivation

The first theme identified was the positive and negative influence that a survivor's *emotions* have on their PA behaviors. On one hand, some survivors described experiencing positive feelings associated with engaging in PA, which was a facilitating factor. On the other hand, some survivors described negative feelings associated with engaging in PA, thus acting as a barrier. Aligned with this finding, medical team members also noted that the association of survivors' emotional relationship with PA influenced their motivation to promote PA. One medical team member noted:

> *…the biggest thing is their own motivation, like if they're not motivated to exercise or to do any physical activity, they're not gonna do it.*

One survivor mentioned:

> *When I'm achy and stiff, I'll go to a stretching type class, and I'll just feel better. I also feel when I'm kind of groggy, I'll go to some of these classes and just getting my blood flowing…my mood improves. I just feel better.*

The second theme identified was the *importance of a routine or habit* as a positive influence on motivation. Survivors noted that once PA became part of their daily routine, it required less conscious decision-making, which reinforced their ability to maintain regular PA. One survivor stated:

> *Just doing it as part of a routine. Say, this is my routine. I get up. I find the class that I want to go to at the center. I click on it and I go to it.*

Discussion

Principal Findings

In this qualitative study informed by the COM-B framework, findings indicate that although PA is widely valued, consistent

XSL•FO
RenderX

engagement remains limited. Although participants frequently described access to PA resources, less than half (5/12, 42%) met the recommended PA guidelines, suggesting a disconnect between available resources and sustained engagement. Rather than reflecting a lack of interest or intent, this pattern highlights the conditions under which PA requires adaptation, reassurance, and contextual support.

Capability-related factors included fluctuating symptoms, knowledge and awareness of PA, and readiness for PA, which shaped both survivors' ability to engage and medical team members' approaches to promoting PA. Opportunity-related influences reflected access to resources and supportive environments, where availability of PA options and social support facilitated engagement, while practical constraints, such as time and environmental barriers, limited feasibility. Motivation operated through both reflective processes, including weighing perceived health benefits alongside safety concerns and intentions to be active, and automatic processes, such as emotional responses and established routines.

The COM-B model conceptualizes behavior as arising from an interacting system of capability, opportunity, and motivation, with capability and opportunity also influencing motivation [18]. Participants' accounts illustrated these cross-domain relationships. Medical team members described family support and peer exercise groups as strengthening survivors' motivation by connecting PA with valued relationships and social roles, illustrating how social opportunity could support reflective motivation. Survivors described treatment-related physical vulnerability as contributing to perceived risk and fear-based avoidance, suggesting that limitations in physical capability could also weaken reflective motivation. These interactions varied across participants rather than reflecting a single, uniform pathway.

Descriptive survey findings complemented these themes by demonstrating variability in PA engagement, with fewer than half of participants (5/12, 42%) meeting recommended PA guidelines and only half (6/12, 50%) classified as active based on the RAPA. Many participants expressed motivation to engage in PA, suggesting that motivation alone may be insufficient to sustain engagement in some contexts, particularly in the absence of supportive conditions or resources. This aligns with prior evidence showing that adherence to PA guidelines among cancer survivors remains low, with previous studies reporting that only 20% meet aerobic PA guidelines [27]. Self-reported health status similarly varied, reinforcing heterogeneity in overall health. Together, these results underscore that PA engagement among survivors and promotion from medical team members represent complex, role-specific behaviors shaped by interacting influences.

Comparison With Prior Literature

These patterns are consistent with prior research suggesting that PA engagement across cancer survivorship is shaped by interrelated conditions that influence feasibility and support. Prior survivorship studies have similarly emphasized that engagement in health-promoting behaviors following a cancer diagnosis is influenced by dynamic and context-dependent situations. For example, qualitative reviews of survivor perspectives describe PA engagement as contingent on a range of factors, including symptoms, environmental conditions, and the variability of reassurance and guidance from health care professionals [28]. With most cancer survivors not achieving recommended PA levels, this reflects ongoing challenges related to clinical practice, resource availability, and behavior change across both survivor and medical team member contexts [19].

Several specific themes identified in our study closely align with those reported in qualitative syntheses of cancer survivors' experiences with PA. Among adults (n=12) with cancer seen in the cardio-oncology clinic, treatment-related side effects, particularly fatigue, emerged as major barriers to PA participation. This finding is consistent with prior research identifying treatment-related symptoms, including fatigue, as predominant barriers to engaging in PA [28,29]. Similarly, environmental concerns were identified as barriers, aligning with findings from qualitative syntheses and systematic reviews, including weather conditions, lack of suitable or adapted PA environments, and financial constraints [28,30]. Psychological concerns were prominent, with fears of worsening symptoms or disease and reduced self-esteem or motivation shaping survivors' willingness to engage in PA, consistent with prior literature describing emotional and cognitive barriers to PA [30]. Support-related factors also emerged as key facilitators, as participants emphasized the value of encouragement, structure, guidance from health care professionals, and group-based settings in supporting PA engagement, alongside frustration when such support was absent, which has been similarly reported in qualitative studies and mixed-method reviews of cancer survivors [28,30].

Prior literature has documented that clinicians' promotion of PA in cancer care is shaped by multiple professional and contextual considerations, rather than by endorsement of PA alone. Qualitative studies of oncology care medical team members indicate that, although clinicians widely recognize the benefits of PA, promotion is influenced by factors such as role clarity, confidence and training, time constraints, availability of referral resources, and concerns related to survivor symptoms and safety [31]. Complementing these findings, Schmitz et al [19] indicated that clinicians often lack clarity regarding their role in assessing, advising, and referring patients for PA, alongside system-level barriers such as limited referral pathways, insufficient resources, and challenges integrating PA promotion into routine care. Importantly, this work emphasizes that oncology clinicians are not expected to provide detailed PA prescriptions but instead play a key role in encouraging PA and directing survivors to appropriate professionals and programs (eg, cancer rehabilitation services, exercise physiologists, and community-based programs), highlighting the need for accessible referral pathways within cancer care [19].

Collectively, prior work shows that both survivor engagement and clinician promotion of PA are shaped by aligned, or misaligned, conditions within the clinical encounter. This is consistent with a recent qualitative study applying the COM-B framework and Theoretical Domains Framework to examine PA integration in breast cancer care from both patient and health care professional perspectives, which similarly found that

XSL•FO
RenderX

patient-level barriers (eg, misconceptions and fatigue) and professional-level barriers (eg, need for clearer protocols and training) shape integration of PA into routine oncology care [32]. By integrating survivor and medical team member perspectives, this study shows that PA engagement is not simply an individual survivor behavior but a negotiated clinical behavior that is jointly experienced and navigated throughout the survivorship journey.

Implications for Intervention Design

Findings from this study indicate that promoting PA in cancer survivorship may benefit from flexible, responsive approaches rather than relying on standardized or one-time recommendations, as both survivors and medical team members described PA engagement as fluctuating over time. These findings highlight the potential value of interventions that facilitate ongoing communication between survivors and medical team members, allowing recommendations to be tailored and revisited as clinical circumstances change. Such approaches may be particularly relevant in clinical contexts where PA has the potential to influence patient outcomes, including treatment tolerance, recovery, and rehabilitation.

For survivors who face symptom burden, safety concerns, or environmental barriers, future intervention development could examine adapted home-based programs that provide feasible alternatives to traditional PA settings. This approach is supported by evidence that home-based PA interventions can provide accessible alternatives for cancer survivors facing feasibility barriers and that interventions incorporating frequent counseling are associated with larger improvements in fatigue, suggesting the importance of pairing flexible PA resources with ongoing support that can be revisited and adapted [33]. These results indicate that interventions should support clinicians in translating general PA guidance into feasible, individualized recommendations that account for survivors' potential barriers. Consistent with this approach, evidence from implementation research suggests that strategies incorporating individualized and interactive elements, such as tailored counseling and active survivor-medical team member communication, are more effective at increasing PA uptake during and after cancer treatment than one-time or generic recommendations [34]. However, the feasibility of delivering individualized counseling in routine oncology care may be limited by time constraints, staff availability, and variability in referral resources. To address this gap, future interventions should examine low-burden implementation strategies that preserve individualized support while fitting within existing clinical workflows.

Access to resources also emerged as an important factor shaping PA engagement. Survivors frequently described community spaces, home-based options, and digital tools as facilitating adaptation and continuity of PA, while medical team members noted variability in survivors' access to such resources as a consideration during promotion. Both survivors and medical team members described the supportive role that survivors' social environments, including family, friends, and their medical team, can play in facilitating physical engagement. This aligns with qualitative evidence demonstrating that social support functions as a key facilitator of PA among cancer survivors, enabling both adoption and maintenance of PA through mechanisms such as encouragement, accountability, and companionship [35]. In this context, clinicians may play a supportive role by helping survivors identify and make use of feasible, accessible resources, both material and social, that align with their circumstances. Framing PA promotion as an adaptive clinical process that incorporates individualized guidance and practical support may enable more sustained engagement without increasing burden within routine care. Future interventions should examine how resource-matching strategies, including digital PA tools, wearable-supported self-monitoring, group-based remote programs, and curated community referral options, can be tailored to survivors' needs and integrated into survivorship care without increasing burden on survivors or medical team members.

The findings also point to important considerations for the design of PA interventions in cancer survivorship. Interventions that focus exclusively on modifying survivor behavior may be limited if they do not also account for the clinical interactions through which PA is discussed and supported. The results suggest that intervention approaches should allow for flexibility, recognizing that survivor capacity and circumstances often change over time rather than remaining stable. Given that participants described clinical encounters as key settings where PA is discussed and negotiated, future intervention development should examine approaches that account for the barriers identified in our study, including symptom burden, safety concerns, competing demands, environmental constraints, social support needs from family and medical team members, and variability in clinical guidance. Population-based evidence indicates that fewer than half of cancer survivors report receiving PA counseling from a medical team member, yet those who do receive counseling are more likely to engage in higher levels of PA, highlighting the survivor-medical team member interaction as a meaningful opportunity for intervention [36].

Intervention approaches may also benefit from attending to how counseling is translated into action, including connection to feasible PA resources and tools, such as home-based options, community spaces, and digital supports. Additionally, leveraging survivors' existing social environments may further support engagement and align with evidence that social support facilitates adoption and maintenance of PA among cancer survivors [35]. Overall, interventions that align survivor experiences, clinician guidance, and available supports are more likely to be effective than those targeting any single element in isolation.

Implications for Future Research

Future research should continue to examine PA engagement and promotion as dynamic, context-dependent processes that unfold throughout survivorship. Longitudinal qualitative or mixed-methods studies may be particularly valuable for understanding how survivors' and clinicians' assessments of feasibility, support needs, and priorities shift over time in response to changing symptoms, treatment phases, and life circumstances. Studies that intentionally integrate survivor and medical team perspectives within shared care contexts could further illuminate how clinical interactions shape PA

XSL•FO
RenderX

engagement. Larger-scale qualitative or mixed-methods studies, including those conducted across multiple clinical settings or health systems, would be useful for assessing the transferability and generalizability of these findings to more diverse survivorship populations. In addition, research that evaluates strategies designed to support survivor-medical team member PA counseling within routine care may help strengthen translation of evidence into practice. Greater attention to variability across care settings, resource availability, and survivor populations will also be important for informing equitable and scalable approaches to PA promotion in survivorship.

Implications for Digital Health and Referral Pathways

Given the relevance of digital health approaches to supporting PA in survivorship, several implications for digital tools and referral pathways emerged from this study. Both survivors and medical team members described interest in accessible referral pathways connecting survivors to appropriate professionals and programs (eg, cancer rehabilitation services, exercise physiologists, and community-based programs) [19]. Recent feasibility studies offering PA interventions provide examples of such approaches, including brief PA communication and clinician referral pathways involving exercise professionals, as well as electronic triage tools paired with brief counseling and referral to PA or rehabilitation services [37,38]. For survivors facing symptom burden, safety concerns, or environmental barriers, digital and technology-supported PA resources, including group-based videoconference (tele-exercise) programs, wearable technology–based interventions, and app-based tracking tools, offer feasible and scalable alternatives to traditional PA settings [39,40]. Future interventions should examine how resource-matching strategies, including digital PA tools, wearable-supported self-monitoring, group-based tele-exercise programs, and curated community referral options, can be tailored to survivors' needs and integrated into survivorship care without increasing burden on survivors or medical team members.

Strengths and Limitations

This study has several strengths. First, it incorporated both survivor and medical team perspectives from a single setting, providing a more comprehensive view of PA engagement and promotion in cancer survivorship than studies focused on a single group. Examining these perspectives together enabled the identification of shared challenges and complementary insights within clinical interactions. Second, the focused, theory-informed interview structure supported in-depth exploration of PA engagement and promotion within a cardio-oncology setting, where PA may be shaped by not only cancer survivorship concerns but also cardiovascular risk, treatment-related symptoms, and perceptions of safety. The use of the COM-B framework to guide data collection and analysis supported systematic examination of capability, opportunity, and motivation as interacting influences on PA engagement and promotion. Although the interview guide also addressed symptom tracking, those findings are reported separately in another manuscript. Additionally, descriptive survey data provided useful context regarding participants' self-reported PA behaviors and cardiovascular health characteristics, although these data were not intended to validate or explain individual qualitative themes. Lastly, conducting interviews remotely helped reduce logistical and health-related barriers to participation, potentially increasing accessibility for individuals who may have had difficulty attending in-person interviews.

The interpretation of these findings should consider several limitations. First, the study relied on self-reported perceptions and experiences, which reflect participants' accounts of PA engagement and promotion rather than directly observed behaviors or objectively measured outcomes. PA measures were self-reported and descriptive in nature, which may be subject to recall or reporting bias. The cardiovascular health composite reported here reflects a modified, 6-domain adaptation of the American Heart Association's LE8 framework, excluding blood glucose (unavailable for all participants) and blood lipids (available for only 4 of 12 participants), and should not be interpreted as equivalent to the validated 8-domain LE8 score. A sensitivity analysis among the subset of participants with lipid data indicated that including this domain did not shift composite scores uniformly (mean change=1.0 points; range −3.1 to +4.4), suggesting that the original variable-denominator approach could have introduced variability tied to data availability rather than true cardiovascular health status. All interviews were conducted remotely, which may have shaped the nature of participant responses and introduced selection bias favoring individuals with greater comfort using technology or access to remote communication platforms. The modest gift card incentive (US $25) provided to survivor participants, while intended to offset time and participation burden, may have also influenced recruitment speed or introduced a degree of selection bias toward individuals more motivated by the incentive or with greater flexibility to participate. This potential influence was not formally assessed. The sample size was relatively small and, while appropriate for qualitative inquiry, may not capture the full range of perspectives across cancer types, treatment stages, or sociodemographic backgrounds. This diagnostic, treatment, and professional heterogeneity strengthened our ability to identify behavioral themes that generalize across clinical contexts, though it also precluded stratified analysis by diagnosis, treatment type, or professional role. Theme prevalence may vary across these subgroups. Survivor participants tended to have higher levels of income and educational attainment and were predominantly White and non-Hispanic, which may limit the transferability of the findings to more socioeconomically and racially or ethnically diverse survivorship populations. Barriers, such as weather, time, and gym access, reflect this relatively affluent sample. Structural social determinants of health barriers common in marginalized populations (eg, transportation, neighborhood safety, cost, and job flexibility) were unlikely to be captured and warrant a dedicated study. In addition, participants were recruited from a single clinical setting, which may limit transferability to other survivorship care environments with different resources, workflows, or practice norms. Relatedly, because participants were recruited from a specialized cardio-oncology clinic rather than a general oncology setting, both survivors and medical team members may have had greater baseline awareness of and engagement with lifestyle interventions than would be expected

XSL•FO
RenderX

in general oncology populations, which may have influenced the prominence of themes related to knowledge, motivation, and support for PA. Finally, although the COM-B framework provided a useful structure for organizing data collection and analysis, its application may have highlighted certain behavioral influences while de-emphasizing others. Alternative approaches could have yielded different interpretive themes.

Conclusion

This qualitative study explored how PA engagement and promotion are experienced and navigated in cancer survivorship from both survivor and medical team perspectives. Across interviews, PA was widely recognized as valuable, yet participation and promotion were influenced by changing symptoms, competing priorities, perceived safety, available resources, and the circumstances of clinical encounters. Patterns of PA were not static; instead, engagement reflected ongoing adjustment to health status and life context rather than simple adherence to recommendations. These findings highlight PA as a behavior that evolves over time within survivorship rather than one that can be adequately supported through singular or standardized guidance.

By applying the COM-B framework, this study offers insight into how capability, opportunity, and motivation interact across survivor and clinical contexts to shape PA engagement and promotion. The findings underscore the importance of clinical conversations, social and material supports, and adaptive approaches that accommodate fluctuation rather than expecting consistency. Attention to how PA is discussed, supported, and resourced may help bridge the gap between evidence and sustained engagement in daily life. Collectively, these insights contribute to a growing emphasis on flexible, context-aware strategies to support PA throughout the survivorship trajectory and may inform more responsive approaches to clinical practice, intervention design, and future research.

Acknowledgments

We thank the participants who shared their time and insights and the American Heart Association for allowing us to use the Life's Essential 8 approach to evaluate cardiovascular health. The authors would also like to thank the Lifestyles Lab team members for their contributions to data management and administrative coordination.

During the preparation of this manuscript, the authors used CoPilot (Microsoft 365) as a writing support tool to proofread the text for grammar. All content was written, critically reviewed, and edited by the authors, who take full responsibility for the final manuscript.

Funding

This study was supported by discretionary research funds provided to RMB through the Ohio State University Comprehensive Cancer Center (OSUCCC).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

MKT and RMB conceived and designed the study, conducted interviews, and coordinated the primary interpretation of findings. MKT led manuscript writing. MKT and RMB oversaw data collection and conducted the final analysis. MKT, RS, SMSG, and VP coordinated data management, performed preliminary analyses, and assisted with manuscript writing. All authors contributed to study design, interpretation of the results, and critical revision of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Medical team member and survivor participant interview guides.
[PDF File, 305 KB - cancer_v12i1e104332_app1.pdf]

Multimedia Appendix 2
Analytic matrix structure template.
[DOCX File, 19 KB - cancer_v12i1e104332_app2.docx]

Multimedia Appendix 3
Supplementary survey results.
[DOCX File, 19 KB - cancer_v12i1e104332_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 4
Representative quotes from medical team members and survivor participants.
[DOCX File, 44 KB - cancer_v12i1e104332_app4.docx]

References

1. Bray F, Laversanne M, Weiderpass E, Soerjomataram I. The ever-increasing importance of cancer as a leading cause of premature death worldwide. Cancer 2021 Aug 15;127(16):3029-3030. [doi: 10.1002/cncr.33587] [Medline: 34086348]
2. Ng HS, Meng R, Marin TS, et al. Cardiovascular mortality in people with cancer compared to the general population: a systematic review and meta-analysis. Cancer Med 2024 Aug;13(15):e70057. [doi: 10.1002/cam4.70057] [Medline: 39096123]
3. Wilcox NS, Amit U, Reibel JB, Berlin E, Howell K, Ky B. Cardiovascular disease and cancer: shared risk factors and mechanisms. Nat Rev Cardiol 2024 Sep;21(9):617-631. [doi: 10.1038/s41569-024-01017-x] [Medline: 38600368]
4. Sturgeon KM, Deng L, Bluethmann SM, et al. A population-based study of cardiovascular disease mortality risk in US cancer patients. Eur Heart J 2019 Dec 21;40(48):3889-3897. [doi: 10.1093/eurheartj/ehz766] [Medline: 31761945]
5. Zaorsky NG, Churilla TM, Egleston BL, et al. Causes of death among cancer patients. Ann Oncol 2017 Feb 1;28(2):400-407. [doi: 10.1093/annonc/mdw604] [Medline: 27831506]
6. Jain D, Russell RR, Schwartz RG, Panjrath GS, Aronow W. Cardiac complications of cancer therapy: pathophysiology, identification, prevention, treatment, and future directions. Curr Cardiol Rep 2017 May;19(5):36. [doi: 10.1007/s11886-017-0846-x] [Medline: 28374177]
7. Armenian SH, Xu L, Ky B, et al. Cardiovascular disease among survivors of adult-onset cancer: a community-based retrospective cohort study. JCO 2016 Apr 1;34(10):1122-1130. [doi: 10.1200/JCO.2015.64.0409]
8. Rock CL, Thomson CA, Sullivan KR, et al. American Cancer Society nutrition and physical activity guideline for cancer survivors. CA Cancer J Clin 2022 May;72(3):230-262. [doi: 10.3322/caac.21719] [Medline: 35294043]
9. Schmitz KH, Holtzman J, Courneya KS, Mâsse LC, Duval S, Kane R. Controlled physical activity trials in cancer survivors: a systematic review and meta-analysis. Cancer Epidemiol Biomarkers Prev 2005 Jul;14(7):1588-1595. [doi: 10.1158/1055-9965.EPI-04-0703] [Medline: 16030088]
10. Mora S, Cook N, Buring JE, Ridker PM, Lee IM. Physical activity and reduced risk of cardiovascular events: potential mediating mechanisms. Circulation 2007 Nov 6;116(19):2110-2118. [doi: 10.1161/CIRCULATIONAHA.107.729939] [Medline: 17967770]
11. Campbell KL, Winters-Stone KM, Wiskemann J, et al. Exercise guidelines for cancer survivors: consensus statement from international multidisciplinary roundtable. Med Sci Sports Exerc 2019 Nov;51(11):2375-2390. [doi: 10.1249/MSS.0000000000002116] [Medline: 31626055]
12. Clinton SK, Giovannucci EL, Hursting SD. The World Cancer Research Fund/American Institute for Cancer Research third expert report on diet, nutrition, physical activity, and cancer: impact and future directions. J Nutr 2020 Apr 1;150(4):663-671. [doi: 10.1093/jn/nxz268] [Medline: 31758189]
13. Baughman C, Norman K, Mukamal K. Adherence to American Cancer Society nutrition and physical activity guidelines among cancer survivors. JAMA Oncol 2024 Jun 1;10(6):789-792. [doi: 10.1001/jamaoncol.2024.0470] [Medline: 38635238]
14. Haussmann A, Ungar N, Gabrian M, et al. Are healthcare professionals being left in the lurch? The role of structural barriers and information resources to promote physical activity to cancer patients. Support Care Cancer 2018 Dec;26(12):4087-4096. [doi: 10.1007/s00520-018-4279-x] [Medline: 29934683]
15. Keogh JWL, Olsen A, Climstein M, Sargeant S, Jones L. Benefits and barriers of cancer practitioners discussing physical activity with their cancer patients. J Canc Educ 2017 Mar;32(1):11-15. [doi: 10.1007/s13187-015-0893-1]
16. Benzo RM, Tetrick MK, Singh R, et al. Cardiovascular symptom tracking among patients with cancer in cardio-oncology care: qualitative study using the capability, opportunity, motivation-behavior (COM-B) framework. JMIR Cancer 2026 Jul 23;12:e100279. [doi: 10.2196/100279] [Medline: 42492086]
17. Lyon AR, López-Fernández T, Couch LS, et al. 2022 ESC Guidelines on cardio-oncology developed in collaboration with the European Hematology Association (EHA), the European Society for Therapeutic Radiology and Oncology (ESTRO) and the International Cardio-Oncology Society (IC-OS). Eur Heart J 2022 Nov 1;43(41):4229-4361. [doi: 10.1093/eurheartj/ehac244] [Medline: 36017568]
18. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
19. Schmitz KH, Campbell AM, Stuiver MM, et al. Exercise is medicine in oncology: engaging clinicians to help patients move through cancer. CA A Cancer J Clinicians 2019 Nov;69(6):468-484. [doi: 10.3322/caac.21579]
20. Topolski TD, LoGerfo J, Patrick DL, Williams B, Walwick J, Patrick MB. The rapid assessment of physical activity (RAPA) among older adults. Prev Chronic Dis 2006 Oct;3(4):A118. [Medline: 16978493]
21. PRAPARE (protocol for responding to and assessing patients' assets, risks, and experiences). URL: https://prapare.org [accessed 2026-09-01]
22. Lloyd-Jones DM, Allen NB, Anderson CAM, et al. Life's essential 8: updating and enhancing the American Heart Association's construct of cardiovascular health: a presidential advisory from the American Heart Association. Circulation 2022 Aug 2;146(5):e18-e43. [doi: 10.1161/CIR.0000000000001078] [Medline: 35766027]

XSL•FO
RenderX

23. Cerwinske LA, Rasmussen HE, Lipson S, Volgman AS, Tangney CC. Evaluation of a dietary screener: the Mediterranean Eating Pattern for Americans tool. J Hum Nutr Diet 2017 Oct;30(5):596-603. [doi: 10.1111/jhn.12451] [Medline: 28168764]
24. Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. Psychiatry Res 2019 Oct;280:112516. [doi: 10.1016/j.psychres.2019.112516] [Medline: 31437661]
25. St George SM, Harkness AR, Rodriguez-Diaz CE, Weinstein ER, Pavia V, Hamilton AB. Applying rapid qualitative analysis for health equity: lessons learned using "EARS" with Latino communities. Int J Qual Methods 2023;22:16094069231164938. [doi: 10.1177/16094069231164938] [Medline: 38463016]
26. Kowalski CP, Nevedal AL, Finley EP, et al. Planning for and Assessing Rigor in Rapid Qualitative Analysis (PARRQA): a consensus-based framework for designing, conducting, and reporting. Implementation Sci 2024 Oct 11;19(1):71. [doi: 10.1186/s13012-024-01397-1]
27. Coletta AM, Marquez G, Thomas P, et al. Clinical factors associated with adherence to aerobic and resistance physical activity guidelines among cancer prevention patients and survivors. PLoS ONE 2019;14(8):e0220814. [doi: 10.1371/journal.pone.0220814] [Medline: 31369653]
28. Finch A, Benham A. Patient attitudes and experiences towards exercise during oncological treatment. A qualitative systematic review. Support Care Cancer 2024 Jul 12;32(8):509. [doi: 10.1007/s00520-024-08649-2] [Medline: 38992238]
29. Frikkel J, Götte M, Beckmann M, et al. Fatigue, barriers to physical activity and predictors for motivation to exercise in advanced Cancer patients. BMC Palliat Care 2020 Mar 31;19(1):43. [doi: 10.1186/s12904-020-00542-z] [Medline: 32234027]
30. Clifford BK, Mizrahi D, Sandler CX, et al. Barriers and facilitators of exercise experienced by cancer survivors: a mixed methods systematic review. Support Care Cancer 2018 Mar;26(3):685-700. [doi: 10.1007/s00520-017-3964-5] [Medline: 29185105]
31. Kumar G, Chaudhary P, Ganti AK, Kim J, Smith LM, Su D. Perceived barriers, facilitators, and needs related to promoting physical activity in cancer care: qualitative insights from oncology care providers. Cancers (Basel) 2025 Jul 9;17(14):2281. [doi: 10.3390/cancers17142281] [Medline: 40723164]
32. Piazzon N, Carrouel F, Gérard A, Ringot A, Cortet M, Verot E. Barriers and enablers to integrating physical activity in breast cancer care: a qualitative study using the TDF and COM-B model. Support Care Cancer 2026 Mar 2;34(3):261. [doi: 10.1007/s00520-026-10469-5] [Medline: 41770281]
33. Huizinga F, Westerink NDL, Berendsen AJ, et al. Home-based physical activity to alleviate fatigue in cancer survivors: a systematic review and meta-analysis. Med Sci Sports Exerc 2021 Dec 1;53(12):2661-2674. [doi: 10.1249/MSS.0000000000002735] [Medline: 34649267]
34. IJsbrandy C, Ottevanger PB, Tsekou Diogeni M, Gerritsen WR, van Harten WH, Hermens R. Review: effectiveness of implementation strategies to increase physical activity uptake during and after cancer treatment. Crit Rev Oncol Hematol 2018 Feb;122:157-163. [doi: 10.1016/j.critrevonc.2017.09.005] [Medline: 29458784]
35. McDonough MH, Beselt LJ, Kronlund LJ, et al. Social support and physical activity for cancer survivors: a qualitative review and meta-study. J Cancer Surviv 2021 Oct;15(5):713-728. [doi: 10.1007/s11764-020-00963-y] [Medline: 33128705]
36. Tarasenko YN, Miller EA, Chen C, Schoenberg NE. Physical activity levels and counseling by health care providers in cancer survivors. Prev Med 2017 Jun;99:211-217. [doi: 10.1016/j.ypmed.2017.01.010] [Medline: 28131780]
37. Caperchione CM, English M, Sharp P, et al. Exploring the practicality and acceptability a brief exercise communication and clinician referral pathway in cancer care: a feasibility study. BMC Health Serv Res 2023 Sep 22;23(1):1023. [doi: 10.1186/s12913-023-10003-x] [Medline: 37740170]
38. Schmitz KH, Chongaway A, Saeed A, et al. An initiative to implement a triage and referral system to make exercise and rehabilitation referrals standard of care in oncology. Support Care Cancer 2024 Apr 2;32(4):259. [doi: 10.1007/s00520-024-08457-8] [Medline: 38561546]
39. Lynch BM, Nguyen NH, Moore MM, et al. A randomized controlled trial of a wearable technology-based intervention for increasing moderate to vigorous physical activity and reducing sedentary behavior in breast cancer survivors: the ACTIVATE trial. Cancer 2019 Aug 15;125(16):2846-2855. [doi: 10.1002/cncr.32143] [Medline: 31012970]
40. Leach HJ, Crisafio ME, Howell MJ, Nicklawsky A, Marker RJ. A group-based, videoconference-delivered physical activity program for cancer survivors. Transl J ACSM 2023 Apr;8(2). [doi: 10.1249/TJX.0000000000000221]

Abbreviations

ACS: American Cancer Society
AICR: American Institute for Cancer Research
COM-B: Capability, Opportunity, Motivation-Behavior
CVD: cardiovascular disease
LE8: Life's Essential 8
MEPA: Modified Mediterranean Eating Pattern for Americans
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PRAPARE: Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences

XSL•FO
RenderX

RAPA: Rapid Assessment of Physical Activity
RQA: rapid qualitative analysis

Edited by M Balcarras; submitted 10.Jun.2026; peer-reviewed by F Carrouel, K Jermsittiparsert; revised version received 24.Aug.2026; accepted 25.Aug.2026; published 11.Sep.2026.

Please cite as:
Tetrick MK, Addison D, Singh R, St. George SM, Pavia V, Gogineni A, Washington P, Ghazi SM, Preston O, Fisher JL, Benzo RM Assessment of Physical Activity Engagement and Promotion Among Cancer Survivors in Cardio-Oncology Care: Qualitative Study JMIR Cancer 2026;12:e104332
URL: https://cancer.jmir.org/2026/1/e104332
doi:10.2196/104332

© Macy K Tetrick, Daniel Addison, Rujul Singh, Sara M St. George, Vanina Pavia, Anvitha Gogineni, Peter Washington, Sanam M Ghazi, Olivia Preston, James L Fisher, Roberto M Benzo. Originally published in JMIR Cancer (https://cancer.jmir.org), 11.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Quality, Empathy, and Readability of AI Chatbot Responses to the Survivorship Needs of Adolescents and Young Adults With Melanoma: Evaluation Study

Jordan Lily Jafarnia[1*], BA; Priscilla Lynne Haff[1*], BS; Reece Philip Moore[1], BS; Alyssa Leigh Osheim[1], BBA; Katherine McKenna Riley[1], BA; Sabrina Zheng[2], BS; Michael Roth[3], MD; Madeleine Hines Salge[4], MD; Kelly Carter Nelson[4], MD

[1]UTHealth Houston, McGovern Medical School, 6431 Fannin St, Houston, TX, United States

[2]Paul L. Foster School of Medicine at Texas Tech University Health Sciences Center El Paso, El Paso, TX, United States

[3]Division of Pediatrics, University of Texas MD Anderson Cancer Center, Houston, TX, United States

[4]Department of Dermatology, Division of Internal Medicine, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX, United States

[*]these authors contributed equally

Corresponding Author:
Kelly Carter Nelson, MD
Department of Dermatology, Division of Internal Medicine, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX, United States

Abstract

Background: Melanoma, a highly aggressive form of skin cancer, is the second most common type of cancer for adolescent and young adult (AYA, ages 15-39 years) patients. AYA patients with melanoma may turn to internet sources, especially artificial intelligence (AI) chatbots, to manage uncertainty about prognosis and treatment.

Objective: This study aims to evaluate the quality, empathy, and readability of responses generated by leading AI chatbots when addressing the top unmet needs of AYA patients with melanoma receiving treatment.

Methods: Our research team recently surveyed 152 AYA patients with melanoma using the Needs Assessment Service Bridge, a validated instrument that assesses psychosocial needs for AYA patients with cancer. The survey identified the top 5 needs for advanced AYA patients with melanoma receiving treatment. Each need was reframed into a question and brief clinical history, then entered into each chatbot by 5 individuals who cleared their prequestion and postquestion history. Chatbot responses were evaluated to assess information quality (Global Quality Score [GQS] and DISCERN), accessibility and readability (GQS, Flesch Kincaid Grade Level, Flesch Reading Ease), and perceived empathy (Perceived Empathy of Technology Scale [PETS], including domains of Emotional Responsiveness [PETS-ER], Understanding and Trust [PETS-UT]).

Results: Across 75 chatbot responses, ChatGPT achieved the highest average quality (mean GQS 4.42, SD 0.32; mean DISCERN 3.24, SD 0.31) and empathy (mean PETS-ER 5.35, SD 1.85; mean PETS-UT 6.36, SD 1.83), though with greater variability. Copilot produced the lowest quality and empathy scores, while Gemini responses were consistently midrange. PETS-UT exceeded PETS-ER across all models, suggesting stronger cognitive empathy than emotional responsiveness. Readability analysis showed outputs exceeded the average US reading level (mean Flesch Kincaid Grade Level 11.82, SD 1.44; mean FRE 38.60, SD 9.00), limiting accessibility. The most readable responses were found in question 2, which also scored higher in quality and empathy, whereas questions 4 and 5 produced the most complex, difficult-to-read responses corresponding with lower quality and empathy ratings.

Conclusions: AI chatbots can provide moderately accurate and supportive responses to needs of AYA patients with melanoma, but outputs are inconsistent, written above the recommended reading level for health information, and limited in empathy. Question framing strongly influenced chatbot performance, with more emotional prompts drawing greater empathy, and readability aligning with both quality and empathy. Chatbot use in this population should remain adjunctive, with further research needed to standardize quality, improve readability, and enhance empathetic communication.

(JMIR Cancer 2026;12:e84234) doi:10.2196/84234

KEYWORDS

melanoma; young adult; artificial intelligence; natural language processing; empathy; readability; adolescent

XSL•FO
RenderX

Introduction

Melanoma, a malignant tumor of melanocytes, is an aggressive form of skin cancer due to its high metastatic potential and associated mortality. Despite being most frequently diagnosed in older adults, cutaneous melanoma accounts for approximately 7% of newly diagnosed cancers among adolescents and young adults (AYAs, age 15 - 39 years) [1]. A melanoma diagnosis and subsequent treatment can be highly disruptive for AYA patients, with implications for physical health, emotional well-being, and social function. The severity of the disease, coupled with the uncertainty of treatment outcomes, often introduces feelings of fear and isolation [2]. In seeking clarity in this uncertainty, patients may turn to the internet and social media for answers, education, and support. When seeking answers through internet interfaces, patients are likely to leverage artificial intelligence (AI)–driven chatbots, such as Chat GPT, Microsoft Copilot, and Google Gemini [3].

AI chatbots use natural language processing to generate conversational responses to user queries and have been shown to provide health information in an accessible and interactive manner [4]. Their availability, objectivity, and capacity to adapt personalized responses to diverse communication styles make them appealing to patients navigating difficult diagnoses [5]. While some studies have evaluated the most frequently asked chatbot questions regarding specific forms of cancer (ie, breast, leukemia, prostate), no published research to our knowledge has systematically examined their ability to address AYAs' specific concerns regarding melanoma treatment and survivorship [6].

A recent survey study of AYAs with melanoma identified several unmet needs for patients in different phases of cancer care using the validated Needs Assessment Service Bridge (NA-SB) questionnaire [7]. This survey represents a separate study and included 152 AYA respondents, 20 of whom were actively receiving systemic treatment at the time of participation; detailed methods and results are reported elsewhere [7]. For AYA patients with melanoma currently undergoing systemic treatment, the top 5 unmet needs included the following: (1) what happens after treatment, (2) fear of cancer recurrence, (3) long-term hormonal side effects, (4) long-term systemic side effects, and (5) genetic implications on diagnosis and treatment [7]. In this study, survey findings were used solely to identify patient-prioritized unmet needs for chatbot evaluation.

Building on these empirically derived patient priorities, this study anchors chatbot prompts in validated unmet needs rather than search-derived queries; evaluates perceived empathy using the Perceived Empathy of Technology Scale (PETS), a recently validated instrument not previously applied in oncology chatbot research; and examines how readability, empathy, and information quality interact. Together, this novel approach enables a patient-centered and multidimensional evaluation of chatbot suitability for AYA melanoma care, further building on previous chatbot-oncology research. With these established unmet needs and the increasing use of AI chatbots for AYA patient support, we evaluated the unique response of 3 chatbots to patient-centered questions to determine if these resources are safe and valid for use by AYA patients with melanoma.

Methods

Need Selection and Question Creation

The top 5 most frequently reported needs for AYA patients with melanoma receiving systemic treatment were identified through the NA-SB survey. Once identified, each NA-SB prompt was slightly modified into a question format with a standardized and appropriate brief clinical history (Textbox 1). All questions shared a common stem that identified the patient as an AYA with metastatic melanoma, with tailored follow-up questions added to reflect the specific need category.

Textbox 1. Questions input into each chatbot, based on the 5 most frequently reported needs.

Questions from the perspective of patients currently on treatment:

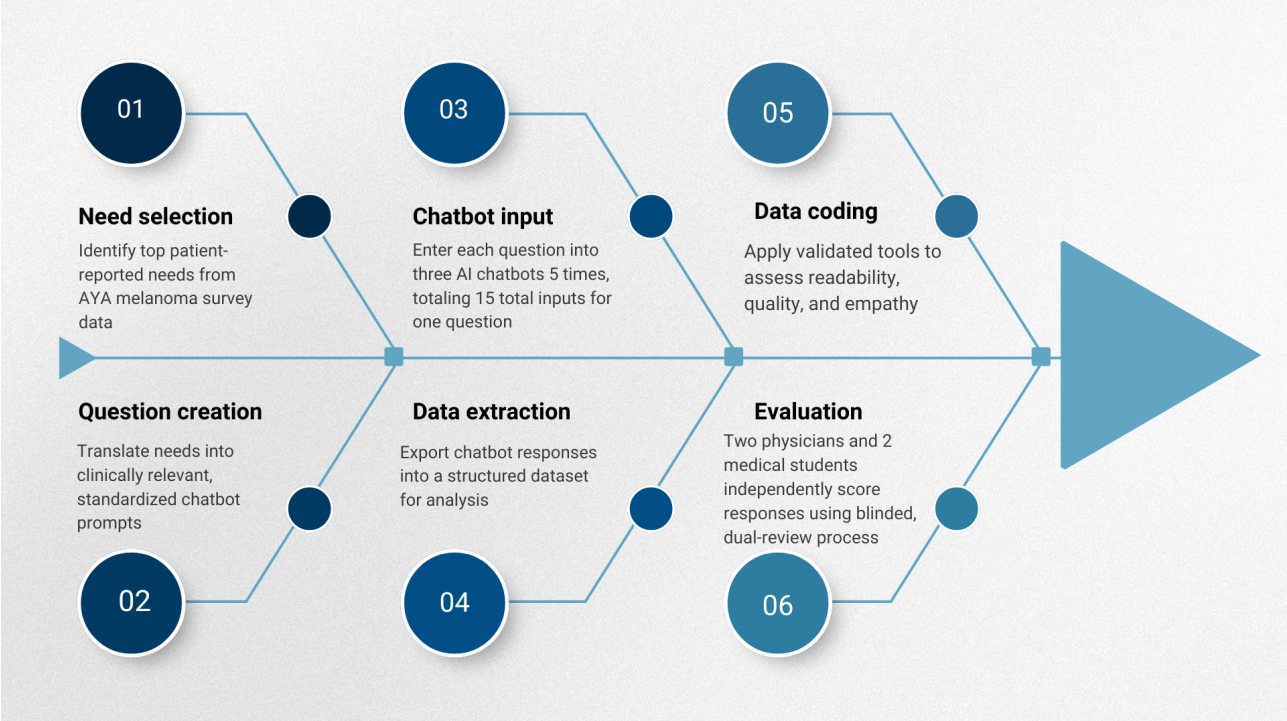
1. I am an adolescent/young adult patient with metastatic melanoma and am currently receiving treatment. I want more information about what will happen when treatment finishes.
2. I am an adolescent/young adult patient with metastatic melanoma and am currently receiving treatment. In the last month, I have become more worried about my cancer spreading. What do I need to know and what steps can I take to cope with this fear?
3. I am an adolescent/young adult patient with metastatic melanoma and am currently receiving treatment. How will my treatment affect long-term hormone changes?
4. I am an adolescent/young adult patient with metastatic melanoma and am currently receiving treatment. I want more information about how my genetics may or may not have impacted my diagnosis and treatment.
5. I am an adolescent/young adult patient with metastatic melanoma and am currently receiving treatment. I want more information about the long-term side effects of treatment.

Study Design

This study evaluated the quality of responses generated by AI chatbots to needs frequently reported by AYA patients with melanoma receiving treatment. Our methodologic approach included six components: (1) need selection, (2) question creation, (3) chatbot input, (4) data extraction, (5) data coding, and (6) evaluation (Figure 1).

XSL•FO
RenderX

Figure 1. Study design workflow. AI: artificial intelligence; AYA: adolescent and young adult.



Ethical Considerations

This study relied exclusively on publicly available chatbot outputs and did not involve interaction with human participants, collection of identifiable private information, or intervention. As such, it did not meet the definition of human participant research and did not require institutional research board approval.

Chatbot Input and Data Extraction

Chat GPT 4.0 (Open AI), Copilot (Microsoft), and Gemini (Google) were selected for evaluation based on their increasing usage among AYAs and advancements in their technology. The free, publicly accessible versions of all chatbots were used to reflect the most accessible chatbot platforms available to patients. The resulting need–related questions, outlined in Textbox 1, were entered into each chatbot 5 times by 4 individuals (RPM, ALO, KMR, and SZ) during a single study period (April 8, 2025, to August 8, 2025). To minimize bias, these team members did not participate in question selection, and AI personalization was avoided through memory clearing between each input, ensuring that responses did not draw on prior prompts or stored contextual data, thus replicating a true AYA patient experience. Responses were copied and exported into an Excel spreadsheet where a separate team completed data coding and evaluation.

Data Coding and Evaluation

Responses were evaluated and coded to assess information quality (Global Quality Score [GQS] and DISCERN), perceived empathy (Perceived Empathy of Technology), and readability (Flesch Kincaid Grade Level [FKGL], Flesch Reading Ease [FRE]; Table 1). Four team members, 2 medical students (JLJ and PLH) and 2 physicians (KCN and MHS), completed evaluations independently, with results hidden until completion to minimize bias. Quality metrics (GQS and DISCERN) were evaluated by physician reviewers due to their clinical expertise, while empathy (PETS) was evaluated by medical student reviewers to approximate a patient-facing perspective. Score interpretation for all tools is provided in Table 1. Reviewers met prior to scoring to align on scale interpretation. Any discrepancy in scores between the raters is a product of the qualitative nature of the rating and is representative of different physicians' interpretation. All analyses are exploratory and descriptive. Duplicate independent scoring with averaging was used to mitigate individual subjectivity, which was considered sufficient for this evaluation using previously validated instruments. Formal interrater reliability statistics (eg, intraclass correlation coefficient or Cohen κ) were not calculated, as estimates based on only 2 raters per domain may be unstable and potentially misleading, particularly for perceptual constructs such as empathy. The results are presented as means, standard deviations (SD), and visual comparisons. No inferential statistical testing was performed, and no claims of statistical significance or superiority were made.

XSL•FO
RenderX

Table . Evaluation metrics for chatbot responses, detailing assessed domains and score interpretation.

Assessment tool	Evaluated domains	Scale	Score interpretation
Global Quality Score [8]	• Accessibility • Quality • Flow • Usefulness	• 5-point Likert scale	• 1=Poor quality, flow, and missing information • 2=Fair • 3=Moderate • 4=Good • 5=Excellent quality, flow, and information
DISCERN [9]	• Accuracy • Quality	• Items 1 - 8 (5-point Likert scale) • Item 16 (summary 5-point Likert scale)	• Low quality information ≤2 • Medium-quality information=2 - 4 • High-quality information ≥4
Perceived Empathy of Technology Scale [10]	• ER[a], UT[b]	• 10 items, 2 domains. Each item is assessed with a 10-point scale and then averaged across 5 items for each domain score (ER and UT)	• 1 - 3=Low empathy • 4 - 6=Moderate empathy • 7 - 10=High empathy
Flesch Kincaid Grade Level [11] (FKGL)	• Comprehension	• Automated formula estimating US school grade level required to understand text	• Numeric grade level (eg, 8=8th-grade reading level)
Flesch Reading Ease [11] (FRE)	• Readability • Flow	• Automated formula producing a score from 0 to 100, with higher value indicating easier reading	• 0 - 30=Very difficult to read • 31 - 50=Difficult to read • 51 - 60=Fairly difficult to read • 61 - 70=Plain English • 70 - 100=Easy to read

[a]ER: Emotional Responsiveness.

[b]UT: Understanding and Trust.

Quality (GQS and DISCERN)

Chatbot answer quality was evaluated using the GQS and DISCERN scales. The GQS scale evaluates quality, accessibility, flow, and perceived usefulness of information for patients, as judged by a physician (Table S1 in Multimedia Appendix 1). The DISCERN scale assesses the quality of written health information, focusing on reliability, clarity, and explanations of risks and benefits (Table S2 in Multimedia Appendix 1). Two physicians (KCN and MHS) evaluated each response with the QGS and DISCERN scales. The results were then averaged into a final score.

Empathy (PETS)

The PETS, a 10-item instrument designed to assess emotional responsiveness and understanding in human-technology interactions, was used to evaluate perceived empathy in chatbot responses. The PETS was applied without modification. The PETS tool is divided into 2 sections: PETS-ER (6 items) and PETS-UT (4 items; Table S3, Multimedia Appendix 1). The PETS-ER (Emotional Responsiveness) focuses on emotional engagement and support. PETS-UT (Understanding and Trust) measures “cognitive empathy” or the ability of the Chatbot to understand the user’s perspective. Each item was individually scored, then averaged to give domain-specific scores for each input. Two medical student reviewers (JLJ and PLH) performed scoring and highlighted key examples for qualitative evaluation.

Readability (FKGL, FRE, and Word Count)

To evaluate each output’s readability, the FKGL and FRE were utilized. FKGL estimates the US grade level required to comprehend a response, while FRE scores readability from 0 to 100 (higher=easier to read). Scores for both instruments were calculated in Microsoft Word’s spelling and grammar tool.

Results

Descriptive metrics of chatbot performance are reported in Table 2, stratified by question with overall summary statistics for each chatbot. Individual input scoring is available in Multimedia Appendix 2.

XSL•FO
RenderX

Table . Descriptive quality, empathy, and readability metrics for chatbot responses by question and overall.

Question and chatbot	GQS[a], mean (SD)	DISCERN, mean (SD)	PETS-ER[b], mean (SD)	PETS-UT[c], mean (SD)	FRE[d], mean (SD)	FKGL[e], mean (SD)	Word count, mean (SD)
Question and chatbot							
1							
ChatGPT	4.50 (0.35)	3.34 (0.30)	6.18 (1.02)	7.80 (0.73)	39.20 (4.77)	12.62 (1.71)	538.80 (66.43)
Copilot	4.00 (0.00)	2.75 (0.09)	5.70 (0.79)	6.45 (0.47)	37.32 (5.68)	11.82 (0.82)	344.80 (54.70)
Gemini	4.00 (0.00)	3.09 (0.14)	5.87 (0.68)	6.50 (1.01)	46.42 (1.20)	11.10 (0.22)	779.80 (69.88)
2							
ChatGPT	4.60 (0.22	3.35 (0.12)	7.53 (0.53)	7.60 (0.80)	54.84 (3.43)	9.10 (0.64)	522.40 (78.60)
Copilot	4.50 (0.35)	3.15 (0.15)	6.67 (0.57)	6.65 (0.38)	46.42 (4.37)	10.72 (0.72)	431.80 (30.91)
Gemini	4.20 (0.27)	3.09 (0.21)	6.67 (0.57)	6.65 (0.38)	52.68 (4.91)	9.80 (0.99)	728.00 (58.94)
3							
ChatGPT	4.00 (0.00)	3.00 (0.00)	3.70 (1.31)	4.18 (0.85)	28.58 (4.78)	11.98 (0.47)	389.20 (48.84)
Copilot	4.10 (0.42)	2.75 (0.10)	4.13 (0.40)	4.10 (0.42)	31.82 (7.22)	12.38 (1.35)	315.20 (47.10)
Gemini	4.50 (0.00)	3.05 (0.11)	4.13 (0.40)	4.10 (0.42)	34.38 (2.51)	13.12 (0.62)	717.80 (71.51)
4							
ChatGPT	4.50 (0.00)	2.94 (0.19)	4.33 (0.41)	6.97 (0.55)	37.62 (1.24)	10.84 (0.30)	545.20 (40.76)
Copilot	3.40 (0.65)	2.55 (0.14)	3.10 (0.37)	3.15 (0.34)	31.72 (4.47)	13.40 (0.65)	360.20 (22.22)
Gemini	4.00 (0.00)	2.81 (0.14)	3.10 (0.37)	3.15 (0.34)	40.60 (4.59	12.16 (0.70)	703.00 (45.08)
5							
ChatGPT	4.40 (0.42)	3.26 (0.50)	5.03 (2.38)	5.30 (2.35)	29.36 (6.90)	12.78 (0.97)	415.60 (178.75)
Copilot	3.70 (0.27)	2.74 (0.19)	2.20 (0.45)	2.70 (0.11)	33.38 (6.52)	12.74 (0.72)	330.80 (34.47)
Gemini	4.20 (0.27)	2.76 (0.03)	2.20 (0.45)	2.70 (0.11)	34.68 (3.20)	12.74 (0.57)	713.60 (50.51)
Overall ratings across all questions							
ChatGPT	4.42 (0.32)	3.24 (0.31)	5.35 (1.85)	6.36 (1.83)	38.9 (10.5)	11.9 (1.64)	482 (111)
Copilot	3.94 (0.53)	2.79 (0.24)	4.36 (1.74)	4.61 (1.72)	36.1 (7.7)	12.3 (1.23)	356 (55)
Gemini	4.18 (0.24)	2.96 (0.19)	4.40 (1.76)	4.62 (1.76)	41.8 (7.9)	11.8 (1.37)	728 (61)

[a]GQS: Global Quality Score.

[b]PETS-ER: Perceived Empathy of Technology Scale-Emotional Responsiveness.

[c]PETS-UT: Perceived Empathy of Technology Scale-Understanding and Trust.

[d]FRE: Flesch Reading Ease.

[e]FKGL: Flesch-Kincaid Grade Level.

Quality (GQS and DISCERN)

ChatGPT demonstrated the highest average information quality scores, achieving the highest GQS and DISCERN ratings for 4 out of the 5 questions (Figures 2 and 3). Copilot responses showed the lowest average quality scores, while Gemini responses generally fell between the 2 other models.

XSL•FO
RenderX

Figure 2. Global Quality Score (GQS) scores comparing each chatbot by question input.

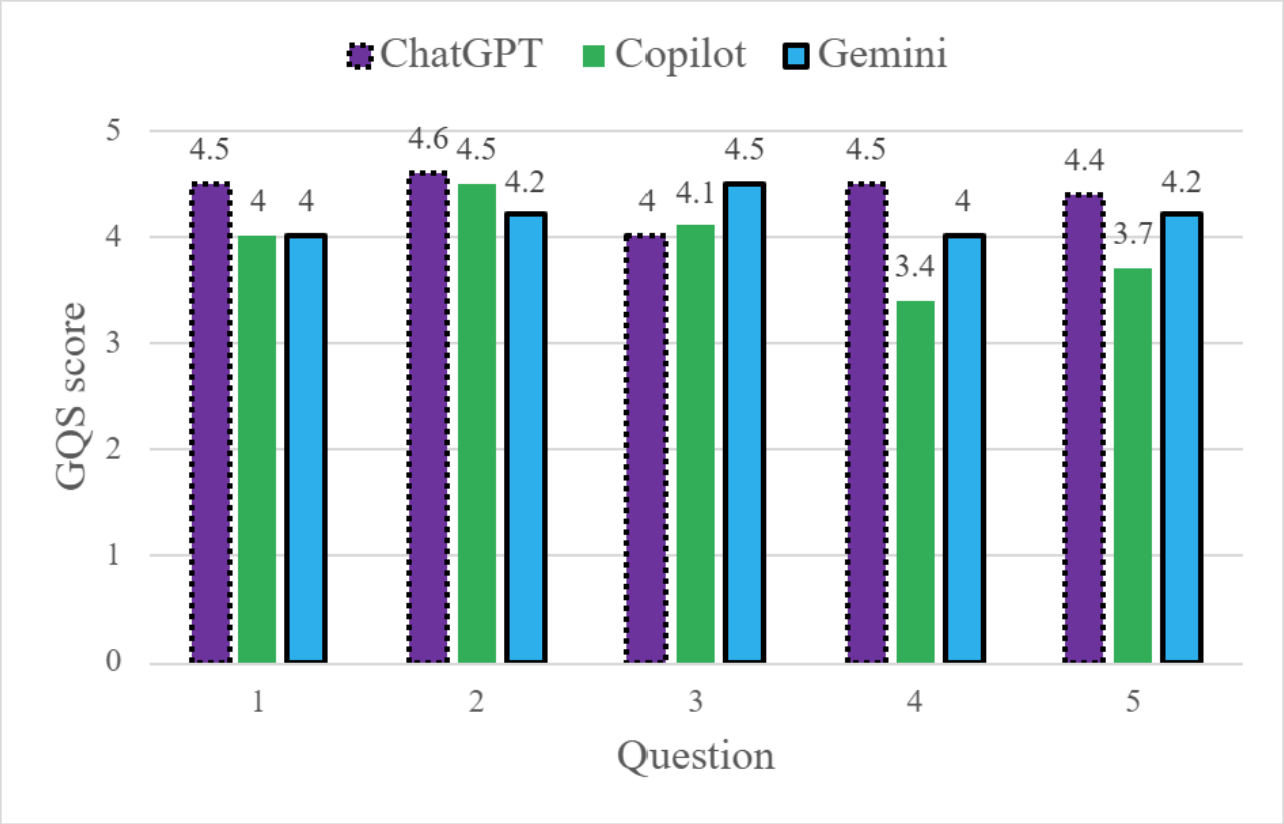
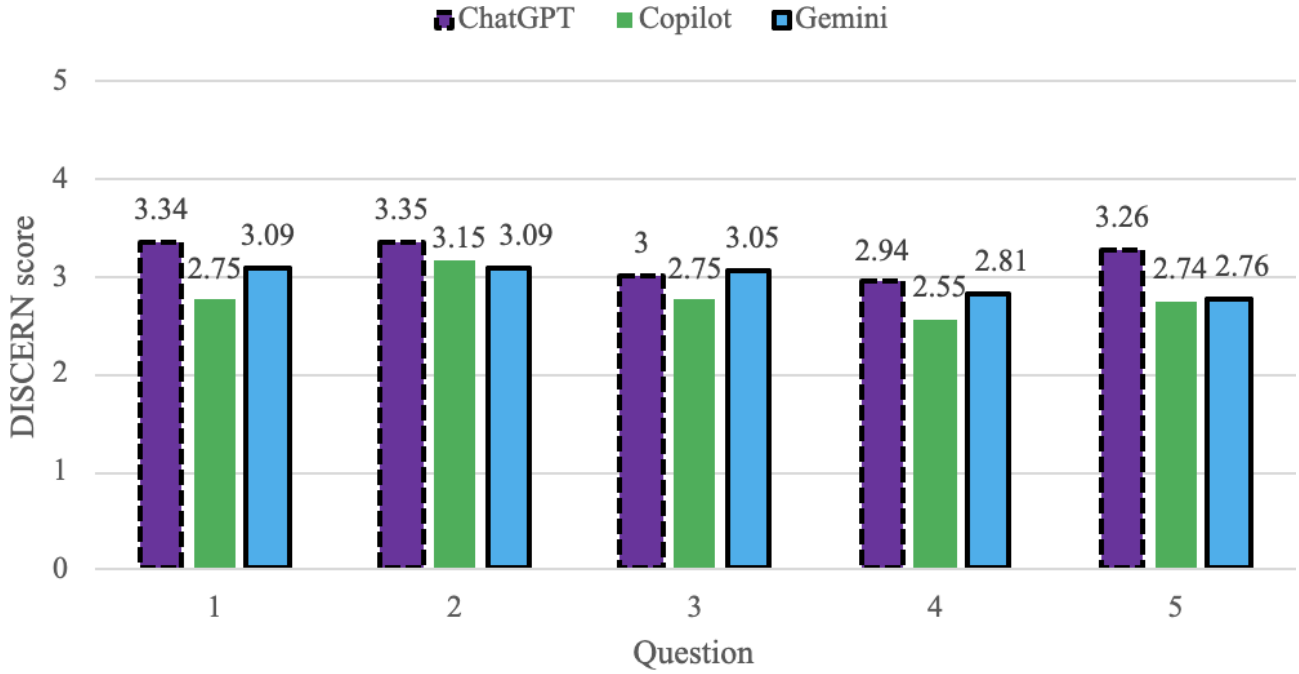


Figure 3. DISCERN scores comparing each chatbot by question input.



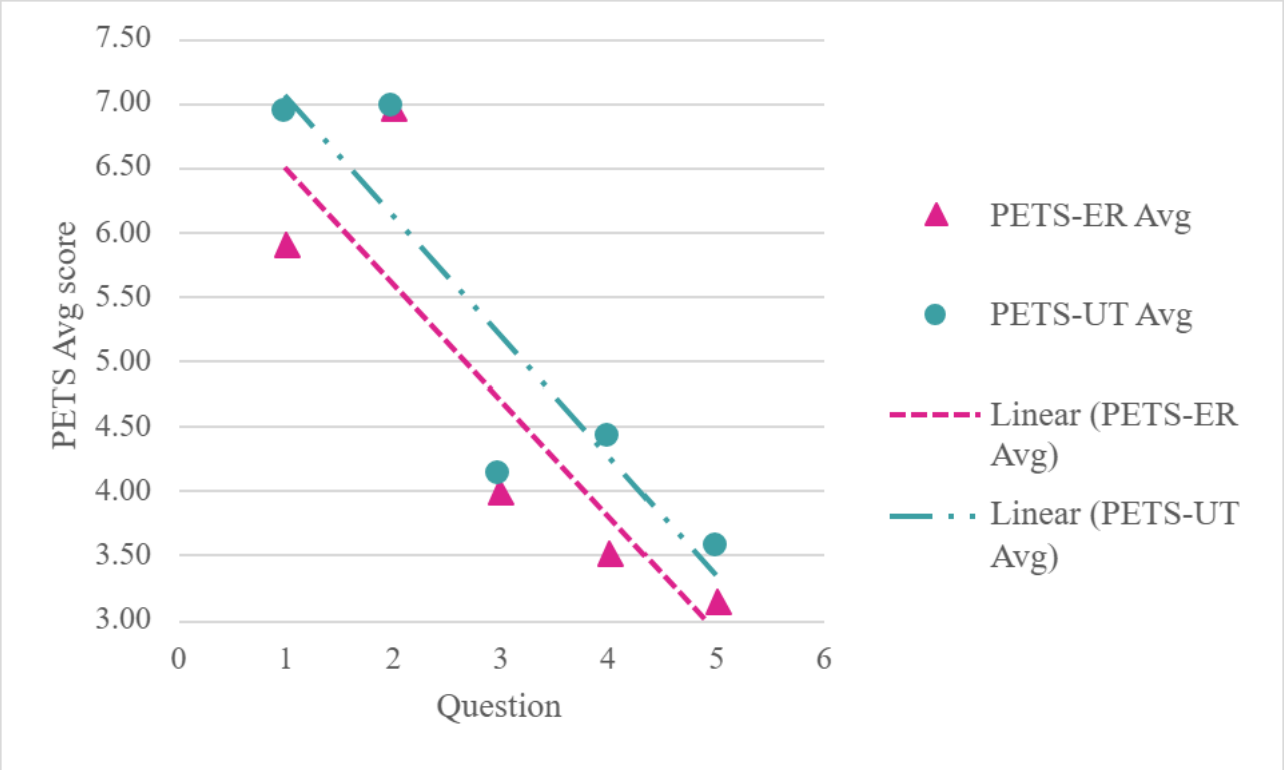
Quality variability differed across chatbots (Table 2). Copilot responses demonstrated greater variability in quality scores across question, while Gemini responses showed more consistent quality patterns. ChatGPT exhibited moderate variability for quality metrics. At the question level, Questions 4 and 5 were associated with greater variability in quality scores across chatbots, while Questions 2 and 3 yielded more similar scores across models (see Multimedia Appendix 3 for all SD). Overall, descriptive rank ordering based on average scores and variability suggests differences in consistency and distribution of information quality across chatbots.

Empathy (PETS-ER and PETS-UT)

PETS scores varied at the question and chatbot level (Table 2). Questions 1 and 2 are associated with the highest perceived empathy, and Question 5 demonstrates the lowest empathy scores (Figure 4).

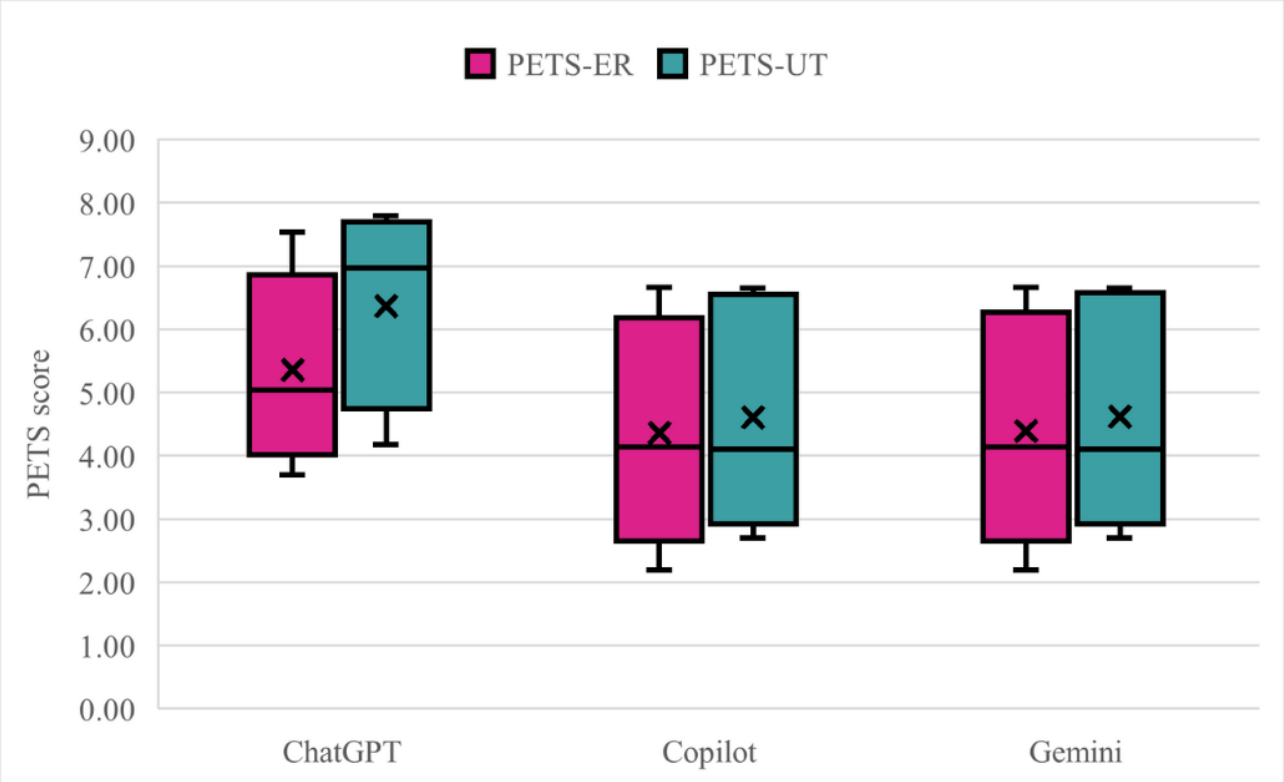
XSL•FO
RenderX

Figure 4. Average Perceived Empathy of Technology Scale (PETS) across chatbot by question. Avg: average; ER: Emotional Responsiveness; UT: Understanding and Trust.



When empathy scores were summarized by chatbot, descriptive differences in average perceived empathy were observed across models (Figure 5). ChatGPT responses demonstrated higher average empathy scores compared to Copilot and Gemini, which demonstrated lower averages. Across all models, PETS-UT scores exceeded PETS-ER scores, suggesting stronger cognitive empathy and perspective taking rather than emotional expression.

Figure 5. Box and whisker plots showing Perceived Empathy of Technology Scale-Emotional Responsiveness (PETS-ER) and Perceived Empathy of Technology Scale-Understanding and Trust (PETS-UT) scores across chatbots, illustrating limited overall perceived empathy compared to understandability.



XSL•FO
RenderX

ChatGPT showed the largest UT-ER gap (mean 6.36, SD 1.83 vs 5.35, SD 1.85), indicating moderate-to-high levels of understanding, but only moderate emotional connection. Copilot and Gemini had similar low-moderate ER and UT averages, with limited perceived empathy overall. Variability patterns further distinguish AI models (Figure 5). ChatGPT produced the widest variation in empathy scores. In contrast, Gemini and Copilot showed slightly more predictable patterns of empathy.

Readability (FRE, FKGL, and Word Count)

All outputs were written at an eighth-grade reading level or higher (range 8.2 - 14.6, mean 11.82, SD 1.44, equivalent to high school senior level). The average FRE score across all AI chatbot responses was 38.60, SD 9.00, indicating difficult readability overall. Gemini tended to produce the longest but most readable responses. Copilot consistently produced shorter responses associated with lower average readability. ChatGPT produced responses with the most variability in readability and word count. Question 5 responses were the least readable (mean grade level 12.76, SD 0.71), while question 2 responses were the most readable (mean grade level 9.9, SD 1.01).

Discussion

Quality

ChatGPT responses were associated with higher average quality scores and moderate variability across multiple questions. Gemini responses demonstrated moderate quality scores with greater consistency, while Copilot responses exhibited greater variability and the lowest quality scores, even producing a factually incorrect answer regarding genetic mutations and the associated risk of developing melanoma (see Copilot 4_1 in Multimedia Appendix 4). In health communication, information quality standardization is critical because it allows for less confusion and increases patient-provider trust [12].

Quality is not only model dependent but also question reliant. Questions 1 and 2 received the highest quality evaluations, whereas Questions 4 and 5 produced lower quality scores. This pattern suggests that chatbots tend to provide more consistent responses to straightforward or well-established topics. In contrast, questions regarding more complex science with sparse and/or unclear literature, such as recurrence risk, long-term treatment effects, and genetic implications, produced less reliable responses, often without communicating their limitations. These complex topics often represent areas of greatest concern for AYA patients with melanoma, which underscores the importance of cautious interpretation of chatbot-generated information.

The review of lower-quality outputs revealed common pitfalls. Responses with low-quality scores were often oversimplified, incomplete, or presented in bullet-point style responses, with insufficient detail to address the question. The lack of actionable information was most apparent in discussions of treatment, side effects, and hormone changes. For example, when a patient asked how treatment would affect hormones, responses accurately identified potential consequences, such as adrenal insufficiency, but failed to describe symptoms. From a patient perspective, clinically beneficial communication translates these answers into recognizable manifestations at home (ie, fatigue, dizziness, or other changes). Accuracy was further undermined by the lack of transparent source citation and inconsistent use of safeguards, such as disclaimers. Instead, statements were presented as definitive, risking patient confusion when conflicting with other resources or provider guidance.

Conversely, higher-quality responses demonstrated more complete and organized content. They cited credible resources, gave disclaimers, and emphasized the importance of deferring to the health care team for individualized guidance. These responses were detailed and well-tailored to the questions, without an overwhelming word count, which improved readability and digestion of information. They had clear directions to access specific resources for further support and actionable suggestions for the patient. These qualities serve as a reference standard for future versions of AI-generated health information.

Although the overall quality scores across all chatbot responses were moderate to high, persistent gaps in sourcing and uncertainty disclosures raise serious patient safety concerns. For AYA patients with melanoma undergoing treatment, inconsistent quality can undermine trust and lead to confusion at a vulnerable stage of care. To be reliable in clinical contexts, AI systems must deliver outputs that are not only accurate and detailed but also verifiable, transparent, and responsibly framed.

Empathy

Across the 3 chatbot platforms, a central pattern emerged: systems struggled to balance clinical accuracy with emotional support. ChatGPT comparatively showed the highest empathetic capacity, while Gemini and Copilot tended to be neutral and lacking emotion. The qualitative review of responses identified a set of best practices that distinguished qualities of higher-empathy outputs from weak ones (Textbox 2). Detailed examples supporting these best practices are provided in Multimedia Appendix 5.

XSL•FO
RenderX

Textbox 2. Key best practices for artificial intelligence (AI) chatbot responses to express empathy.

Best practices (Perceived Empathy of Technology Scale [PETS] scores)

- Open with empathy and understanding
- Asks follow-up and clarification questions
- Symptoms and side effects in plain language
- Medical disclaimers and boundaries
- Cites credible resources
- Specific actionable coping strategies
- Adolescents and young adults–specific
- Demonstrates interest in the patient
- Provides empathy in all contexts

Common pitfalls (↓ PETS scores)

- Robotic or purely clinical tone
- One-way information delivery
- Lists of side effects or clinical jargon
- Overstep scope
- Lack citations or specific resources
- Generic or overwhelming advice
- Adult-centric or generic framing
- Lack of interest or personalization
- Overlooks unspoken emotional needs

AYA patients with melanoma have unique needs compared to adults; however, chatbots often default to general or even pediatric framing [13]. When AYA patient-specific needs were addressed, it was often with a surface-level list rather than a meaningful explanation. Systems were better equipped to name emotions rather than express them, as demonstrated by higher UT than ER scores, and empathy was also prompt-dependent. Chatbot responses generally failed to infer patient distress if not explicitly stated (as seen in questions 3-5), unlike a human clinician who can use nonverbal cues and proactively address unspoken emotional needs even in the face of an analytical question. The chatbot lacks the emotional intelligence to deliver difficult news and interpret patient reactions. Responses receiving low PETS scores risk coming across as dismissive or cold, which may invalidate patient concerns and discourage them from seeking support or adhering to care recommendations in the future when they most need it.

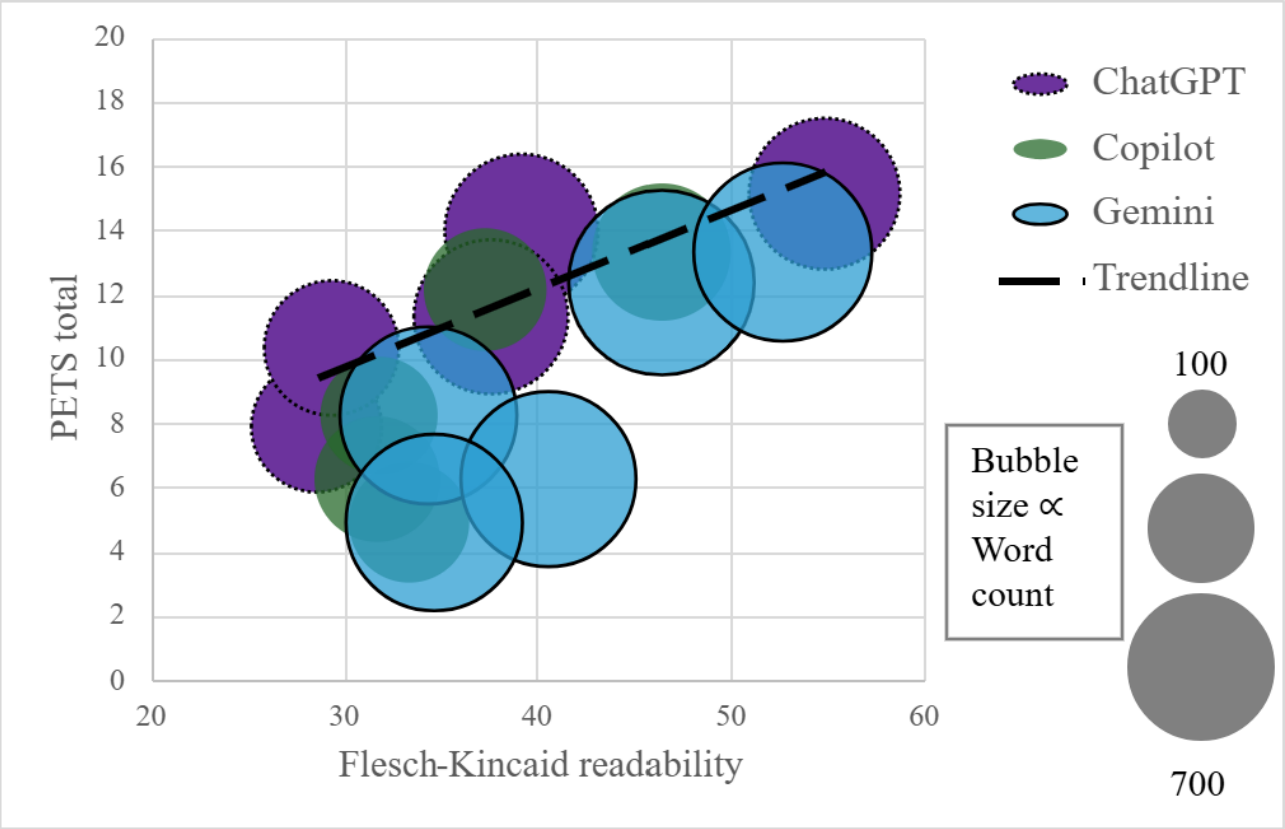
Human providers naturally contextualize information based on patient age, life stage, current treatments, and psychosocial needs, a skill beyond a chatbot's reach. Since our 5 questions represent domains of greatest emotional and informational needs of AYAs with melanoma, the irregularity of empathy suggests that current AI systems may not adequately support this unique patient population.

Readability

Response length and readability varied across chatbots. Gemini responses were longer on average with higher readability scores, whereas ChatGPT responses demonstrated wider variability, and Copilot responses were shorter with lower average readability scores. Responses that were easier to read tended to be perceived as higher quality (GQS/DISCERN) and more empathetic (PETS) as seen in Figure 6.

XSL•FO
RenderX

Figure 6. The relationship between Flesch-Kincaid readability and total Perceived Empathy of Technology (PETS) (empathy) scores across all chatbot responses. Each bubble represents a chatbot, with size proportional to overall word count. While word count shows little association with either readability or empathy, responses with higher readability tend to receive higher empathy scores.



Readability analysis revealed that all chatbot responses exceeded the US average reading level (7th-8th grade), with outputs averaging at the 11th-12th grade level. This creates a barrier for accessibility. These findings echo prior research that health care documents are often written at a level too high for general populations, reducing inclusivity and comprehension [14]. Unlike clinicians, who are trained to use strategies such as follow-up questions and the "teach-back" method, chatbots do not check understanding unless prompted. This may highlight a key limitation of AI chatbots for medical communication. They draw primarily from online sources, which are often already written at inflated reading levels, and thus lead to reproduced complexity by AI outputs [15]. Clinicians, by contrast, are trained to adapt explanations to patient literacy, developmental stage, and clinical context.

Readability is particularly important for AYAs, whose health literacy may be affected by education disruption, treatment-related cognitive impairment, and emotional distress. Cancer-related cognitive impairment affects approximately 25% to 28% of AYA survivors and may affect attention, memory, and processing speed, increasing the cognitive burden of complex medical information [16]. The National Comprehensive Cancer Network Guidelines for AYA Oncology emphasize that information should be delivered in an age-appropriate and developmentally appropriate manner, particularly in the context of disrupted education and psychosocial stressors [17]. While no evidence-based guidelines specify an optimal reading level for AYA oncology populations, the commonly cited 8th-grade target is extrapolated from general health literacy recommendations and represents a pragmatic benchmark rather than an AYA-specific threshold.

AI may therefore generate content that is technically accurate but inaccessible, potentially leading to confusion or disengagement. For proper implementation, chatbots should aim to target their output responses of health care information to an 8th-grade level and incorporate interactive strategies that mimic provider feedback loops. Without these adjustments, the use of AI to answer the needs of AYA patients with melanoma remains limited.

Patient Safety and Clinical Implications

Although AI chatbots may provide AYA patients with adequate education, support, and resources regarding their diagnosis, they cannot replace clinician input and should be used cautiously. The potential dangers associated with patients using AI in health care are multifaceted, with key risks including misinformation, unsafe recommendations, lack of understanding of the questions being asked, and insensitivity. Given that AYAs are the primary users of AI compared to other age groups, they may be particularly vulnerable to these risks due to their vulnerability to algorithmic bias and variable access to clinical guidance [18]. These risks are inferred from content characteristics of chatbot responses and were not measured as patient outcomes in this study.

AI use in health care has been shown to spread misinformation [19]. In our study, we found that chatbot explanations regarding the causes of metastatic melanoma and recommendations for next steps may include inaccurate or misleading information.

XSL•FO
RenderX

Responses across all 3 platforms often focused on sun avoidance and using sun protection factor to prevent melanoma recurrence or secondary malignancy. Although this advice may be relevant for certain patient populations, it is not applicable to all patients, particularly for individuals with skin of color. As a result, responses lack inclusivity for diverse patient communities. Additionally, some chatbot responses encouraged AYA patients to explore clinical trials available for their conditions. While well-intended, this response is particularly unsafe to AYA patients, as it may increase the risk of false hope or confusion if patients are encouraged to consider trials that are not medically appropriate. Some responses even offered to interpret lab and genetic testing results. While AI chatbots have the potential to explain pathology reports and simplify medical jargon to patients, their interpretations are not flawless and should be used carefully with clinician oversight [20]. Encouraging vulnerable patients to input their private health information to AI systems may expose them to breaches of data privacy and security, potentially leading to unauthorized access or misuse of sensitive medical information [21].

Limitations

This study evaluated a limited number of single-turn responses (5 question and 75 single-turn responses), which may not fully capture chatbot performance. AI chatbots can refine responses through follow-up prompts, so the single-turn design does not reflect extended, multiturn conversations. This limitation may influence the perceived quality, empathy, and readability of chatbot responses. Studies testing multiturn chats may be beneficial in the future. In addition, chatbot responses are sensitive to prompt phrasing and contextual detail, and the use of subjective scoring instruments introduces potential variability in interpretation. Each response was independently evaluated by 2 reviewers per domain, with scores averaged to reduce individual rater bias. For transparency, we have included all outputs in Multimedia Appendices 4, 6 and 7 as well as an example of the grading process with score justifications in Multimedia Appendix 8. Formal interrater reliability statistics were not calculated, which limits quantitative assessment of scoring consistency. Although duplicate independent scoring with averaging was used to reduce subjectivity, future studies with larger and cross-disciplinary reviewer pools should incorporate formal reliability testing to further strengthen methodological rigor. Additionally, the separation of reviewer roles by domain may introduce systematic differences in scoring interpretation.

Future Directions

Moving forward, we aim to use responses from this dataset to develop resources within melanoma oncology clinics that answer these questions and provide higher-quality and safer responses for AYA patients with melanoma. By providing age-appropriate education and support tools, we hope to improve the treatment experience and the overall well-being for AYAs navigating melanoma care. This will help expand the currently limited resources available for AYAs currently receiving treatment and hopefully bring more attention to the unaddressed survivorship needs of patients. Future clinic-integrated resources would be developed with clinician oversight and human evaluation to ensure accuracy, safety, and appropriateness for patient use, rather than relying on autonomous AI fine-tuning. If this study were to be replicated or expanded upon, it would be helpful to incorporate intraclass correlation coefficient or Cohen κ into the data analysis to correct for discrepancies in the grading system. Additionally, this study establishes a patient-centered framework that can be applied to future research evaluating AI chatbot communication around emerging and experimental melanoma therapies, including novel immunotherapies and mRNA-based cancer vaccines. Our findings suggest that chatbot performance declines in areas with limited or evolving scientific literature, indicating that AI responses on new treatments or clinical trials may be particularly vulnerable to misinformation or oversimplification. Future studies should specifically assess this phenomenon to generate realistic expectations for AYAs navigating the rapidly evolving melanoma therapeutic landscape.

Conclusion

This study provides an evaluation of the utility of AI to answer the most pressing needs of AYA patients with melanoma currently receiving treatment. Overall, chatbot responses were often informative but demonstrated persistent gaps in quality, empathy, and readability. Patient safety may be at risk due to the inability of AI systems to integrate patient context, emotional nuance, and conversational reciprocity into answers. Question wording, emotional phrasing, and clarity were predictive of the responses received. When questions addressed topics covered by multiple publications, AI responses were more consistent, with higher quality. However, when questions addressed controversial or rapidly changing topics, AI responses were often vague. Quality, empathy, and readability were all interconnected, as higher-rated responses incorporated clear and concise wording, showed interest in the patient's medical status, provided tangible resources and references, and recognized the limits of AI, encouraging further conversation with medical providers. This research should be replicated and can provide a reference standard for future AI chatbot responses for AYA patients with melanoma.

Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) in the research process. According to the GAIDeT taxonomy (2025), data collection was delegated to GAI tools under full human supervision. The GAI tools used were ChatGPT (GPT-4o), Gemini, and Copilot. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

XSL•FO
RenderX

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The dataset generated during this study, including all chatbot outputs and scoring, is available in the Multimedia Appendices. No human participant data were collected.

Authors' Contributions

JLJ and PLH are co-first authors and contributed equally to this work. JLJ, PLH, and KCN contributed to conceptualization and methodology; JLJ and PLH contributed to question development, empathy, readability scoring, and manuscript writing. RPM, SZ, ALO, and KMR performed question input and data collection. MR provided project oversight and expert clinical opinion. MHS contributed to manuscript writing and editing, quality scoring, and response review. KCN served as principal investigator; provided project oversight, conceptualization, methodology, contributed to data collection; and participated in manuscript writing and editing. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Global Quality Score (GQS), DISCERN, and Perceived Empathy of Technology Scale (PETS) scoring scales utilized for this study.
[DOCX File, 17 KB - cancer_v12i1e84234_app1.docx]

Multimedia Appendix 2
Chatbot response scores for each input.
[DOCX File, 31 KB - cancer_v12i1e84234_app2.docx]

Multimedia Appendix 3
Calculated standard deviations of response scores.
[DOCX File, 19 KB - cancer_v12i1e84234_app3.docx]

Multimedia Appendix 4
All Microsoft Copilot outputs.
[DOCX File, 75 KB - cancer_v12i1e84234_app4.docx]

Multimedia Appendix 5
Expanded best practices table.
[DOCX File, 16 KB - cancer_v12i1e84234_app5.docx]

Multimedia Appendix 6
All Google Gemini outputs.
[DOCX File, 74 KB - cancer_v12i1e84234_app6.docx]

Multimedia Appendix 7
All ChatGPT outputs.
[DOCX File, 77 KB - cancer_v12i1e84234_app7.docx]

Multimedia Appendix 8
Example of response score breakdown.
[DOCX File, 16 KB - cancer_v12i1e84234_app8.docx]

References

1. National Cancer Institute. Cancer stat facts: cancer among adolescents and young adults (AYAs). SEER Program. URL: https://seer.cancer.gov/statfacts/html/aya.html [accessed 2025-08-23]

XSL•FO
RenderX

2. McInally W, Gray-Brunton C, Chouliara Z, Kyle RG. Life Interrupted: experiences of adolescents, young adults and their family living with malignant melanoma. J Adv Nurs 2021 Sep;77(9):3867-3879. [doi: 10.1111/jan.14959] [Medline: 34245464]
3. Sidoti O, McClain C. 34% of U.S. adults have used ChatGPT, about double the share in 2023. Pew Research Center. URL: https://www.pewresearch.org/short-reads/2025/06/25/34-of-us-adults-have-used-chatgpt-about-double-the-share-in-2023/ [accessed 2025-08-23]
4. Clark M, Bailey S. Chatbots in health care: connecting patients to information. Can J Health Technol 2024;4(1). [doi: 10.51731/cjht.2024.818]
5. Kim HK. The effects of artificial intelligence chatbots on women's health: a systematic review and meta-analysis. Healthcare (Basel) 2024 Feb 23;12(5):534. [doi: 10.3390/healthcare12050534] [Medline: 38470645]
6. Pan A, Musheyev D, Bockelman D, Loeb S, Kabarriti AE. Assessment of artificial intelligence chatbot responses to top searched queries about cancer. JAMA Oncol 2023 Oct 1;9(10):1437-1440. [doi: 10.1001/jamaoncol.2023.2947] [Medline: 37615960]
7. Moore R, Ahmad N, Kadosh M, et al. Identifying gaps and needs in care for adolescents and young adults with melanoma. Support Care Cancer 2025 Oct 7;33(10):913. [doi: 10.1007/s00520-025-10001-1] [Medline: 41057735]
8. Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. Am J Gastroenterol 2007 Sep;102(9):2070-2077. [doi: 10.1111/j.1572-0241.2007.01325.x] [Medline: 17511753]
9. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111. [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
10. Schmidmaier M, Rupp J, Cvetanova D, Mayer S. Perceived empathy of technology scale (PETS): measuring empathy of systems toward the user. Presented at: CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems; May 11, 2024. [doi: 10.1145/3613904.364203]
11. Ley P, Florio T. The use of readability formulas in health care. Psychol Health Med 1996 Feb;1(1):7-28. [doi: 10.1080/13548509608400003]
12. Chauhan V, Sagar M. Healthcare decision making and choice: an empirical model of patient confusion. Manag Decis 2023 Nov 7;61(11):3454-3474. [doi: 10.1108/MD-11-2022-1488]
13. Ahmad N, Kashyap A, Taylor M, et al. Quality of life among adolescents and young adults with melanoma: a systematic review. J Adolesc Young Adult Oncol 2026 Feb;15(1):1-13. [doi: 10.1089/jayao.2025.0029] [Medline: 40470545]
14. Espinosa J, Lucerna A, Schuitema H. Healthcare documents: how readable? A study of simple strategies to decrease reading grade level and increase reading ease of healthcare documents. Mathews J Emergency Med 2022;7(1):44. [doi: 10.30654/MJEM.10044]
15. How ChatGPT and our foundation models are developed. OpenAI Help Center. URL: https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed [accessed 2025-09-10]
16. NCCN clinical practice guidelines in oncology: adolescent and young adult (AYA) oncology. : National Comprehensive Cancer Network (NCCN); 2025 URL: https://www.nccn.org/professionals/physician_gls/pdf/aya.pdf [accessed 2026-01-06]
17. Vizer LM, Mikles SP, Piepmeier AT. Cancer-related cognitive impairment in survivors of adolescent and young adult non-central nervous system cancer: a scoping review. Psychooncology 2022 Aug;31(8):1275-1285. [doi: 10.1002/pon.5980] [Medline: 35726379]
18. Carolan S, Martin AW, Gong GC, Borja S. The state of consumer AI. : MENLO Ventures; 2025 URL: https://menlovc.com/wp-content/uploads/2025/11/menlo_ventures_consumer_ai_report-2025.pdf [accessed 2026-03-06]
19. Denecke K, Lopez-Campos G, Rivera-Romero O, Gabarron E. The unexpected harms of artificial intelligence in healthcare: reflections on four real-world cases. Stud Health Technol Inform 2025 May 2;325:55-60. [doi: 10.3233/SHTI250219] [Medline: 40326654]
20. Steimetz E, Minkowitz J, Gabutan EC, et al. Use of artificial intelligence chatbots in interpretation of pathology reports. JAMA Netw Open 2024 May 1;7(5):e2412767. [doi: 10.1001/jamanetworkopen.2024.12767] [Medline: 38776080]
21. Li J. Security implications of AI chatbots in health care. J Med Internet Res 2023 Nov 28;25:e47551. [doi: 10.2196/47551] [Medline: 38015597]

Abbreviations

AI: artificial intelligence
AYA: adolescents and young adults
ER: Emotional Responsiveness
FKGL: Flesch-Kincaid Grade Level
FRE: Flesch Reading Ease
GQS: Global Quality Score
NA-SB: Needs Assessment and Service Bridge

XSL•FO
RenderX

PETS: Perceived Empathy of Technology Scale
UT: Understanding and Trust

Edited by M Balcarras; submitted 21.Sep.2025; peer-reviewed by D Chrimes, M Chatzimina, T Fabrizio; revised version received 13.Feb.2026; accepted 25.Feb.2026; published 26.Mar.2026.

Please cite as:
Jafarnia JL, Haff PL, Moore RP, Osheim AL, Riley KM, Zheng S, Roth M, Salge MH, Nelson KC
Quality, Empathy, and Readability of AI Chatbot Responses to the Survivorship Needs of Adolescents and Young Adults With Melanoma: Evaluation Study
JMIR Cancer 2026;12:e84234
URL: https://cancer.jmir.org/2026/1/e84234
doi:10.2196/84234

© Jordan Lily Jafarnia, Priscilla Lynne Haff, Reece Philip Moore, Alyssa Leigh Osheim, Katherine McKenna Riley, Sabrina Zheng, Michael Roth, Madeleine Hines Salge, Kelly Carter Nelson. Originally published in JMIR Cancer (https://cancer.jmir.org), 26.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Stakeholder Perspectives on Cancer Survivors' Return to Work and Well-Being: Qualitative Interview Study

Sara L Fällman[1,2], PhD; Katarina Wijk[3,4,5], PhD; Anna Efverman[1], PhD; Linda Eklund[1,6,7], MSc; Per Fessé[1,3], PhD; Sven Trygged[8], PhD; Maria Fjell[1,9], PhD

[1]Department of Caring Science, Faculty of Health and Occupational Studies, University of Gävle, Kungsbäcksvägen 47, Gävle, Gävleborg, Sweden

[2]Department of Work Life and Social Welfare, Faculty of Caring Science, Work Life and Social Welfare, University of Borås, Borås, Västra Götaland, Sweden

[3]Centre for Research and Development, Region Gävleborg, Uppsala University, Uppsala, Uppsala, Sweden

[4]Department of Occupational Health, Psychology and Sports Sciences, Faculty of Health and Occupational Studies, University of Gävle, Gävle, Gävleborg, Sweden

[5]Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden

[6]Center for Clinical Research Dalarna, Uppsala University, Uppsala, Sweden

[7]Department of Occupational Therapy and Healthcare Counselling, Region Dalarna, Ludvika, Sweden

[8]Department of Social Work, Criminology, and Public Health Science, Faculty of Health and Occupational Studies, University of Gävle, Gävle, Gävleborg, Sweden

[9]Division of Nursing, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Sara L Fällman, PhD
Department of Caring Science, Faculty of Health and Occupational Studies, University of Gävle, Kungsbäcksvägen 47, Gävle, Gävleborg, Sweden

Abstract

Background: About 4 in 10 people in Sweden get cancer during their lifetime, and approximately half of them will be diagnosed during their working life. As cancer survival rates improve, a growing number of individuals face challenges in returning to work following treatment. This increases the demand for effective return-to-work (RTW) strategies. Despite existing rehabilitation frameworks, cancer survivors often encounter barriers to sustainable work reintegration.

Objective: This study aimed to investigate stakeholders' perceptions of facilitators and barriers to RTW among cancer survivors, including factors promoting work-related well-being during the RTW process.

Methods: During the development phase of a work-oriented rehabilitation intervention, semistructured interviews were conducted, with 25 stakeholders involved in the RTW process: health care professionals (n=12), social insurance officers (n=7), employers (n=5), and an employment service agency officer (n=1). Data were analyzed using qualitative content analysis.

Results: Five overarching themes were identified that influenced RTW: collaboration and clear division of responsibilities, balancing individual adaptations, reducing structural barriers through support, views and expectations of the individual regarding RTW, and the emotional significance of work. Key facilitators included flexible work arrangements, individualized adaptations, a strong desire to work, and the emotional value of work. Barriers comprised lack of collaboration among stakeholders, particularly the absence of a clear division of responsibilities, as well as structural barriers, such as sick leave bureaucracy and financial obstacles. The role of rehabilitation coordinators was highlighted as pivotal in bridging gaps between stakeholders and ensuring continuity in care.

Conclusions: RTW for cancer survivors is a complex, multifactorial process requiring coordinated efforts across health care, insurance, and employment sectors. Enhancing collaboration, clarifying stakeholder responsibilities, and implementing flexible, individualized support structures are essential to facilitating cancer survivors' RTW. Additionally, including a designated coordinator in the process is proposed. More support during the early phase of RTW is necessary to reduce the risk of long-term labor market exclusion. To help cancer survivors' RTW, clear role definitions and shared responsibilities among stakeholders are essential. Flexibility in the RTW process helps individuals reintegrate at their own pace, reduces isolation, and promotes social connection. These are key considerations for future policies and support measures.

(JMIR Cancer 2026;12:e89954) doi:10.2196/89954

KEYWORDS

cancer survivors; rehabilitation; return to work; sickness absence; work participation

XSL•FO
RenderX

Introduction

When developing work-oriented rehabilitation interventions, it is crucial to consider the experiences of stakeholders in their working life. Understanding barriers and facilitators can enhance successful implementation of such interventions into practice [1,2]. Over the past 30 years, the global incidence of cancer among individuals younger than 50 years has risen by 79%. This trend is particularly pronounced in high-income countries, influenced by lifestyle factors and enhanced diagnostic methods [3,4]. Concurrently, advancement in cancer treatment has markedly improved survival rates for many cancer types, especially those affecting working-age populations [5]. In Sweden, it is noteworthy that almost half of all cancer cases are diagnosed during people's working years [6]. Consequently, cancer survivors, individuals diagnosed with cancer who are still alive, are likely to remain active on the labor market for many years to come [7]. This reality underscores the importance of facilitating return to work (RTW) postdiagnosis [8-11]. Furthermore, cancer diagnoses have significant economic consequences for both individuals and society. Patients frequently experience considerable financial strain due to lost income from prolonged sick leave or reduced working hours [9]. On a societal level, the economic impact of cancer-related work absence in Sweden is substantial, estimated at approximately US $4 million annually, which is comparable with the cost associated with outpatient cancer care [12].

Although cancer treatment has improved survival, many survivors experience long-term physical, cognitive, and psychological effects that can hinder individuals' work ability. Although more cancer survivors return to work [13], many face functional limitations [14,15], and about 25% of cancer survivors did not return even 14 years postdiagnosis [16]. Among those who do not return, many work part-time [17] or attribute their absence to mental illness, chronic pain, or musculoskeletal disorders [18]. In the Nordic countries, 14% of cancer survivors received disability pension 7 years after diagnosis [7], and many reported reduced work ability (K Tödt et al, unpublished data, 2026). Previous studies of common cancer types in Sweden, including breast, colorectal, and prostate cancer, have demonstrated that many survivors experience prolonged work absence following diagnosis and treatment. More recent evidence from a nationwide registered-based prospective cohort study showed that, even after returning to work, cancer survivors had a substantially higher risk of future labor market marginalization [19]. Specifically, they exhibited a 1.6-fold increased risk of new sick leave episodes, a 2-fold higher risk of receiving a disability pension, and a 1.3-fold increased risk of unemployment compared with matched individuals with no cancer [19]. These findings underscore the need for a comprehensive rehabilitation approach, viewing RTW as a complex process requiring coordinated medical, psychosocial, and work-oriented rehabilitation [20].

Prior research has shown that many cancer survivors are motivated to maintain or resume employment during and after treatment [13,21]. Prolonged absence from work is associated with reduced quality of life [22] and may erode dignity, social identity, and financial security [23,24]. These challenges have intensified calls for more effective multiple-stakeholder interventions to support occupational reintegration [8-11].

Motivation to return to work is shaped by both personal and contextual factors. Social support from family, peers, colleagues, and employers reinforces a sense of belonging and security, thereby facilitating labor market reintegration [8,11]. Equally critical are workplace conditions: flexible work arrangements, task modifications, and organizational cultures oriented toward inclusivity and health promotion, which are consistently identified as facilitators of RTW [7,9]. Collectively, the evidence underscores the importance of supportive networks and enabling organizational structures to reclaim their professional role following treatment.

Theories concerning cancer and work emphasize the importance of tailoring rehabilitation interventions to the individual cognitive and emotional resources of survivors, as well as to the specific demands they face related to their illness, treatment, work, and personal circumstances [20]. In non–cancer populations with, for example, musculoskeletal or mental illness disorders, the most promising interventions for RTW were found to be multidisciplinary, involving several stakeholders [11,25]. Reviews have shown positive effects of cancer rehabilitation in general, but several areas remain underdeveloped, particularly regarding RTW [15,26]. Cancer rehabilitation seeks to prevent and alleviate cancer-related consequences [15,26], while work-oriented rehabilitation specifically addresses those affecting work participation [10].

Approaches to cancer rehabilitation and RTW vary considerably across European countries. Differences in health care and social insurance systems, as well as in the distribution of responsibilities among stakeholders, shape both the design of interventions and how they are perceived by employers, health care providers, and insurance agencies [16,27,28]. Research highlights the importance of adapting rehabilitation strategies to each sociopolitical and institutional context, while also addressing the specific needs of cancer survivors and involved stakeholders [8,9,11]. The Swedish context illustrates this well. Employers are liable for wage compensation during the first initial 14 days of sick leave, after which financial benefits are administered by the Swedish Social Insurance Agency. Employers are responsible for workplace-related interventions, whereas the Swedish Social Insurance Agency coordinates actions among relevant authorities and stakeholders.

Important stakeholders regarding work participation include health care professionals in cancer, primary and occupational health care, managers, and social insurance officers [29-31]. Previous research has stressed that cancer survivors often require more comprehensive support from these diverse working-life stakeholders to facilitate a sustainable RTW process [21].

To enhance intersectoral collaboration, the Act on Coordination of Rehabilitation Efforts (Act no. 2019:1297) mandates the establishment of rehabilitation coordination. These professionals, commonly occupational therapists, physiotherapists, social workers, or registered nurses, serve as intermediaries between patients and key actors within the RTW process [32]. Their responsibilities include providing individualized support, ensuring communication across organizational boundaries, and

XSL•FO
RenderX

facilitating the development of personalized rehabilitation plans [33]. At the core of the Swedish RTW system lie systematic assessments of work ability and structured measures aimed at promoting workplace reintegration. Collectively, these arrangements are intended to provide a coordinated and equitable approach to supporting individuals on sick leave.

Although the Act on Coordination of Rehabilitation Efforts requires regions to provide coordinated rehabilitation, these are often delivered through rehabilitation coordinators whose roles and visibility vary across settings [32,33]. Research suggests that coordinators may support RTW processes through continuity and coordination [20], but evidence of their effectiveness in improving RTW outcomes is inconsistent [10]. Although rehabilitation coordinators represent one important stakeholder in RTW processes, this study adopts a broader perspective to capture how facilitators and barriers emerge across multiple interacting actors.

Previous research has identified various facilitators and barriers to RTW after cancer [9,34,35], but much of this evidence is based on heterogeneous contexts [9] and overlooks how institutional conditions shape the RTW processes [16,36]. This is particularly relevant in Sweden, where RTW involves multiple actors across health care, workplaces, and the social insurance systems [17,19]. Key stakeholders' perspectives also remain underexplored [23,31,37], especially regarding how these actors interact in practice [38], an important gap given that RTW is a multiactor process. Recent structural changes, such as the introduction of rehabilitation coordinators, have further reshaped RTW processes [32,33], yet limited empirical knowledge exists about how stakeholders experience these changes [21,39]. Against this backdrop, this study was conducted to generate context-specific, empirical insights into facilitators and barriers to RTW after cancer from a multistakeholder perspective. The study contributes by (1) integrating multiple stakeholder perspectives, and (2) informing the development of a tailored intervention.

Hence, it is essential to further explore and describe the perspectives of stakeholders themselves, as their roles, perceptions, and interactions are central to shaping the conditions under which cancer survivors attempt to resume working life. The aim of this study was to investigate working-life stakeholders' perceptions of both facilitators and barriers to RTW among cancer survivors, including factors for work-related well-being during the RTW process.

Methods

Study Design

This qualitative study was a part of the development phase of a complex intervention following the updated Medical Research Council framework [40] regarding work-oriented rehabilitation. A qualitative interview design was chosen to gain an in-depth understanding of stakeholders' experiences and perceptions of cancer survivors' RTW process. Semistructured interviews were conducted with 25 stakeholders representing various roles in the RTW context during spring 2023. The qualitative interview approach provided flexibility and allowed participants to elaborate on issues of personal and professional relevance [41].

Sample participants were recruited using purposive sampling to ensure variation in characteristics, geographic locations, and work settings (specialized care, primary care, and public and private sector). Purposive sampling involves the intentional selection of information-rich cases to capture diverse experiences related to the studied phenomenon [42]. The sampling aimed to capture variation both between stakeholder groups and within groups comprising multiple roles. For example, the health care professional group included several relevant clinical disciplines (physicians, nurses, occupational therapists, and physiotherapists). The sampling strategy also reflected the stakeholders commonly involved in RTW processes for cancer survivors, resulting in higher representation of health care professionals and representatives from the Swedish Social Insurance Agency, while representation from Employment Agency officers was more limited (n=1). Rehabilitation coordinators were included as one of several stakeholder groups; however, the study was not designed to specifically examine their role but rather to explore facilitators and barriers to RTW across multiple stakeholders.

Participants represented 3 regions in central Sweden, involved in a larger research program developing an intervention to support cancer survivors' RTW. Stakeholders included health care practitioners (registered nurse, occupational therapist, physician, counselor, and rehabilitation coordinator), social insurance officers, employment service agency representatives, and first-line managers from private and public employers. Selection was based on experience supporting cancer survivors in maintaining or returning to work during cancer treatment. Inclusion criteria were employment within a stakeholder group relevant to RTW processes (ie, health care professionals, social insurance officers, employment service officers, social workers, or first-line managers) and experience of supporting cancer survivors in RTW. No specific exclusion criteria were applied; however, all participants were required to be able to participate in an interview conducted in Swedish. The study followed COREQ (Consolidated Criteria for Reporting Qualitative Research) [43].

Setting

The study was conducted within the Swedish RTW system, a tax-funded framework involving health care services, employers, and the Swedish Social Insurance Agency. Employers provide sick pay during the initial 2 weeks of sick leave, after which the Swedish Social Insurance Agency assesses work capacity, administers benefits, and coordinates rehabilitation in collaboration with health care providers. Employers are also responsible for workplace-based rehabilitation, while additional support, such as rehabilitation, may be provided by the Swedish Public Employment Service (Act no. 2019:1297).

RTW processes typically involve ongoing interaction between these stakeholders, where health care providers assess work ability, employers implement workplace adjustments, and the Social Insurance Agency coordinates measures and monitors eligibility for benefits. For individuals with cancer, the system includes specific regulations and coordinates efforts intended

XSL•FO
RenderX

to support a gradual and individualized RTW process, although the level of coordination may vary.

Data Collection

A study coordinator recruited participants in 2 stages. First, brief study information was emailed to key organizational contacts relevant to the study, who were asked to forward the invitation to potential participants. Interested individuals contacted author SLF for further information. Subsequently, comprehensive information was provided, eligibility criteria were verified, and participants were informed about the purpose of the study. Initially, 26 individuals expressed interest in participating; however, 1 participant withdrew prior to the interview due to personal reasons, resulting in 25 completed interviews. Participants provided informed consent and, immediately before the interview, completed a brief web-based survey describing their demographic characteristics (Table 1). Participants' age ranged from 39 to 65 years, and those who were employers had between 8 and 46 employees (median 43).

Table . The participating working-life stakeholders.

Variable	Participants (n=25)
Age (years), median (25th to 75th percentile)	48 (43 - 59)
Gender, n (%)	
Men	4 (16)
Women	21 (84)
Occupational sector, n (%)	
Health care[a]	12 (48)
Social Insurance Agency	7 (28)
Employment Service Agency	1 (4)
Employer[b]	5 (20)
Years of occupational experience[c], median (25th-75th percentile)	18 (5 - 25)
Frequency of working with people with cancer, n (%)	
Several times a day	7 (28)
Once a day	2 (8)
Once a week	4 (16)
Once a month	4 (16)
Several times a year	4[d] (16)
Less often than once a year	4[e] (16)
Knowledge about the national guideline recommendations on cancer rehabilitation[f], n (%)	
No, I don't know about it	6 (24)
Yes, but I have not read it	14 (56)
Yes, I have read it	5 (20)

[a]Their occupations were nurse (n=3), occupational therapist (n=3), physician (n=3), counselor (n=1), and rehabilitation coordinator (n=2). The health care settings were specialized care (n=9) (eg, oncology: n=4), primary care (n=2), and occupational health care (n=1).

[b]The employers worked in the public sector (n=3), private sector in a middle-sized company (n=1), or private sector in a big company (n=1).

[c]In their present occupation.

[d]Being health care professionals within primary care or rehabilitation clinics.

[e]Being employers.

[f]Distributed by the Confederation of Regional Cancer Centres [44].

The authors SLF, LE, PF, and MF interviewed the stakeholders individually, one-on-one in a video meeting. The interviewers comprised a district nurse, an occupational therapist, and 2 oncology nurses; 3 were women and 1 was a man. The interviewers had both professional experience in clinical cancer care and experience in conducting qualitative interviews. A semistructured interview guide was followed (Textbox 1), asking follow-up questions as needed to deepen or expand responses, in line with qualitative interview methodology [41]. Interviews were audio-recorded, lasting between 20 and 71 minutes (median 42 minutes).

XSL•FO
RenderX

Textbox 1. The semistructured interview guide.

At the beginning of the interview, the interviewer introduced themselves and briefly described their professional background. The interviewer presented a case study based on a fictional patient "Kim" on whom the informants were asked to base their reasoning. The first questions were connected to this case, and the case description formed the basis for the interview. In the last questions of the interview, we moved away from the Kim case and asked interviewees to talk more generally about the questions. When necessary, field notes were made. A detailed description of the case used during the interviews is provided in Multimedia Appendix 1.

Interview questions

- What challenges do you anticipate regarding work in relation to having cancer?
- Could you tell me if you have a role in supporting individuals' work and occupational health during or after cancer treatment, and if so, what does this involve?
- Can you tell me which other actors have roles here?
- What do you believe could promote work and health here, and what do you believe could hinder work and health?
- Does Kim, in this example, need any support regarding work and health? If so, please describe it.

Moving beyond the case of Kim

- If anything were possible, what kind of support would need to exist to support people's work and health during and after treatment?
- As a final question, we have now discussed...(your role, other stakeholders, collaboration, and support related to work and health). Is there anything else you would like to add regarding work or health in relation to work for individuals during or after cancer treatment?

Data Analysis

According to Elo and Kyngäs [45], the interviews were transcribed verbatim and analyzed using qualitative content analysis. Initially, transcripts were read several times by authors SLF, KW, and MF to gain familiarity. No dedicated qualitative analysis software was used. Instead, data were managed in Microsoft Excel to organize transcripts, track codes, and collate categories. Data handling followed strict standards of integrity and confidentiality.

Transcripts were evenly distributed among these authors, who independently extracted meaning units relevant to the study aim and generated initial categories and subcategories. Subsequently, these authors collaboratively reviewed their analyses, refined categories and subcategories, and reached consensus on findings. All 25 interviews were analyzed as a single dataset. No separate analyses were conducted by stakeholder groups; instead, themes were developed based on patterns identified across the entire material. During coding, sector-related nuances were noted and considered in the interpretation, but no formal comparison between stakeholder groups was performed. An example of the analytic structure is illustrated by the data unit "Most people want to work. I have to say it's an extremely exciting experience working with people, that everyone is so eager to get back to work. I think so." which was coded as "Most people want to work," grouped into the subcategory "Desire to work," and further abstracted into the category "Emotional significance of work." Trustworthiness was ensured through analyst triangulation to enhance credibility and reduce interpretative bias [42]. Dependability and confirmability were supported through transparent documentation and regular discussion within the research team [46]. Reflexivity and methodological transparency were emphasized to strengthen credibility and transferability, in line with qualitative research standards [41]. The research team included members with backgrounds in nursing, health-related research, and occupational therapy, which contributed to multiple professional perspectives during data interpretation. Reflexive discussions throughout the analytic process increased awareness of potential preunderstandings and supported interpretative rigor.

Ethical Considerations

The study was approved by the Swedish Ethical Review Authority (registration numbers 2022-03686-01, 2023-05671-02) and conducted in compliance with the Declaration of Helsinki (World Medical Association, 2013). Informed consent was obtained from all participants. Participants did not receive any financial or material compensation for their participation.

Results

Overview

Barriers and facilitating factors regarding RTW were identified not only in the individual adaptations made by cancer survivors but also in cooperation among stakeholders. Structural issues, including financial obstacles and sick leave regulations, also influenced RTW outcomes. Highlighted factors promoting work-related well-being were flexibility and personalized adaptations. Categories and subcategories are presented in Table 2.

XSL•FO
RenderX

Table . Categories and subcategories of stakeholders' perceptions of both facilitators and barriers to work among cancer survivors in a return to work process.

Category	Subcategory
Collaboration and clear division of responsibilities	• Desire for developed and improved cooperation • Who is responsible for what? • Clear division of responsibilities
Balancing individual adaptations	• Adaptations to the work situation • Flexibility in RTW[a] • A balancing act
Reducing structural barriers through support	• Financial obstacles to RTW • Challenging sick leave bureaucracy
Views and expectations of the individual effect of RTW	• Courage to push and set demands • Inclusion in the workplace
Emotional significance of work	• Desire to work • Existential questions take focus

[a]RTW: return to work.

Collaboration and Clear Division of Responsibilities

The interviews revealed not only a strong desire for more developed and improved collaboration among stakeholders but also a lack of awareness about available support and confusion over roles and responsibilities. Stakeholders emphasized the need for a structured collaboration plan to clarify roles and enhance teamwork. This, according to them, would facilitate collaboration and ensure that each party had a clear understanding of their respective roles and responsibilities.

Desire for Developed and Improved Cooperation

The informants stressed that cooperation between stakeholders sometimes had been lacking or functioned poorly, negatively affecting the RTW process. They described an impermeable silo mentality among stakeholders who were expected to collaborate, causing irritation and frustration over this lack of coordination, emphasizing that it had hindered efficiency in helping individuals return to work.

> *I often get nasty answers and a feeling they think I'm not knowledgeable and that I disrupt contact with healthcare. That can be a barrier to returning to work.* [12, Social Insurance Agency]

Throughout the interviews, a desire to increase and improve cooperation emerged. Informants strongly believed that closer collaboration is crucial for addressing the challenges and would lead to better, more coordinated outcomes. They noted that greater cooperation would have benefited cancer survivors' RTW and resulted in better, faster, and more cost-effective care. Furthermore, informants also argued for a unified body to oversee the overall coordination.

They felt that moving away from siloed work structures toward team-based collaboration would be better for everyone and more cost-effective for society. Teamwork was seen as promoting continuity and reducing vulnerability, since remaining members could ensure stability even if 1 key person was replaced.

> *We need more cooperation in teams. The healthcare system would save a lot of money if we got out of our silos, if we worked in teams, which I've had the benefit of doing at times. Because then you really can help people, compared to getting help from a physical therapist here, then a doctor and then a nurse, an occupational therapist, a psychologist. And then they don't talk to each other.* [6, occupational therapist]

Informants working in inpatient hospital care described well-functioning collaborations and established team models, but in other areas there was a total lack of collaboration between specialties. They emphasized the importance of enhanced collaboration, suggesting specialized outpatient teams with in-depth knowledge of diseases, as this might be difficult for primary care to manage on their own. They also recommended appointing a dedicated coordinator to oversee and support the entire RTW process and stressed the need for clear consensus and direction from higher leadership levels. Rehabilitation coordinators were highlighted as important facilitators of collaboration and serve as a link supporting cancer survivors' RTW.

Who Is Responsible for What?

Among the informants, there was a perceived lack of knowledge about who was responsible for what, which they believed hindered the RTW process. More broadly, they noted an insufficient understanding of the tasks and roles of different professional groups and their potential contributions. This lack of clarity and knowledge was seen as a cause of failed initiatives, thereby negatively affecting the process. Furthermore, interviews revealed a lack of trust in the competence of certain individuals expected to participate in the sick leave process.

> *...something I wonder about is that the person who's supposed to decide on the sickness benefit isn't required to have any knowledge at all of healthcare but is responsible for coordination.* [19, manager]

XSL•FO
RenderX

Informants emphasized frequent mistreatment by the Social Insurance Agency. Some described such negative experiences that they questioned the agency's role and purpose in the sick leave process. They believed that this situation imposed further barriers to the RTW process for cancer survivors.

> *The collaboration between actors is dysfunctional—the Social Insurance Agency ignores us when we contact them; I'm not sure what they're good for.* [10, physician]

On the other hand, it was mentioned on several occasions that Social Insurance Agency employees felt that they were treated poorly, particularly by doctors.

> *Well, it's just that you feel you don't know anything and that you're a bother. Unfortunately, with time I've gotten quite a few sarcastic answers when I've...or supplemented a doctor's certificate that I should have understood.* [12, Social Insurance Agency]

They described experiencing a sense of being a disturbance and sometimes having been given substandard sick leave documentation to manage. When they asked supplementary questions, they felt that their inquiries were considered foolish and that they were expected to know more.

Clear Division of Responsibilities

During the interviews, the informants reported a lack of division of responsibilities and expressed a desire for greater clarity around who was responsible for what. They emphasized that lack of clarity regarding division of responsibilities prolonged the RTW process or, in some cases, prevented it altogether.

> *I said she should do work training, but her employer was on vacation, so I had to put her on full-time sick leave // --- // I had to change her sick leave several times. If the Social Insurance Agency had contacted me instead, it would have been much simpler than having a patient try to explain.* [8, physician]

There was a desire for a holistic approach to RTW that also considered social aspects. Doctors reported having control over medical issues but noted that the social perspective received far less attention due to time constraints. They believed that the process could have been facilitated and improved if a more holistic approach had been applied. Informants repeatedly stated that a specialized unit focusing on RTW for this group would have facilitated the process. The need for such a unit was further emphasized by those who described the situation as complex. They explained that doctors were sometimes expected to issue sickness certificates even when they were no longer involved in the patient's treatment and had limited knowledge of, or contact with, the patient.

> *Then there's another aspect I think is sometimes a bit complicated. It's that a lot of follow-ups occur through me, where they don't have a doctor to contact, but anyway it's the doctor who's supposed to be behind the sick leave, which can sometimes cause confusion.* [2, occupational therapist]

To improve efficiency and clarify responsibilities, informants suggested involving certain functions earlier in the process and assigning them a more defined role as a consistent link throughout the process. They emphasized the need for a dedicated person to follow up with the patient beyond the acute stage, indicating that the rehabilitation coordinator was well suited for this role.

Balancing Individual Adaptations

The stakeholders reported that a flexible RTW approach was required to facilitate cancer survivors' RTW process during and after cancer treatment, where tasks were adapted to cancer survivors' abilities and needs. They said the process can be compared with a balancing act that requires continual adjustment in how it is set up.

Adaptations to the Work Situation

To facilitate a smooth and sustainable RTW process, informants reported tailoring the work situation to individuals' needs. They emphasized adapting tasks and establishing a plan that ensures long-term sustainability for cancer survivors. Gradually increasing workload, based on a plan developed with the cancer survivor, was seen as essential. They also stated that adaptations may be necessary if cancer survivors are unable to handle excessive sensory input or work in large groups.

Informants noted that it is the manager's responsibility to modify the work situation and ensure a workable solution. Specific adaptations could include adjusted working hours or a condensed schedule for additional recovery time. Informants indicated that, although most employers recognized the need and were willing to make individual adjustments, ongoing challenges persisted. The time required to identify alternative solutions for employees varied, and in recent years the process has sometimes been more challenging and restrictive. Some informants perceived a reluctance among employers to adapt to cancer survivors' working situation.

> *Maybe returning to work full-time isn't doable, so instead you get to see if there are alternative solutions. But you're always thinking along with colleagues and maybe others, so that you always have a plan, continuous conversations, follow-ups, so that all parties are agreed on the direction we're going in.* [21, manager]

There were different opinions about how flexible the system is in helping cancer survivors return to work after treatment. Some felt that making individual adjustments had become easier, while others thought that it had become more difficult due to changes in legislation. The bottom line, however, is that everyone felt that being able to make flexible adjustments was important.

Flexibility in RTW

Although work tasks were adapted to cancer survivors' needs, informants noted that plans often required rebalancing as the survivors' requirements changed after returning to work. Flexible tasks include the possibility to take short breaks during the day, while adjustment and shielding were sometimes necessary for those with treatment-related immune weakness,

XSL•FO
RenderX

increasing vulnerability to infection. Consequently, cancer survivors might need to influence aspects of their work situation to ensure their own health and safety.

> *Well, sometimes people's immune system is suppressed due to treatment, or lack of treatment. So, then we've...Well, not discouraged her, but said you can do that from home during this period, that would be good, to avoid possible risks, like.* [21, manager]

Employees may also have the flexibility to be at the workplace at times that suit them, with the option to work remotely and participate in meetings virtually.

A Balancing Act

Regarding the RTW process, the informants emphasized that it is a balancing act that affects how successful RTW and the process will be. They noted that planning is easy, but that checking progress along the way is often neglected. That neglect often undermines a sustainable RTW. Without regular check-ins, the likelihood of a successful transition decreases. It is crucial to navigate the process by finding the right balance between pushing forward and pulling back when needed.

> *In my experience, it's easy to push things and go too fast.* [19, manager]

The informants found it important to be aware of physically and mentally demanding tasks and to monitor cancer survivors' progress, checking how they manage these responsibilities. Furthermore, they emphasized the need to avoid putting excessive pressure on cancer survivors, while still valuing and using their skills and expertise. Finding this balance is essential and plays a significant role in ensuring a sustainable RTW.

Reducing Structural Barriers Through Support

One area the interviewees felt had worsened concerned the structural barriers that could arise in connection with RTW. Among other things, they thought that regulation of working hours for RTW was a problem. They also pointed to the financial situation that could arise when employees wanted to return to work but for various reasons were unable to perform their regular tasks, as this could lead to increased costs. Furthermore, they stressed that lack of qualified access was an obstacle to creating a "smooth" RTW process.

Financial Obstacles to Returning to Work

Informants reported that patients' financial situation could be a barrier to RTW. A key challenge was that employees were often unable to perform their pre-illness tasks. Even when employees could not carry out their duties fully, employers were still required to cover the costs, creating difficult circumstances. This could result in a situation where patients wanted to return to work, but employers could not afford to accommodate them. Conflicts sometimes arose when employees believed that they could work while employers doubted their ability.

Refusals or dismissals in such cases were described as more common among larger employers who, despite greater resources, still struggle to make necessary adjustments. Similarly, physically demanding occupations, such as industrial work, often faced considerable challenges in providing suitable accommodation. Nonetheless, many employers, despite lacking financial support, allowed employees to return even when not fully able to work, demonstrating notable commitment and fostering a more inclusive work environment.

> *I think it's pretty great that so many employers take them back, because they don't gain anything; they're just helping out. People need different degrees of support, but just having a supervisor, a specific person, I think that's a good foundation for things working out.* [18, Employment Service Agency]

One proposal that emerged in the interviews was that employers who accepted employees who were not yet fully able to work should receive financial support. The informants emphasized the importance of individuals feeling needed and having a role, without feeling like a financial burden on the employer.

Challenging the Sick Leave Bureaucracy

There was a broad consensus among informants that simplicity and flexibility in sick leave regulations were health-promoting, although not all agree. While many acknowledged recent simplification efforts by the Social Insurance Agency, others remained critical. They also highlighted the difficulties cancer survivors face when dealing with RTW regulations. One example concerned the requirement to distribute reduced working hours evenly across the week, which they viewed as impractical for cancer survivors able to work only brief periods; for example, 25% of working time equated to less than 2 hours a day. The informants described the regulation as inflexible in relation to fluctuating health and argued that RTW would be facilitated if each cancer survivor could determine their own scheduling. The regulation was considered complicated, and its lack of flexibility could cause cancer survivors to choose full-time sick leave instead of trying to balance work with sick leave.

> *Today it's possible to be on sick leave part-time and work part-time, but then every time the person gets worse, a new application for a change in sick leave is needed and that can...there's so much administration around it, even the person has to keep track, so in the end you don't do it.* [14, Social Insurance Agency]

Informants noted that sick leave was generously granted during the first year, with minimal administrative demands to avoid burdening cancer survivors. Few questions were asked by the Swedish Social Insurance Agency during that period, and the prevailing attitude was that cancer survivors should not be bothered with routine bureaucracy, which could add unnecessary stress during an already challenging time. There were also certain circumstances that allowed cancer survivors to be on sick leave for special reasons. Furthermore, if the treatment is ongoing, this patient group does not need to be assessed for all work in Sweden.

> *As long as cancer treatment is ongoing, the person doesn't have to be assessed in relation to all jobs in Sweden. So that exception we have is great.* [16, Social Insurance Agency]

XSL•FO
RenderX

The advantages and disadvantages of the ease of getting sick leave during the first year were discussed as something that could prolong sick leave. However, they believed that the real problem comes after patients have been on sick leave for 550 days and are fully covered by social insurance. At this point, cancer survivors need to be matched against the entire labor market, which can be a problem for many.

Views and Expectations of the Individual Effect of RTW

In the interviews, stakeholders on the labor market emphasized that, in several respects, individual conditions affect RTW. They said that this was a matter of providing support, having empathy, and calibrating when a push is needed to help the individual summon up the strength or courage to return to work. It is also about helping the individual feel included in the social interaction that a workplace can provide.

Courage to Push and Make Demands

There was a perception that individuals on sick leave sometimes needed encouragement to avoid getting stuck in their situation. If people have been on sick leave for a long time and have completed all treatment phases, they might need a gentle push out of their comfort zone to regain the confidence to return to work. This encouragement was seen as crucial in determining whether the individual would dare to take the step back into working life.

> *And then the healthcare system plays a very important role and then maybe they need to be the ones who push patients a bit "but now you need to start thinking about your job." Instead of prolonging and prolonging and prolonging the sick leave period.* [16, Social Insurance Agency]

The informants reported finding it easy to fall into the role of feeling sorry for the person on sick leave. However, they did not believe that this was beneficial to people on sick leave, as they could easily adopt that role. Therefore, they emphasized the importance of not allowing the person on sick leave to remain in that condition for too long, as it could lead to passivity. It was mentioned that laying out a clear plan for RTW, with realistic requirements, was beneficial. The informants believed that applying pressure was necessary, although it should be well balanced and based on the patient's situation and well-being.

> *I think having support, understanding and a little push from people in a person's life is helpful. But it should take into account the treatment stage and how difficult the treatment is.* [12, Social Insurance Agency]

The informants also emphasized the need to be less afraid of providing that extra push. This approach, they noted, should be carried out with care and consideration for the patient's individual conditions, while still daring to challenge and motivate the patient when appropriate.

Inclusion in the Workplace

The informants discussed the importance of including cancer survivors in the workplace as part of health promotion work. They described various approaches to achieving this. Some informants described a long and balanced process to maintain social contact with the person on sick leave, such as regular monthly check-ins via SMS text messages to later invite them to lunch or breakfast. Over time, they might encourage the individual to visit the workplace and participate in social events, for example, by attending a shared 30-minute coffee break.

> *We met regularly. At first maybe every three months or so // --- // And then it was more like: How are things going? How do you feel? When do we think it's time to start working?* [22, manager]

The informants emphasized the importance of people on sick leave maintaining contact with their work group and participating in social activities. Otherwise, they noted, it could feel like walls were being built, making it difficult to reconnect after a long absence. Managers also described how they would check in with the person on sick leave beforehand to understand their preferences regarding communication about their cancer diagnosis. They had experience not only with coworkers who wanted to talk openly about their diagnosis but also with those who found it too difficult to discuss in social settings. By gathering this information, managers could help prepare other employees for interacting in a way that aligned with the sick person's wishes. They highlighted the need for continuous adaptation, as preferences varied across individuals.

> *I've sometimes just said "but can't you just go back right after treatment and bring some cake and sit in the breakroom for 30 minutes just to get it over with?" so your first day back at work doesn't mean working and taking care of everything. // --- // Most people who keep in touch throughout the whole period find it easier to return.* [7, physician]

The informants also reported that challenging situations could arise, for example, when a replacement had to be brought in for the person on sick leave. This could make the person on sick leave feel threatened or uncomfortable. The informants stressed the great importance of not forgetting the person on sick leave but instead ensuring that they felt included in what was happening at work, even if they were physically unable to be there. They believed that this was a big step toward making the road back to work short when that time came.

Emotional Significance of Work

In several interviews, stakeholders on the labor market described how individuals' experiences and attitudes toward returning to work are, in many respects, based on emotions. This included the feeling of wanting to work and the social interaction that work can provide. Several concerns were also raised, such as the fear of working or not having the strength to do so.

The Desire to Work

There was a general opinion among informants that cancer survivors had a strong desire to work. In many cases, this was described as a deep-seated and significant aspiration. The informants reasoned that work was not only a source of income, but it also met important social and psychological needs. Being able to continue working creates a sense of belonging,

XSL•FO
RenderX

strengthens identity, and promotes an experience of belonging to a group and being part of the healthy world.

> *Most people want to work. I must say that's a really exciting lesson I learned from working with people, that everyone is so eager to get back to work. In my experience.* [2, occupational therapist]

The informants reported that most people who wanted to continue working or return to work after a cancer diagnosis often succeeded, even if the process was challenging. They said that it could be particularly difficult for those who had more physically demanding work compared with those who had lighter physical jobs. Moreover, they indicated that the status of the profession was important for the possibility of returning to work, although this aspect was not clear-cut.

Existential Questions Come Into Focus

The informants discussed existential questions as factors that could both promote and hinder RTW. The period following treatment was described as challenging for many patients with cancer. In addition to the physical challenges that could arise, existential dimensions could influence an individual's ability to return to work. These dimensions included questioning the very purpose of work, feeling that work was meaningless, or that the job lacked purpose when you were facing uncertainty about your survival.

> *Well, if you get a disease, cancer or something else resulting in sick leave, then the person gets caught up in that process, mostly alone, accepts the disease, but is maybe afraid of not surviving. That's a limiting aspect.* [20, manager]

In addition to existential concerns, the informants discussed the fact that cancer survivors might fear being infected by colleagues. During treatment, cancer survivors may have a weakened immune system, making the fear of infection very real, as an infection could jeopardize their ability to continue treatment. To address this, stakeholders reported adjusting work arrangements to minimize contacts with others, including enabling remote work. Such measures reduce commuting, lower infection risk, and allow survivors to manage their workload according to their capacity.

Discussion

Principal Findings

This study underscores that the RTW process for cancer survivors is a complex and multifaceted challenge, requiring coordinated efforts across the health care, insurance, and employment sectors. The interview findings reveal several key factors that act as either facilitators or barriers to successful RTW. Notably, the significance of interdisciplinary collaboration, clearly defined roles and responsibilities, and flexible support structures emerged as critical areas with considerable potential for improvement. However, these areas also pose challenges due to the different perspectives and priorities among stakeholders.

Building on this overall complexity, the findings underscore the need for enhanced collaboration and clearer role definition within the RTW process for cancer survivors, and in the planning of future RTW interventions. Despite a shared consensus among stakeholders to support individuals in their transition back to work, systemic barriers such as siloed organizational structures, ambiguous responsibilities, and insufficient communication were perceived to impede effective coordination. Similar challenges have been noted in other contexts. Svärd et al [38] observed that stakeholders in RTW processes for individuals with common mental disorders often operate under competing goals and ambiguous mandates, leading to fragmented efforts. In this study, stakeholders expressed expectations of other actors while remaining uncertain about their own responsibilities. This finding is noteworthy given Sweden's rehabilitation coordinator role, designed to support individuals on sick leave and facilitate coordination among stakeholders, including health care professionals, employers, and social insurance officers. Strengthening the mandate and visibility of these coordinators may therefore improve coordination outcomes, and comparable roles could be valuable in other national contexts. Désiron et al [39] likewise emphasized that such coordinators can provide tailored support, maintain continuity, and help bridge gaps between stakeholders.

The reported lack of coordination and cooperation between stakeholders may be linked to how rehabilitation coordinators have been implemented in Sweden. Although such functions have been introduced and are present in many settings, previous research has indicated that their role is not yet consistently established or fully integrated into RTW processes [32,33]. Barriers such as unclear mandates, variable recognition among stakeholders, and limited resources may restrict their ability to fulfill their intended coordinating function [32,33]. This may partly explain why participants in this study continued to describe persistent gaps in coordination.

More broadly, previous research has emphasized the importance of facilitators in changing processes [1]. Strengthening collaboration and clarifying the division of responsibilities among stakeholders are essential for ensuring a smoother RTW experience. For instance, Lau et al [36] highlight that siloed health care systems impede information exchange and collaborative evidence generation, ultimately undermining patient outcomes and system efficiency. The lack of integration among health care providers, social insurance agencies, and employers results in duplicated efforts, delayed interventions, and increased frustration among professionals. In our study, participants consistently emphasized the importance of teamwork and interprofessional collaboration. This statement aligns with findings from Désiron et al [39], who advocated for a dedicated RTW coordinator to improve work-oriented rehabilitation interventions. In a previous interview study by Bilodeau et al [37], employer representatives' perspectives on supporting breast cancer survivors in their RTW process were explored. These findings closely align with stakeholders' perceptions in our study, reinforcing the notion that collaboration challenges, unclear responsibilities, and communication barriers are prevalent across contexts. The highlighted issues of stakeholder collaboration reflect our observations regarding fragmented efforts and unclear mandates.

XSL•FO
RenderX

In addition to these coordination challenges, stakeholders perceived that structural barriers, such as financial challenges and sick leave regulations, must be addressed when promoting cancer survivors' RTW. Financial constraints often discouraged employers from accommodating employees who were not fully productive, reflecting the findings of Fitch and Nicoll [47], who identified reduced income as a significant adverse consequence for cancer survivors. Financial difficulties are often associated with increased psychological distress and reduced quality of life [48].

Employers frequently struggled with balancing economic considerations by providing suitable workplace adjustments, resulting in conflicts, redeployment, or even unemployment for cancer survivors. The complexity and inflexibility of sick leave regulations were also identified as critical issues. Cancer survivors faced significant administrative burdens when adjusting their sick leave arrangements, particularly regarding rules that mandated even distribution of working hours across the week. Such rigid regulations often discouraged survivors from attempting gradual RTW, potentially prolonging full-time sick leave unnecessarily. To address these barriers, simplifying sick leave regulations and providing clear, adaptable guidelines are essential. Additionally, introducing financial incentives for employers who facilitate flexible and partial RTW arrangements could encourage greater inclusivity and reduce economic pressures. These measures could significantly enhance the feasibility of accommodating cancer survivors in workplaces, promoting healthier, sustainable RTW.

Our findings underscore the importance of tailoring work conditions to cancer survivors' individual needs during RTW. Key supportive factors were flexible, personalized arrangements; gradual workload increases; remote work options; and adaptable schedules. Limited, nonpressured testing of work capacity was particularly valued for balancing health and rebuilding confidence. Stakeholders across professions emphasized individualized support, reflecting the complexity of recovery. This aligns with López-Faneca et al [35], who highlight the value of personalized RTW programs, and Bilodeau et al [37], who emphasize humanistic management and flexibility. They also stress the importance of addressing unclear responsibilities and expectation gaps, as these may hinder coordination. Together, these findings underscore the need for structured measures such as clear communication, emotional support, respect for privacy, and tailored interventions. Overall, our findings reinforce the need for individualized, flexible, and coordinated RTW approaches to promote recovery, sustainable employment, and workplace inclusion.

Beyond workplace adaptations, stakeholders considered cancer type, health status, treatment duration, and family circumstances as influential for cancer survivors' motivation and ability to return to work. Personal attitudes, expectations, and the emotional meaning of work were also emphasized. A workplace where individuals feel included and not defined by illness was viewed as essential, and encouraging workplace contact and inclusion was seen as an important aspect of stakeholder support, aligning with reviews indicating broad consensus that RTW should be encouraged when health permits [49]. Stakeholders further highlighted health-promoting factors, such as social belonging and adjusted workload, as important for sustainable RTW. This is in line with a previous review evaluating the importance of personal and social factors to sustainable RTW after ill health due to mental disorders or musculoskeletal disorders [34]. Furthermore, stakeholders stressed the need to gently give cancer survivors a push back to work, always with the cancer survivors' well-being in mind. One study showed that strengthening people's self-efficacy regarding RTW is an important factor for successful and more rapid RTW [50]. Earlier research has also shed light on the perceived importance of maintaining contact with work during sickness absence [51]. However, uncertainty about when to initiate communicating with employees and the fear of exacerbating cancer survivors' condition by making contact were described by stakeholders in earlier research and perceived as barriers to RTW [51].

Extending these findings, the results showed that stakeholders clearly perceived that RTW can promote multiple dimensions of well-being for cancer survivors on sick leave. Social benefits include reestablishing connections with colleagues, a sense of belonging, and receiving collective support, while economic benefits involve transitioning from sick leave benefits to earned income. This multifaceted value of RTW is an important finding in this study. However, stakeholders may need to consider that RTW can serve as a coping mechanism for managing existential reflections associated with a potentially life-threatening illness. The transition back to work can help survivors regain a sense of normalcy and control [52], yet it may also mask underlying emotional distress [53]. Research shows that cancer survivors often experience heightened anxiety and depression after treatment ends, partly due to reduced contact with health care providers and lingering fears of recurrence [13]. These psychological challenges can be exacerbated in cases where recovery is delayed or complicated, increasing the risk of emotional setbacks. Therefore, RTW interventions must address both physical and emotional dimensions of recovery [35] and involve stakeholders not only from the health care and the Social Insurance Agency but also from employers, who are generally not involved in developing RTW interventions [54].

Taken together, these findings are in line with previous research on RTW across a range of chronic conditions [7,25,34], which consistently highlights the importance of coordination, communication, and role clarity among stakeholders involved in the RTW process [20,38]. Similar barriers, such as fragmented responsibilities, limited collaboration between organizations, and challenges in navigating complex systems, have been described in studies of other chronic disorders [36,50,54], suggesting that many of the identified themes reflect more general structural and organizational challenges in RTW processes.

At the same time, some aspects appear to be more specific to cancer. These include the unpredictability of the disease trajectory [9,53], the impact of treatment-related side effects such as fatigue and cognitive impairments [13,20,52], and the need for flexible and individually adapted RTW processes over time [10,21]. Such factors may place additional demands on both coordination and workplace adaptations and highlight the importance of addressing cancer-related needs within RTW interventions. Taken together, these findings suggest that while

XSL•FO
RenderX

several barriers and facilitators are shared across diagnostic groups, effective RTW support for cancer survivors also requires attention to condition-specific challenges [22,34].

The findings also align with the model by Feuerstein et al [20], which emphasizes the need to tailor rehabilitation to the individual cognitive and emotional resources of cancer survivors and to the contextual demands they face. Although the analysis was inductive and not theoretically driven, the results underscore the importance of coordinated, individualized support across multiple systems. The complexity of the RTW process—requiring interdisciplinary collaboration, clearly defined roles, and flexible support structures—reflects the view of Feuerstein et al [20] that effective rehabilitation must integrate both personal and structural factors. Framed within this model, the study advances understanding of how rehabilitation practices and policies can better support survivors' transition back to work.

Finally, the findings of this study have also informed the development of a planned intervention. In particular, the identified challenges related to coordination and communication suggest a need to strengthen structured collaboration between stakeholders. The importance of role clarity highlights the need to define responsibilities more explicitly within the intervention. Furthermore, the identified need for individualized and flexible support indicates that the intervention should allow for adaptation to individual circumstances and changing needs over time. Finally, the findings underline the importance of addressing both workplace-related and health care–related aspects within a coordinated framework.

Strengths and Limitations

The study has several strengths, including the methodological approach, which was considered appropriate for understanding and describing perceptions. In addition, a purposive sampling strategy and multiregional representation ensured broad stakeholder perspectives and increased contextual diversity and potential transferability. The sample size was considered sufficient, reflected in the rich interview data. One explanation could be the use of a fictional case during semistructured interviews. This may have helped the participants reflect on realistic scenarios, while the semistructured format allowed for depth in responses. One limitation of the study is that some stakeholder groups were underrepresented, which may have limited the diversity and comprehensiveness of perspectives. The sample comprised fewer participants from the Employment Service Agency compared with other stakeholder groups, which may limit the depth of insights from this sector. However, this may reflect the limited involvement of this stakeholder group in RTW processes for cancer survivors within the Swedish system. Nevertheless, the sampling strategy was designed to maximize variation across key roles rather than to ensure equal group sizes, and we gathered substantial input from representatives in health care, social insurance, and employer organizations. Moreover, stakeholders were asked about cancer survivors as if they were a homogeneous group. This may be a limitation, because although rehabilitation and the conditions for RTW after cancer vary depending on the cancer diagnosis and severity, the tasks to be performed at work also vary.

A further limitation of the study is that the study did not specifically investigate stakeholders' experiences of rehabilitation coordinators or their expectations regarding this function. Given the ongoing implementation of coordination efforts, future research should examine how rehabilitation coordination is perceived and how it can be further developed to support RTW processes.

In addition, cancer survivors were not included as participants. As key actors in RTW processes, their perspectives are essential for a comprehensive understanding. However, the study was designed to focus on stakeholders in working life, and the findings should be interpreted in light of this scope. Previous studies by our research group have explored cancer survivors' perspectives, which complement the present findings.

Implications and Further Research

Building on these findings, a clearer definition of roles and shared responsibilities among working-life stakeholders is essential to effectively support the RTW process among cancer survivors. Flexibility within the RTW process is crucial for enabling individuals to resume work at their own pace, facilitating their transition out of illness-related isolation, and promoting social interaction with colleagues. These are important aspects to consider in future policy development, interventions, and support measures. Additional research is needed to explore cancer survivors' personal RTW experiences, with the goal of obtaining a more comprehensive understanding of the factors that influence the RTW process. To support interpretation and guide future research, Table 3 summarizes the implications of the findings for the Swedish RTW system.

XSL•FO
RenderX

Table . Summary of the implications of the findings for the return to work system.

Findings	Implications
A desire for collaboration and clear division of responsibilities	• Ensure that each stakeholder has a clear understanding of their respective roles and responsibilities. • Ensure clarity regarding the division of responsibilities. • Appoint a dedicated person to oversee and support the entire RTW[a] process.
Balancing individual adaptations	• Gradually increase workload tailored to needs and work conditions. • Flexible RTW plan, developed with the cancer survivor. • Allow, if possible, adaptations to the work situation.
Reducing structural barriers through support	• Stakeholders emphasized the need for a structured collaboration plan to clarify roles and enhance teamwork. • Promote awareness about available support. • Check progress along the way.

[a]RTW: return to work.

Conclusions

This study highlights strategies that facilitate a successful RTW process among cancer survivors from the perspectives of working-life stakeholders. The findings underline the importance of individualized and tailored support, as well as the need for clear policies to guide stakeholders' roles and responsibilities. Clarifying the rehabilitation coordinator's role in improving stakeholder cooperation, and enhancing knowledge of cancer survivors' specific workplace needs, can collectively facilitate a smoother and more sustainable RTW transition.

Acknowledgments

The authors acknowledge the use of artificial intelligence (AI) tools, including ChatGPT and Bing, solely for enhancing the English language and correcting grammar in the manuscript. No AI tools were used for data collection, analyses, or scientific interpretation. The authors retain full responsibility for the manuscript's content. Informed consent was obtained from all individual participants included in the study.

Funding

The study was supported by the Swedish Research Council for Health, Working Life and Welfare (Forte), grant 2021-01546.

Data Availability

The data generated during and/or analyzed during this study are not publicly available for privacy or ethical reasons, but requests will be considered by the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: All authors
Methodology: SLF, AE, LE, PF, MF
Investigation: SLF, LE, PF, MF
Formal analysis: SLF, KW, AE, MF
Writing – original draft: SLF, KW, AE, MF,
Writing – review & editing: All authors
Supervision: AE

Conflicts of Interest

None declared.

Multimedia Appendix 1
Case description used in the interviews to guide discussions on return to work, translated from Swedish.
[DOCX File, 39 KB - cancer_v12i1e89954_app1.docx]

Checklist 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.

XSL•FO
RenderX

[PDF File, 437 KB - cancer_v12i1e89954_app2.pdf]

References

1. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003 Oct 11;362(9391):1225-1230. [doi: 10.1016/S0140-6736(03)14546-1] [Medline: 14568747]
2. Wang Y, Wong ELY, Nilsen P, Chung VCH, Tian Y, Yeoh EK. A scoping review of implementation science theories, models, and frameworks—an appraisal of purpose, characteristics, usability, applicability, and testability. Implement Sci 2023 Sep 19;18(1):43. [doi: 10.1186/s13012-023-01296-x] [Medline: 37726779]
3. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
4. Zhao J, Xu L, Sun J, et al. Global trends in incidence, death, burden and risk factors of early-onset cancer from 1990 to 2019. BMJ Oncol 2023;2(1):e000049. [doi: 10.1136/bmjonc-2023-000049] [Medline: 39886513]
5. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018 Mar 17;391(10125):1023-1075. [doi: 10.1016/S0140-6736(17)33326-3] [Medline: 29395269]
6. Statistics on cancer incidence. National Board of Health and Welfare. 2024. URL: https://www.socialstyrelsen.se/contentassets/8fb02af2b1cd48edb6468c80835015c4/2025-11-9875.pdf [accessed 2025-12-16]
7. Nyberg ST, Airaksinen J, Pentti J, et al. Predicting work disability among people with chronic conditions: a prospective cohort study. Sci Rep 2023 Apr 18;13(1):6334. [doi: 10.1038/s41598-023-33120-3] [Medline: 37072462]
8. Pergolotti M, Wood KC, Kendig T, Love K, Mayo S. Cancer rehabilitation services for older women with breast cancer: Impact on health-related quality of life outcomes. J Am Geriatr Soc 2024 Aug;72(8):2402-2411. [doi: 10.1111/jgs.19021] [Medline: 38884258]
9. de Boer AGEM, de Wind A, Coenen P, et al. Cancer survivors and adverse work outcomes: associated factors and supportive interventions. Br Med Bull 2023 Apr 5;145(1):60-71. [doi: 10.1093/bmb/ldac028] [Medline: 36372773]
10. de Boer AG, Tamminga SJ, Boschman JS, Hoving JL. Non-medical interventions to enhance return to work for people with cancer. Cochrane Database Syst Rev 2024 Mar 5;3(3):CD007569. [doi: 10.1002/14651858.CD007569.pub4] [Medline: 38441440]
11. Kluit L, Hoving JL, Jamaludin FS, et al. Effectiveness of clinical healthcare interventions for enhancing the work participation of patients with various health conditions: a synthesis of systematic reviews. BMJ Open 2025 Feb 20;15(2):e094201. [doi: 10.1136/bmjopen-2024-094201] [Medline: 39979058]
12. Lundqvist A, Andersson E, Steen Carlsson K. Kostnader för cancer idag och år 2040 [Costs of cancer in Sweden today and in 2040]. : IHE—The Swedish Institute for Health Economics; 2016 URL: https://ihe.se/rapport/kostnad-for-cancer/ [accessed 2026-04-07]
13. Wu W, Yackel HD, Salner A, et al. Work-health balance of cancer survivors returning to work: a meta-ethnography. Eur J Oncol Nurs 2024 Feb;68:102482. [doi: 10.1016/j.ejon.2023.102482] [Medline: 38070450]
14. Seifart U, Schmielau J. Return to work of cancer survivors. Oncol Res Treat 2017;40(12):760-763. [doi: 10.1159/000485079] [Medline: 29183032]
15. Hart NH, Stout NL, Haywood D, et al. World Health Organization package of interventions for rehabilitation for cancer: a MASCC-endorsed resource for global action to address unmet rehabilitation needs of people affected by cancer. Support Care Cancer 2024 Jun 7;32(7):417. [doi: 10.1007/s00520-024-08569-1] [Medline: 38847912]
16. de Boer AG, Torp S, Popa A, et al. Long-term work retention after treatment for cancer: a systematic review and meta-analysis. J Cancer Surviv 2020 Apr;14(2):135-150. [doi: 10.1007/s11764-020-00862-2] [Medline: 32162193]
17. Chen L, Alexanderson KAE. Trajectories of sickness absence and disability pension before and after colorectal cancer: a Swedish longitudinal population-based matched cohort study. PLoS One 2021;16(1):e0245246. [doi: 10.1371/journal.pone.0245246]
18. Plym A, Johansson ALV, Bower H, et al. Causes of sick leave, disability pension, and death following a breast cancer diagnosis in women of working age. Breast 2019 Jun;45:48-55. [doi: 10.1016/j.breast.2019.02.012] [Medline: 30852409]
19. Hiyoshi A, Alexanderson K, Tinghög P, Cao Y, Fall K, Montgomery S. Future sick leave, disability pension, and unemployment among patients with cancer after returning to work: Swedish register-based matched prospective cohort study. Cancer 2025 Jan 1;131(1):e35580. [doi: 10.1002/cncr.35580] [Medline: 39377486]
20. Feuerstein M, Todd BL, Moskowitz MC, et al. Work in cancer survivors: a model for practice and research. J Cancer Surviv 2010 Dec;4(4):415-437. [doi: 10.1007/s11764-010-0154-6] [Medline: 20945110]
21. Eklund L, Engström M, de Boer AGEM, et al. Work-related experiences and needs while undergoing curative cancer therapy: exploring the target population's perspective during the developmental phase of a work-oriented intervention. Support Care Cancer 2025 Jun 17;33(7):589. [doi: 10.1007/s00520-025-09624-1] [Medline: 40523991]
22. Tamminga SJ, Jansen LP, Frings-Dresen MHW, de Boer AGEM. Long-term employment status and quality of life after cancer: a longitudinal prospective cohort study from diagnosis up to and including 5 years post diagnosis. Work 2020;66(4):901-907. [doi: 10.3233/WOR-203234] [Medline: 32925145]

XSL•FO
RenderX

23. Fitch MI, Sharp L, Hanly P, Longo CJ. Experiencing financial toxicity associated with cancer in publicly funded healthcare systems: a systematic review of qualitative studies. J Cancer Surviv 2022 Apr;16(2):314-328. [doi: 10.1007/s11764-021-01025-7] [Medline: 33723742]
24. Marmot M, Allen J, Bell R, Bloomer E, Goldblatt P, Consortium for the European Review of Social Determinants of Health and the Health Divide. WHO European review of social determinants of health and the health divide. Lancet 2012 Sep 15;380(9846):1011-1029. [doi: 10.1016/S0140-6736(12)61228-8] [Medline: 22964159]
25. Tingulstad A, Meneses-Echavez J, Evensen LH, Bjerk M, Berg RC. Effectiveness of work-related interventions for return to work in people on sick leave: a systematic review and meta-analysis of randomized controlled trials. Syst Rev 2022 Sep 5;11(1):192. [doi: 10.1186/s13643-022-02055-7] [Medline: 36064472]
26. Hunter EG, Gibson RW, Arbesman M, D'Amico M. Systematic review of occupational therapy and adult cancer rehabilitation: part 2. Impact of multidisciplinary rehabilitation and psychosocial, sexuality, and return-to-work interventions. Am J Occup Ther 2017;71(2):7102100040p1-7102100040p8. [doi: 10.5014/ajot.2017.023572] [Medline: 28218586]
27. Schellack S, Breidenbach C, Rick O, Kowalski C. Predictive factors for return to work among cancer survivors in Europe: a systematic review. Crit Rev Oncol Hematol 2024 Aug;200:104422. [doi: 10.1016/j.critrevonc.2024.104422] [Medline: 38897312]
28. Mewes JC, Steuten LMG, Groeneveld IF, et al. Return-to-work intervention for cancer survivors: budget impact and allocation of costs and returns in the Netherlands and six major EU-countries. BMC Cancer 2015 Nov 12;15(1):899. [doi: 10.1186/s12885-015-1912-7] [Medline: 26560707]
29. Lim CYS, Laidsaar-Powell RC, Young JM, et al. Work: saviour or struggle? A qualitative study examining employment and finances in colorectal cancer survivors living with advanced cancer. Support Care Cancer 2022 Nov;30(11):9057-9069. [doi: 10.1007/s00520-022-07307-9] [Medline: 35972645]
30. Liu WH, Fox J, Yates P. Work-related experiences of prostate cancer survivors in Australia: a qualitative study. BMC Public Health 2023 Sep 16;23(1):1806. [doi: 10.1186/s12889-023-16706-4] [Medline: 37716940]
31. Caron M, Durand MJ, Tremblay D. Perceptions of breast cancer survivors on the supporting practices of their supervisors in the return-to-work process: a qualitative descriptive study. J Occup Rehabil 2018 Mar;28(1):89-96. [doi: 10.1007/s10926-017-9698-x] [Medline: 28271399]
32. Koordineringsinsatser inom hälso- och sjukvården. En nationell kartläggning 2021 och jämförelse med 2018 [Coordination initiatives within healthcare: a national survey in 2021 and comparison with 2018]. The Swedish Association of Local Authorities and Regions. 2022. URL: https://extra.skr.se/download/18.1d1b45d17f44a93c2f60b7e/1647012929421/SKR-Koordineringsinsatser-inom-hälso-%20och-sjukvården.pdf [accessed 2022-03-30]
33. Berglund E, Friberg E, Engblom M, Svärd V. Physicians' experience of and collaboration with return-to-work coordinators in healthcare: a cross-sectional study in Sweden. Disabil Rehabil 2024 Sep;46(18):4120-4128. [doi: 10.1080/09638288.2023.2261851] [Medline: 37772755]
34. Etuknwa A, Daniels K, Eib C. Sustainable return to work: a systematic review focusing on personal and social factors. J Occup Rehabil 2019 Dec;29(4):679-700. [doi: 10.1007/s10926-019-09832-7] [Medline: 30767151]
35. López-Faneca L, Ruiz-Frutos C, Gómez-Salgado J, et al. Prognostic factors affecting return to work in cancer patients: a systematic review. Occup Environ Med 2025 Aug 25;82(6):305-312. [doi: 10.1136/oemed-2025-110212] [Medline: 40803828]
36. Lau RS, Boesen ME, Richer L, Hill MD. Siloed mentality, health system suboptimization and the healthcare symphony: a Canadian perspective. Health Res Policy Syst 2024 Jul 17;22(1):87. [doi: 10.1186/s12961-024-01168-w] [Medline: 39020412]
37. Bilodeau K, Gouin MM, Fadhlaoui A, Porro B. Supporting the return to work of breast cancer survivors: perspectives from Canadian employer representatives. J Cancer Surviv 2024 Aug;18(4):1384-1392. [doi: 10.1007/s11764-023-01382-5] [Medline: 37140676]
38. Svärd V, Arapovic Johansson Z, Holmlund L, Hellman T, Kwak L, Björk Brämberg E. Collaboration in the return-to-work process after sick leave due to common mental disorders: a qualitative study of stakeholders' views on goals and roles. BMC Public Health 2024 Jun 11;24(1):1567. [doi: 10.1186/s12889-024-19063-y] [Medline: 38862931]
39. Désiron H, Simons B, Spooren A, et al. Practice-based evidence to support return to work in cancer patients. Front Rehabil Sci 2022;3:819369. [doi: 10.3389/fresc.2022.819369] [Medline: 36189048]
40. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. Int J Nurs Stud 2024 Jun;154:104705. [doi: 10.1016/j.ijnurstu.2024.104705] [Medline: 38564982]
41. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet 2001 Aug 11;358(9280):483-488. [doi: 10.1016/S0140-6736(01)05627-6] [Medline: 11513933]
42. Patton MQ. Qualitative Research & Evaluation Methods: Integrating Theory and Practice, 4th edition: SAGE Publications, Inc; 2015.
43. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

XSL•FO
RenderX

44. Nationellt vårdprogram för cancerrehabilitering [National clinical cancer care guidelines]. Confederation of Regional Cancer Centres. 2025. URL: https://cancercentrum.se/rehabiliteringpalliativvard/cancerrehabilitering/vardprogram.7797.html [accessed 2026-04-20]
45. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
46. Lincoln YS, Guba EG, Pilotta JJ. Naturalistic inquiry. Int J Intercult Relat 1985 Jan;9(4):438-439. [doi: 10.1016/0147-1767(85)90062-8]
47. Fitch MI, Nicoll I. Returning to work after cancer: survivors', caregivers', and employers' perspectives. Psychooncology 2019 Apr;28(4):792-798. [doi: 10.1002/pon.5021] [Medline: 30720242]
48. Smith GL, Lopez-Olivo MA, Advani PG, et al. Financial burdens of cancer treatment: a systematic review of risk factors and outcomes. J Natl Compr Canc Netw 2019 Oct 1;17(10):1184-1192. [doi: 10.6004/jnccn.2019.7305] [Medline: 31590147]
49. McPeake J, Mikkelsen ME, Quasim T, et al. Return to employment after critical illness and its association with psychosocial outcomes. A systematic review and meta-analysis. Ann Am Thorac Soc 2019 Oct;16(10):1304-1311. [doi: 10.1513/AnnalsATS.201903-248OC] [Medline: 31184500]
50. Khosa SK, Jelley W. Stakeholder involvement and systematic review of elements that determine a successful return to work for injured workers: a scoping review. J Allied Health Sci 2024;22(3):13 [FREE Full text]
51. Popa AE. Moving beyond the patients' views on the process of return to work after cancer: a qualitative evidence synthesis on articles published between 2008 and 2017. Work 2022;72(4):1299-1310. [doi: 10.3233/WOR-210554] [Medline: 35723146]
52. Gershfeld-Litvin A, Vishnia O, Hanalis-Miller T. Recovering or working: women's experiences of working while coping with cancer: a qualitative study. Support Care Cancer 2025 Mar 17;33(4):289. [doi: 10.1007/s00520-025-09349-1] [Medline: 40095125]
53. Galica J, Alsius A, Walker L, et al. Returning to work after cancer treatment: an exploratory sequential mixed-methods study guided by transitions theory. Cancer Nurs 2025 Jan 3;2025. [doi: 10.1097/NCC.0000000000001449] [Medline: 39778099]
54. Lamore K, Dubois T, Rothe U, et al. Return to work interventions for cancer survivors: a systematic review and a methodological critique. Int J Environ Res Public Health 2019 Apr 14;16(8):1343. [doi: 10.3390/ijerph16081343] [Medline: 31014004]

Abbreviation

AI: Artificial Intelligence
COREQ: Consolidated Criteria for Reporting Qualitative Research
RTW: return to work

Edited by M Balcarras; submitted 22.Dec.2025; peer-reviewed by B Schwarz, V White; revised version received 31.Mar.2026; accepted 01.Apr.2026; published 11.May.2026.

Please cite as:
Fällman SL, Wijk K, Efverman A, Eklund L, Fessé P, Trygged S, Fjell M
Stakeholder Perspectives on Cancer Survivors' Return to Work and Well-Being: Qualitative Interview Study
JMIR Cancer 2026;12:e89954
URL: https://cancer.jmir.org/2026/1/e89954
doi:10.2196/89954

© Sara L Fällman, Katarina Wijk, Anna Efverman, Linda Eklund, Per Fessé, Sven Trygged, Maria Fjell. Originally published in JMIR Cancer (https://cancer.jmir.org), 11.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Correlates of Fitness Tracker Ownership and Use in Cancer Survivors: Cross-Sectional Survey

Roberto M Benzo[1,2], PhD; James L Fisher[3], PhD; Macy K Tetrick[1], BS; Rujul Singh[1,4], BS; Alex B Osei[5], BS; Jessica Krok-Schoen[2,6], PhD; Brett S Nickerson[2,6], PhD; Electra D Paskett[1,2], PhD; Ali Kargarandehkordi[7,8], PhD; Sachin Kumar[9], PhD; Sara M St George[10,11], PhD; Michael F Rayo[12], PhD; Frank J Penedo[11,13], PhD; Peter Washington[8], PhD

[1]Division of Cancer Prevention and Control, Department of Internal Medicine, College of Medicine, The Ohio State University, Columbus, OH, United States

[2]The Ohio State University Comprehensive Cancer Center, The Ohio State University Wexner Medical Center, Columbus, OH, United States

[3]Arthur G. James Cancer Hospital, Columbus, OH, United States

[4]Grossman School of Medicine, New York University, New York, NY, United States

[5]Department of Design, The Ohio State University, Columbus, OH, United States

[6]School of Health and Rehabilitation Sciences, The Ohio State University, Columbus, OH, United States

[7]Department of Information and Computer Sciences, University of Hawai i at Mānoa, Honolulu, HI, United States

[8]Division of Clinical Informatics and Digital Transformation (DoC-IT), Department of Medicine, University of California, San Francisco, CA, United States

[9]Department of Integrated Systems Engineering, College of Engineering, The Ohio State University, Columbus, OH, United States

[10]Department of Public Health Sciences, Miller School of Medicine, University of Miami, Miami, FL, United States

[11]Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL, United States

[12]Department of Computer Sciences and Engineering, College of Engineering, The Ohio State University, Columbus, OH, United States

[13]Department of Psychology, University of Miami, Miami, FL, United States

Corresponding Author:
Roberto M Benzo, PhD
Division of Cancer Prevention and Control, Department of Internal Medicine
College of Medicine
The Ohio State University
Suite 200
3650 Olentangy River Rd
Columbus, OH, 43124
United States
Phone: 1 4127281156
Email: roberto.benzo@osumc.edu

Abstract

Background: Consumer fitness tracker devices offer scalable opportunities to monitor real-world behavior and support health in cancer survivorship. However, adoption and sustained use outside structured research settings remain incompletely characterized, limiting their integration into survivorship care.

Objective: The primary objectives were to describe the prevalence of fitness tracker ownership and use patterns among cancer survivors and to identify sociodemographic, psychosocial, and usability-related correlates of device ownership and frequent use.

Methods: We conducted a cross-sectional survey of 893 cancer survivors enrolled in the Total Cancer Care protocol at a comprehensive cancer center. Participants completed an adapted online questionnaire assessing fitness tracker ownership, frequency of use, and perceived barriers and facilitators. Multivariable logistic regression models were used to identify sociodemographic, psychosocial, and usability-related correlates of device ownership and frequent use.

Results: More than half of the participants (506/893, 56.7%) reported owning a fitness tracker, and among owners, 82.2% (416/506) reported frequent use (most days or every day), including 71.3% (361/506) who wore the device every day. The most commonly used devices were the Apple iWatch (272/506, 53.8%) and Fitbit (147/506, 29.1%). In multivariable analyses, fitness tracker ownership was independently associated with sex, household income, and cancer site. Male participants had lower odds

XSL•FO
RenderX

of ownership (adjusted odds ratio [aOR] 0.61, 95% CI 0.42-0.87; *P*=.006), while higher household income was associated with greater ownership (US $50,000-$99,999: aOR 1.70, 95% CI 1.13-2.57; *P*=.01; ≥US $100,000: aOR 3.64, 95% CI 2.41-5.50; *P*<.001). Ownership also differed by cancer site. Among owners, frequent use was concurrently and inversely associated with self-reported device discomfort (*P*<.001), low motivation (*P*<.001), information overload (*P*=.04), and limited app integration (*P*=.007).

Conclusions: In this single-center, predominantly White, higher-income sample of cancer survivors, fitness tracker ownership was common and patterned by demographic and socioeconomic characteristics, while sustained engagement among device owners was associated with psychosocial and usability factors. These findings suggest that scalable, fitness tracker–enabled survivorship care will need to address both structural disparities in ownership as well as behavioral readiness and user experience to ensure clinically meaningful implementation.

(JMIR Cancer 2026;12:e92876) doi:10.2196/92876

KEYWORDS

digital health; cancer survivorship; wearable devices; physical activity; behavior change; usability; mobile health; mHealth; fitness tracker; wearable activity monitors; health equity; US adult population; sociodemographic determinants of physical activity; wearables; United States; observational studies; sociodemographic determinants; physical activity data

Introduction

Fitness tracker sensing technologies have become increasingly prominent in oncology research as tools for continuously capturing real-world physiological and behavioral data, including physical activity (PA), heart rate, heart rate variability, sleep, and circadian rest-activity rhythms. These devices offer a potentially scalable approach for monitoring patient recovery, functional status, and emerging health risks in daily life, complementing infrequent clinic-based assessments [1-3]. Their expanding role reflects a broader shift toward digital health measurement in cancer survivorship and the need for objective data streams that can inform timely intervention and long-term self-management.

PA is among the most frequently assessed metrics in fitness tracker–enabled oncology research because of its established relevance to cancer survivorship outcomes. Moreover, exercise is safe before, during, and after treatment and is associated with improved quality of life, reduced fatigue, and better physical function [3,4]. Observational studies consistently link higher postdiagnosis PA with lower risks of cancer-specific and all-cause mortality across multiple cancer types [3,5]. However, many survivors experience declines in PA during treatment that persist into survivorship, and exercise promotion is not routinely embedded in oncology care [6]. Reflecting this gap, Stout et al [7] highlighted the need for patient-centered pathways that support repeated assessment of functional status and timely referral for personalized exercise interventions. In this context, fitness tracker devices may complement clinic-based performance tests by enabling continuous monitoring of functional proxies, such as habitual and ecologically valid measures of PA walking patterns and heart-rate responses to daily ambulation, as well as emerging estimates of cardiorespiratory fitness (eg, maximum volume of oxygen consumed [VO_2 max] derived from walking-based algorithms), thereby supporting scalable identification of functional decline between visits.

Recent reviews indicate substantial growth in the use of fitness trackers across oncology studies. In a synthesis of 199 trials involving more than 18,500 patients, Chow et al [8] reported that the most commonly used devices were ActiGraph (36% of studies), Fitbit (19% of studies), Garmin (7% of studies), and activPAL (6% of studies) [8]. Activity minutes and step counts were the most frequently collected metrics, with adherence rates reported at or above 80%, when available. Complementary work has emphasized the role of fitness trackers in behavioral interventions: a meta-analysis of randomized controlled trials found that programs incorporating pedometers or activity trackers yield significant increases in moderate-to-vigorous PA, total activity, and daily steps, as well as improvements in fatigue, aerobic fitness, and quality of life relative to usual care [9]. A scoping review concluded that wearable fitness trackers are generally acceptable and can facilitate short-term increases in activity in cancer survivors, although long-term adherence and maintenance remain inconsistent [10].

Tracker-derived signals may also provide insight into broader health outcomes beyond PA. For example, rest-activity rhythm parameters obtained from accelerometry have demonstrated moderate-to-strong associations with survival, functional status, and symptom burden [11]. However, the overall literature remains limited by methodological heterogeneity, reliance on small or trial-selected samples, variability in device algorithms, and minimal integration into routine survivorship care [8,10,11]. These limitations underscore the need to understand the real-world feasibility of relying on fitness tracker devices to inform clinical or behavioral decision-making.

A critical barrier to translation is that little is known about real-world patterns of fitness tracker ownership and use among cancer survivors. Most existing oncology studies recruit participants who agree to use devices provided by the research team, leaving unanswered questions on whether survivors in the general population own fitness trackers, use them consistently, or differ from the broader population in meaningful ways. Moreover, digital health interventions risk reinforcing inequities if fitness tracker access and engagement vary by age, race, ethnicity, socioeconomic status, or cancer history. Characterizing these patterns is essential for determining the feasibility, reach, and equity of tracker-enabled survivorship care.

XSL•FO
RenderX

To address these gaps, this study examined fitness tracker ownership and active use among a large sample of cancer survivors. We identified sociodemographic, clinical, and behavioral correlates of ownership and sustained use, providing data needed to inform the design and implementation of equitable, scalable digital health strategies for cancer survivorship.

Methods

Study Design and Setting

This cross-sectional study examined fitness tracker ownership and use among cancer survivors enrolled in the Total Cancer Care (TCC) protocol at the Ohio State University Comprehensive Cancer Center – James Cancer Hospital and Solove Research Institute. The study followed the CONSORT (Consolidated Standards of Reporting Trials) guidelines. Data were collected between December 2023 and December 2024 using a self-administered online survey delivered through REDCap (Research Electronic Data Capture). A trained research assistant was available over email to address participants' questions.

Ethical Considerations

The study was approved by the Ohio State University Institutional Review Board (IRB #2023C0190) with a waiver of signed consent, and standardized protection of participants' rights and interests was ensured. Online consent was obtained through survey completion. The right to discontinue involvement in the survey at any point was guaranteed to participants. Participants were offered entry into a raffle for US $25 Amazon electronic gift cards as compensation for their participation. Data security and participant confidentiality were protected via several measures: data were fully deidentified, secure and password-protected data storage systems were installed, access to identifying information was implemented, and confidentiality protocols were mandated.

Participants and Recruitment

Eligible participants were adults (≥18 years), English speaking, diagnosed with a neoplasm (breast, prostate, lung, colorectal, head and neck, gynecologic, genitourinary, endocrine, hematologic, thoracic, digestive, nervous system, cutaneous, or soft-tissue/bone sites), and with an active email address in the TCC registry. Individuals with cognitive or psychiatric impairments precluding informed participation were excluded.

Recruitment was conducted by querying the TCC database for all eligible patients who had not been recently contacted for other survey requests (ie, ≥3-4 months since the last contact). The sampling frame included 8568 patients with cancer at different sites: breast (n=2137, 24.9%), thyroid (n=1451, 16.9%), head and neck (n=1670, 19.5%), digestive system (n=661, 7.7%), urinary system (n=727, 8.5%), thorax (n=409, 4.8%), hematologic system (n=533, 6.2%), gynecologic system (n=190, 2.2%), and additional sites (cutaneous, nervous system, soft-tissue/bone, and unknown; n=790, 9.2%). Of these, 7700 (89.9%) individuals had an email address registered with the TCC and were contacted via email invitations sent in monthly batches. Participants received an email invitation with a short study description and a link to a landing webpage with a full description.

Of the 7700 (89.9%) individuals contacted, 895 (11.6%) completed the survey, 152 (2.0%) partially completed the survey, and 6653 (86.4%) did not respond. Two completed surveys were excluded due to missing key demographic variables required for analysis, yielding a final analytical sample of 893 (11.6%) participants. The overall survey response rate (completed or partially completed surveys) was 13.6% (n=1047).

Survey Instrument

The survey was part of the Lifestyle Technology in Cancer Study, a broader multimodule questionnaire designed to assess how cancer survivors engage with mobile health (mHealth) technologies across several domains.

For the fitness tracker component, an instrument developed by Holko et al [12] for Federally Qualified Health Center patients was used and adapted with permission. The adaptation retained the original structure (awareness, ownership, reasons for use and nonuse, perceived barriers, and data-sharing preferences), while updating language to reflect survivorship contexts (eg, anchoring items to recovery and health after cancer treatment) and modern device features, such as sleep tracking, workout guidance, and integration with other apps and devices. Additional items were added to capture survivor-specific considerations, including participants' interests in topics, format, and features for future digital lifestyle interventions. Prior engagement with supportive care services, including nutrition and PA counseling, was assessed in alignment with American Cancer Society survivorship recommendations [3].

The broader Lifestyle Technology in Cancer survey also included modules on mHealth app use, health and lifestyle behaviors, and social and structural determinants of health. These modules were collected to support a series of future analyses and manuscripts focused on different aspects of mHealth engagement; however, they fall outside the scope of the study. Here, the fitness-tracker module and its associated outcomes (ownership and use) are reported, consistent with the analytic aims of this study. All items were administered electronically in English via REDCap, with multiple-choice responses and optional free-text fields. The adapted survey was pilot-tested internally for clarity prior to study launch and can be found in Multimedia Appendix 1.

Measures

The survey assessed multiple domains related to fitness tracker ownership, use, and perceptions. Participants were first asked whether they currently owned a fitness tracker and, if so, to indicate the device brand, duration of ownership, and frequency of wear. Follow-up questions explored reasons for using the tracker, including activity monitoring, sleep and heart rate tracking, goal setting, motivation, and accountability, as well as reasons for nonuse, such as discomfort, technical issues, or loss of interest. Non-owners were asked to identify reasons for not owning a tracker (eg, cost, lack of interest, privacy concerns, or feeling self-motivated without a device) and to indicate whether they were interested in obtaining one in the future. Barriers to and facilitators of use were assessed through

XSL•FO
RenderX

multiple-choice items adapted from Holko et al [12], with optional free-text responses, allowing participants to elaborate on personal challenges and motivations related to wearable engagement.

Sociodemographic and social determinants of health variables were assessed using items adapted from the PRAPARE (Protocol for Responding to and Assessing Patient Assets, Risks, and Experiences) survey, a widely used, nationally standardized screening tool for social and structural determinants of health in clinical settings [13]. Psychosocial measures included perceived stress and confidence in one's ability to take good care of one's health. The confidence item was adapted from the Health Information National Trends Survey (HINTS-5), a nationally administered and validated population-based survey [2].

Statistical Analyses

Descriptive statistics, including means (SEs) and percentages, were used to characterize participants and their reported reasons for owning and using a fitness tracker. To explore associations between various factors and fitness tracker ownership, univariate logistic regression was performed to estimate odds ratios (ORs) and 95% CIs for demographic, socioeconomic, psychosocial, and PA-related variables. Similarly, univariate logistic regression was used to assess associations between demographic characteristics and specific challenges that may hinder fitness tracker usage (eg, concerns about accuracy, discomfort, or lack of motivation). These associations were evaluated in relation to frequent use (ie, using a fitness tracker on most days or every day) versus less frequent use (including the remaining categories). This binary threshold was predetermined to differentiate habitual use from sporadic use, as both "most days" and "every day" signify a consistent integration of the device into daily life. Furthermore, this threshold was selected to align methodologically with the common operationalization of adherence in digital health and the wearable literature [14]. To identify independent predictors of fitness tracker ownership and frequent use, multivariable logistic regression models were constructed. Variables with at least marginal statistical significance ($P<.15$) in univariate analyses were initially included. A stepwise backward elimination approach was then applied: variables were removed sequentially, starting with the least significant, until all remaining variables were either (1) statistically significantly associated with the outcome or (2) demonstrated substantive confounding effects (defined as altering the association of another variable in the model with the outcome by ≥15%). The final multivariable models thus included only those factors that were independently associated with fitness tracker ownership or frequent use. Breast cancer was selected as the reference category for the cancer-site variable because it represented the largest single cancer site in the analytic sample (n=278, 31.1%) and provided the most stable and interpretable comparison group; this choice is descriptive only and does not imply a substantive hypothesis about breast cancer survivors as a normative referent. This variable selection strategy was chosen to accommodate the exploratory aims of the study, which sought to identify correlates of fitness tracker ownership and use across a broad set of candidate domains, for which no single established theoretical framework jointly specifies predictors in cancer survivor populations. A purely confirmatory, theory-driven specification was therefore not feasible, and a data-driven screening approach was selected to allow empirical identification of independently associated factors, while controlling for potential confounding. Multicollinearity among candidate predictors was evaluated using variance inflation factors (VIFs) prior to model fitting; all VIFs were below conventional thresholds (<5), indicating no problematic collinearity. To assess model stability, both multivariable models were refit with all candidate predictors entered simultaneously (Tables S1 and S2 in Multimedia Appendix 2) to confirm that the direction and magnitude of the primary associations were consistent across specifications. Given the cross-sectional design, all models estimated associations and could not establish temporal order or causality. Moreover, the two regression models were crafted with distinct analytical objectives. The ownership model explored a wide array of sociodemographic, clinical, and behavioral factors to pinpoint correlates linked to device ownership. In contrast, the frequent-use model was more narrowly focused, concentrating on modifiable usability and psychological barriers reported by device owners.

Notably, PA was excluded as a potential predictor in the frequent-use model due to this model's emphasis on actionable usability and psychological barriers. Complete-case analysis was used for all regression models; participants with missing data on any variable included in a given model were excluded from that analysis. For household income, which had a substantial nonresponse rate (149/893, 16.7%), "no response" was retained as an explicit category in both univariable and multivariable ownership models rather than excluding these participants in order to preserve the sample size and avoid bias from listwise deletion of a nonrandomly missing variable. Missingness on other covariates was minimal and handled via complete-case exclusion. The multivariable model predicting fitness tracker ownership included 893 participants, and the multivariable model predicting frequent use among device owners included 504 participants.

Results

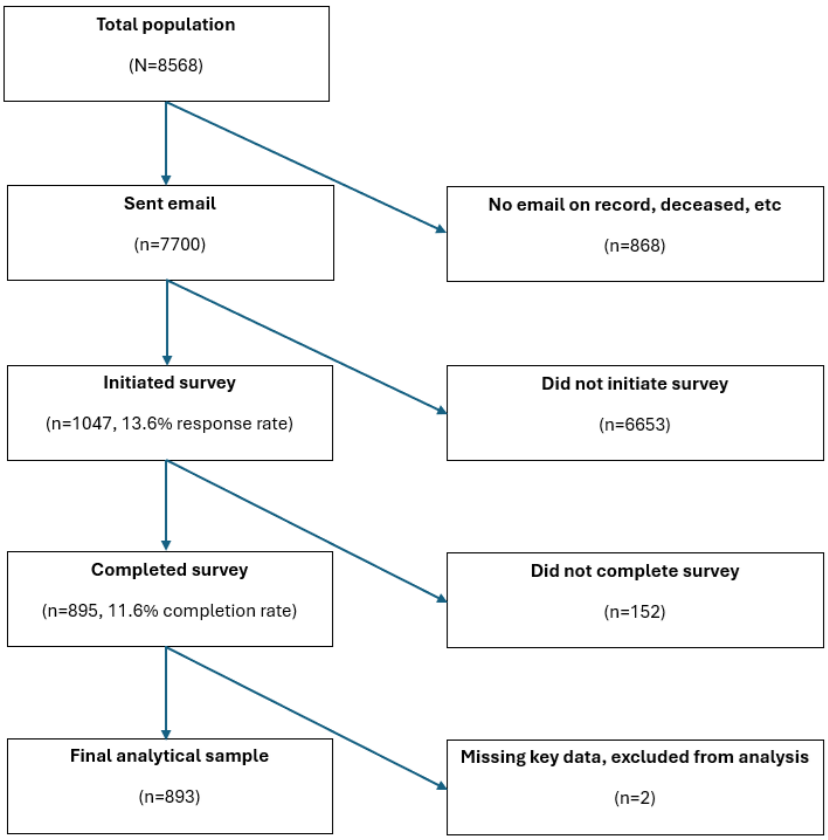
Participant Characteristics

Figure 1 shows the participant selection flowchart based on CONSORT guidelines. A total of 893 cancer survivors were included in the analytic sample (Table 1). The mean age was 61.8 (SE 0.38) years; 68.8% (614/893) identified as female and 93.5% (835/893) as non-Hispanic White. The most common cancer sites were breast (278/893, 31.1%), thyroid (166/893, 18.6%), and oral cavity/pharynx (149/893, 16.7%). The majority of survivors had the *International Classification of Diseases for Oncology* (ICD-O) cancer classification of "malignant" (742/893, 83.1%), and the mean time since diagnosis at the time of survey completion was 10.3 (SE 0.20) years. Approximately one-third of the participants reported annual household incomes exceeding US $100,000 (324/893, 36.3%), and 68.3% (610/893) earned at least a college degree. Regarding psychosocial factors, 20.8% (186/893) participants reported feeling "quite a bit" or "very much" stressed, and 52.3% (467/893) were "very" or

XSL•FO
RenderX

“completely” confident in their ability to take good care of their health.

Figure 1. Participant selection flowchart based on CONSORT guidelines. CONSORT: Consolidated Standards of Reporting Trials.



XSL•FO
RenderX

Table 1. Sample demographic and socioeconomic factors and responses to questions concerning stress and confidence in taking care of health for 893 mHealth[a] participants.

Characteristics	Value
Age[b] (years), mean (SE)	61.8 (0.38)
Age group (years), n (%)	
<50	138 (15.5)
50-59	208 (23.3)
60-69	316 (35.4)
≥70	231 (25.9)
Sex, n (%)	
Female	614 (68.8)
Male	279 (31.2)
Race/ethnicity, n (%)	
White/non-Hispanic	835 (93.5)
Other	58 (6.5)
Cancer site, n (%)	
Breast	278 (31.1)
Thyroid	166 (18.6)
Oral cavity/pharynx	149 (16.7)
Bone marrow	61 (6.8)
All other sites	239 (26.8)
Malignancy level, n (%)	
Malignant	742 (83.1)
In situ	125 (14.0)
Benign/unknown	26 (2.9)
Time since diagnosis (years), mean (SE)	10.3 (0.2)
Time since diagnosis (years), n (%)	
<5	147 (16.5)
5-10	389 (43.6)
>10	357 (40.0)
Household income (US $), n (%)	
0-49,999	171 (19.1)
50,000-99,999	249 (27.9)
≥100,000	324 (36.3)
No response	149 (16.7)
Education, n (%)	
High school diploma/General Educational Development [GED] or less	109 (12.2)
Some college	169 (18.9)
College degree	359 (40.2)
Graduate/professional degree	251 (28.1)
No response	5 (0.6)
Stress level , n (%)	
Not at all	136 (15.2)
A little bit	313 (35.1)

XSL•FO
RenderX

Characteristics	Value
Somewhat	254 (28.4)
Quite a bit/very much	186 (20.8)
No response	4 (0.4)
Confidence about ability to take good care of health, n (%)	
Completely confident	126 (14.1)
Very confident	341 (38.2)
Somewhat confident	316 (35.4)
A little confident/not confident at all	108 (12.1)
No response	2 (0.2)

[a]mHealth: mobile health.

[b]Age is presented as a continuous variable with mean (SE). Percentages may not total 100% due to rounding. Small cell sizes (eg, <5) are presented because the data were fully deidentified and not linked to sensitive or identifiable subgroups. Categories with small response counts (eg, <5) should be interpreted with caution.

Fitness Tracker Ownership

More than half of the respondents (506/893, 56.7%) reported currently owning a fitness tracker, while 43.3% (386/893) did not, and 1 (0.1%) individual declined to answer (Figure 2). The most common devices were the Apple iWatch (272/506, 53.8%), Fitbit (147/506, 29.1%), the Samsung Galaxy Watch (36/506, 7.1%), and the Garmin Wrist Tracker (29/506, 5.7%). Among non-owners, 60.6% (234/386) participants expressed interest in obtaining a tracker, with the most common reasons being to monitor activity (173/386, 44.8%), sleep (148/386, 38.3%), and heart rate (146/386, 37.8%), as well as for encouragement, reminders, and accountability (114/386, 29.5%).

The most frequently endorsed facilitators of ownership were activity tracking (403/506, 79.6%), heart rate tracking (228/506, 45.1%), sleep tracking (164/506, 32.4%), and integration with other apps and devices (148/506, 29.2%), followed by goal setting or progress tracking (112/506, 22.1%). Preventive health tracking (38/506, 7.5%) and food/calorie tracking (31/506, 6.1%) were less commonly reported. Barriers to ownership reported among non-owners included cost (165/386, 42.7%) and lack of interest (138/386, 35.8%), which were the most frequently cited reasons for not having a tracker. Other common barriers were believing they did not need a fitness tracker to stay healthy (78/386, 20.2%) and limited technology skills (73/386, 18.9%). Smaller proportions of participants reported comfort issues (34/386, 8.8%), privacy concerns (33/386, 8.5%), and lack of awareness about fitness trackers (31/386, 8.0%).

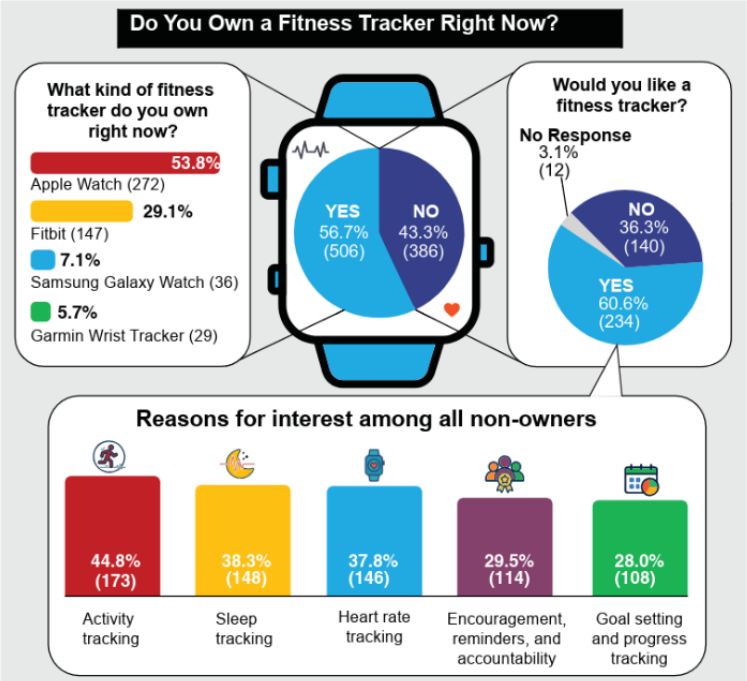
In univariate analyses (Table 2), several demographic and clinical factors were significantly associated with fitness tracker ownership. A younger age, female sex, a higher household income, higher educational attainment, and the cancer site were each associated with greater odds of ownership in unadjusted models. Specifically, male participants had lower odds of ownership compared to females (OR 0.42, 95% CI 0.31-0.56; *P*<.001), while higher household income was positively associated with ownership (US $50,000-$99,999: OR 1.61, 95% CI 1.08-2.38; *P*=.02; ≥US $100,000: OR 3.85, 95% CI 2.61-5.68; *P*<.001).

In the final adjusted multivariable model, fitness tracker ownership remained independently associated with sex, household income, and cancer site. Male participants had 39% lower odds of ownership compared to females (adjusted odds ratio [aOR] 0.61, 95% CI 0.42-0.87; *P*=.006). Higher household income was strongly associated with ownership, with participants reporting annual household incomes of US $50,000-$99,999 (aOR 1.70, 95% CI 1.13-2.57; *P*=.01) and ≥US $100,000 (aOR 3.64, 95% CI 2.41-5.50; *P*<.001) having greater odds of owning a tracker compared to those earning <US $50,000. Ownership also varied by cancer site, with lower odds observed among participants with thyroid cancer (aOR 0.64, 95% CI 0.41-1.00; *P*=.05), oral cavity/pharyngeal cancer (aOR 0.39, 95% CI 0.25-0.63; *P*<.001), bone marrow cancer (aOR 0.49, 95% CI 0.26-0.90; *P*=.02), and other cancer types (aOR 0.48, 95% CI 0.31-0.73; *P*<.001) relative to the reference group. Age was marginally associated with ownership in the adjusted model (aOR 0.99 per year, 95% CI 0.97-1.00; *P*=.05). Education, race/ethnicity, the BMI, perceived stress, and self-rated health confidence were not independently associated with ownership after adjustment.

The demographic homogeneity of the sample (primarily non-Hispanic White, college-educated individuals with higher household incomes, recruited via email) may limit statistical power to detect associations for certain characteristics, particularly race/ethnicity. Consequently, null findings for sociodemographic variables, such as race/ethnicity and education, should be interpreted in this context rather than as evidence that disparities do not exist. Finally, in a sensitivity analysis in which all candidate predictors were entered simultaneously, the direction, magnitude, and significance of associations were largely consistent with the reduced model (Table S1 in Multimedia Appendix 2).

XSL•FO
RenderX

Figure 2. Fitness tracker ownership, interest, and associations among cancer survivors. Participants were asked about their current ownership of fitness trackers, their interest in using one, and specific features or functionalities they find helpful. Over half reported owning a tracker, with the Apple iWatch being most common. Among those interested in owning one, top motivations included activity tracking and health monitoring. Facilitators of ownership included usability features such as integration with other apps, while common barriers were cost and lack of interest. Participants could select multiple responses, but those who chose “no response” for a section could not select other options. The "%" column refers to the percentage of participants who selected each option. Participants could select multiple responses.



Facilitators of Ownership

	%	n=506
Activity Tracking	79.6%	403
Heart Rate Tracking	45.1%	228
Sleep Tracking	32.4%	164
Integrating with Other Apps and Devices	29.2%	148
Encouragement, Reminders, and Accountability	26.3%	133
Goal Setting and Progress Tracking	22.1%	112
Health Insights	11.9%	60

Barriers of Ownership

	%	n=386
Cost	42.7%	165
Lack of interest	35.8%	138
I don't need a fitness tracker to live healthy	20.2%	78
Not tech savvy	18.9%	73
Comfort issues	8.8%	34
Privacy concerns	8.5%	33
Lack of awareness (I don't know what they are or what they do)	8.0%	31

XSL•FO
RenderX

Table 2. Logistic regression predicting fitness tracker ownership.

Variables	Univariable		Multivariable (n=893)[a]	
	OR[b] (95% CI)	*P* value	aOR[c] (95% CI)	*P* value
Age (per year increase)	0.97 (0.96-0.98)	<.001	0.99 (0.97-1.00)	.052
Sex				
Female	Reference	—[d]	Reference	—
Male	0.42 (0.31-0.56)[e]	<.001[e]	0.61 (0.42-0.87)[e]	.006
Race/ethnicity				
White, non-Hispanic	Reference	—	Reference	—
Other	0.77 (0.46-1.31)	.34	—	—
Annual household income (US $)				
0-49,999	Reference	—	—	—
50,000-99,999	1.61 (1.08-2.38)[e]	.02[e]	1.70 (1.13-2.57)[e]	.01
≥100,000	3.85 (2.61-5.68)[e]	<.001[e]	3.64 (2.41-5.50)[e]	<.001[e]
No response	1.36 (0.88-2.12)	.17	1.47 (0.93-2.33)	.10
Education				
High school diploma/General Educational Development (GED) or less	Reference	—	Reference	—
Some college	1.05 (0.65-1.70)	.85	—	—
College degree	2.13 (1.38-3.29)[e]	<.001[e]	—	—
Graduate/professional degree	1.85 (1.17-2.91)[e]	.008[e]	—	—
Cancer site				
Breast	Reference	—	—	—
Thyroid	0.63 (0.42-0.95)[e]	.03[e]	0.64 (0.41-1.00)[e]	.05[e]
Oral cavity/pharynx	0.27 (0.18-0.41)[e]	<.001[e]	0.39 (0.25-0.63)[e]	<.001[e]
Bone marrow	0.32 (0.18-0.56)[e]	<.001[e]	0.49 (0.26-0.90)[e]	.02[e]
All other sites	0.33 (0.23-0.47)[e]	<.001[e]	0.48 (0.31-0.73)[e]	<.001[e]
Perceived stress level				
Not at all	Reference	—	Reference	—
A little bit	1.23 (0.83-1.85)	.306	—	—
Somewhat	1.48 (0.98-2.26)	.07	—	—
Quite a bit/very much	1.43 (0.91-2.23)	.12	—	—
Confidence in self-care				
Completely confident	Reference	—	—	—
Very confident	0.84 (0.56-1.28)	.43	—	—
Somewhat confident	0.79 (0.52-1.20)	.26	—	—
A little/not at all confident	0.80 (0.47-1.34)	.39	—	—
PA[f] frequency				
Every day	Reference	—	—	—
4-5 days/week	0.93 (0.64-1.36)	.72	—	—
2-3 days/week	1.04 (0.70-1.54)	.85	—	—
<2 days/week	0.72 (0.49-1.05)	.09	—	—

XSL•FO
RenderX

Variables	Univariable		Multivariable (n=893)[a]	
	OR[b] (95% CI)	*P* value	aOR[c] (95% CI)	*P* value
BMI category (kg/m²)				
<25.0 (normal weight)	Reference	—	—	—
25.0-29.9 (overweight)	0.89 (0.63-1.26)	.51	—	—
≥30.0 (obese)	0.82 (0.59-1.14)	.24	—	—

[a]The multivariable model included variables with *P*<.15 in univariable analyses; backward elimination retained variables that were statistically significant or demonstrated substantive confounding (≥15% change in another coefficient).

[b]OR: odds ratio.

[c]aOR: adjusted odds ratio.

[d]Variables not retained are indicated by an em dash (—). The univariable sample included all 893 participants; the multivariable model was fit on n=893. Variables included in the multivariable model were mutually adjusted.

[e]Statistical significance at *P*<.05 (two sided).

[f]PA: physical activity.

Use Patterns and Perceived Barriers Among Fitness Tracker Owners

Among participants who owned a fitness tracker, most reported daily use: 71.3% (361/506) wore their device every day, and an additional 10.9% (55/506) wore it most days, while 6.3% (32/506) wore it some days, 6.5% (33/506) rarely used it, 4.5% (23/506) never used it, and 0.4% (2/506) declined to answer the question (Figure 3). For analytic purposes, “never,” “rarely,” and “some days” were grouped as infrequent use (88/506, 17.3%), and “most days” and “every day” were categorized as frequent use (416/506, 82.2%). This binary variable served as the outcome for the logistic regression analyses presented in Table 3.

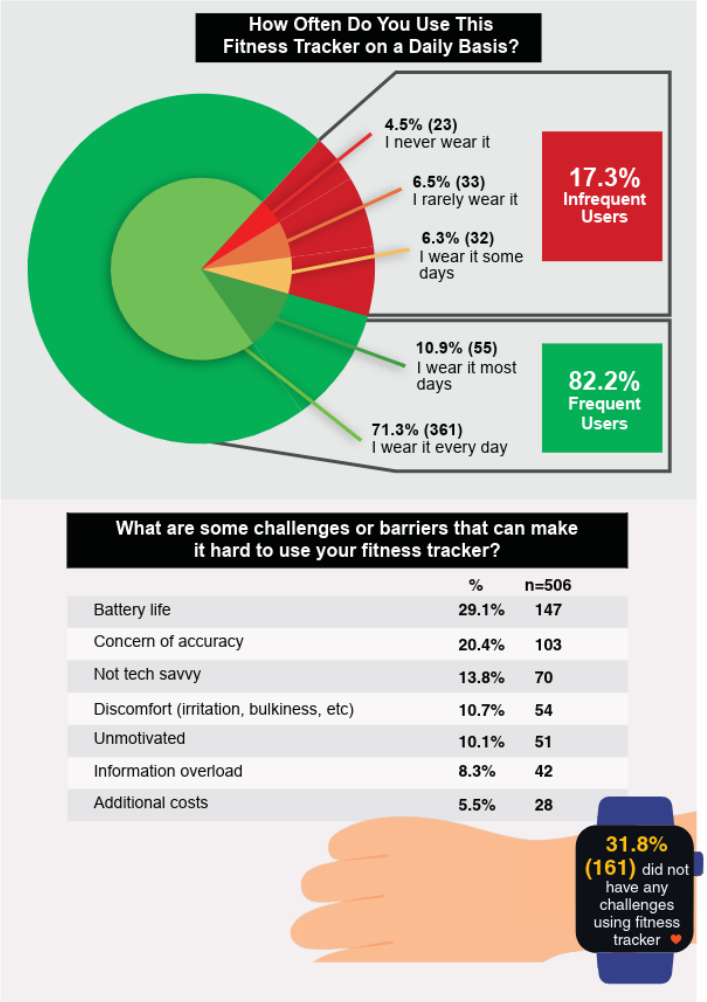
Among current users, the most commonly reported challenges to regular use were battery life (147/506, 29.1%) and concerns about accuracy (103/506, 20.4%). Other barriers included not feeling tech savvy (70/506, 13.8%); discomfort, such as irritation or bulkiness (54/506, 10.7%); low motivation (51/506, 10.1%); information overload (42/506, 8.3%); and additional costs (28/506, 5.5%). Notably, nearly one-third of the owners (161/506, 31.8%) reported having no challenges using their device.

In multivariable models (Table 3), several barriers were significantly associated with lower odds of being a frequent tracker user. In the multivariable model, participants who reported device discomfort or irritation were 75% less likely to use their trackers frequently (aOR 0.25, 95% CI 0.13-0.48; *P*<.001). Those citing information overload were 55% less likely to use their trackers frequently (aOR 0.45, 95% CI 0.21-0.97; *P*=.04). Similarly, participants reporting low motivation were 76% less likely to use their trackers frequently (aOR 0.24, 95% CI 0.12-0.46; *P*<.001). Reports of device malfunction and lack of integration with other apps or devices were each associated with 75% lower odds of frequent use (aOR 0.25, 95% CI 0.10-0.62; *P*=.003; aOR 0.25, 95% CI 0.09-0.69; *P*=.007, respectively). Accuracy concerns, privacy issues, and battery life were not significantly associated with the frequency of use. It is important to note, however, that some barrier items (particularly low motivation and information overload) are conceptually proximate to the outcome of infrequent use and may partially reflect participants’ descriptions of their current behavior rather than independent antecedent factors. Device-level barriers, such as discomfort, malfunction, and limited app integration, are more clearly separable from use frequency, but the cross-sectional design precludes establishing temporal order for any of these associations. Finally, in a sensitivity analysis in which all candidate predictors were entered simultaneously, the direction and magnitude of the primary associations were consistent with the reduced model (Table S2 in Multimedia Appendix 2). Two variables near the conventional significance threshold exhibited minor shifts: concern about accuracy moved from nonsignificant to significant, and information overload moved from significant to nonsignificant. These fluctuations are consistent with the modest size of the infrequent-use group (n=88, 17.3%) that do not alter the substantive conclusions. A sensitivity analysis restricting the outcome to daily use produced largely consistent findings (Table S3 in Multimedia Appendix 2).

Overall, although most participants used their fitness trackers regularly, less frequent use was concurrently associated with self-reported device usability and motivational factors rather than demographic characteristics or privacy concerns. As these factors were measured at the same time as the outcome, they should be interpreted as correlates rather than established antecedent barriers.

XSL•FO
RenderX

Figure 3. Frequency of fitness tracker use and barriers to use among cancer survivors. Participants were asked how often they used their fitness tracker and whether they experienced any challenges with use. Most participants reported wearing their tracker daily, while others indicated less frequent use. Of the five response categories, three reflected infrequent use, while two represented frequent use. Among the identified barriers, the most common were battery life, accuracy concerns, and discomfort. Participants could select more than one challenge. Barriers were only reported if participants either selected specific issues or indicated "I did not have any challenges." Those who chose not to answer or skipped the question were excluded from the barrier summary. The "%" column refers to the percentage of participants who selected each option. Participants could select multiple responses. Those who chose "no response" for a section were not allowed to select additional options. Note: 0.5% of responses are missing due to nonresponse.



XSL•FO
RenderX

Table 3. Logistic regression predicting frequent use[a] of fitness trackers among current owners.

Variables	Univariable		Multivariable (n=504)[b]	
	OR[c] (95% CI)	*P* value	aOR[d] (95% CI)	*P* value
Age (per year increase)	1.00 (0.98-1.02)	.96	—[e]	—
Sex				
Female	Reference	—	Reference	—
Male	0.96 (0.59-1.58)	.88	—	—
Reported barriers to use				
Accuracy concerns	1.77 (0.92-3.40)	.09	1.91 (0.96-3.78)	.07
Discomfort (irritation, bulkiness)	0.30 (0.16-0.56)[f]	<.001[f]	0.25 (0.13-0.48)[f]	<.001[f]
Not tech savvy	1.00 (0.53-2.00)	.94	—	—
Battery life	0.65 (0.40-1.06)	.09	—	—
Information overload	0.43 (0.22-0.87)[f]	.02[f]	0.45 (0.21-0.97)[f]	.04[f]
Low motivation	0.27 (0.15-0.51)[f]	<.001[f]	0.24 (0.12-0.46)[f]	<.001[f]
Privacy concerns	0.57 (0.18-1.84)	.35	—	—
Additional costs	0.62 (0.25-1.50)	.28	—	—
Limited app/device integration	0.24 (0.10-0.59)[f]	.002[f]	0.25 (0.09-0.69)[f]	.007[f]
Device malfunction	0.33 (0.14-0.78)[f]	.01[f]	0.25 (0.10-0.62)[f]	.003[f]
Limited internet access	0.56 (0.14-2.14)	.39	—	—
No one to help me use it	0.91 (0.26-3.28)	.89	—	—
Other	0.57 (0.18-1.84)	.35	—	—

[a]Frequent use was defined as wearing the device every day or most days (416/506, 82.2%); the reference category was infrequent use (rarely, some days, or never; 88/506, 17.3%).

[b]The multivariable model included variables with $P<.15$ in univariable analyses; backward elimination retained variables that were statistically significant or demonstrated substantive confounding (≥15% change in another coefficient).

[c]OR: odds ratio.

[d]aOR: adjusted odds ratio.

[e]Variables not retained are indicated by an em dash (—). All included variables were mutually adjusted.

[f]Statistical significance at $P<.05$ (two sided).

Discussion

Principal Findings

This study included one of the most comprehensive examinations to date of real-world fitness tracker ownership and engagement among cancer survivors. More than half of the respondents reported owning a tracker, yet frequent use (defined as wearing the device most days or every day) was less universal, highlighting important distinctions between device adoption and sustained engagement. Patterns of ownership were associated primarily with demographic and socioeconomic characteristics within this cohort, while continued use among device owners appeared more strongly related to psychosocial and usability-related factors. These prevalence estimates should be interpreted in light of the cohort's characteristics: participants were recruited from a single institution, required an active email address to be contacted, and were predominantly non-Hispanic White, with a college education and a higher household income. Concurrent self-reports of device discomfort, information overload, and technical issues were associated with less consistent engagement in cross-sectional analysis.

Although the sample was relatively homogeneous in terms of race and ethnicity, which may limit the identification of certain access and equity-related barriers, the observed differences in ownership by sex, income, and cancer site highlight significant disparities in access, even within this group. Although this study was underpowered and not ideally suited to detect such patterns, they were still discernible within our limited sample. These patterns are becoming increasingly relevant as oncology moves toward digitally supported survivorship care, where fitness trackers offer a potential data source for monitoring PA, symptoms, and functional recovery.

Comparison With Prior Research

In this study, approximately 57% of cancer survivors reported owning a fitness tracker, a prevalence higher than that reported in several prior population-based studies. This difference likely reflects, in part, the cohort's recruitment approach and

XSL•FO
RenderX

demographic composition rather than a uniform increase in adoption among cancer survivors generally. Despite this caveat, Zhou et al [15], using Health Information National Trends Survey – Surveillance, Epidemiology, and End Results (HINTS-SEER) data, documented increasing adoption of wearables among US cancer survivors over time, with ownership rising from approximately 21% in 2019 to 31% in 2022. Prior work has also documented lower wearable adoption among older cancer survivors. For example, Faro et al [16], in a study with a mean participant age of 74 years, found that only 31.8% of survivors owned a wearable device and that increasing age is inversely associated with ownership.

In addition to overall prevalence, we observed differences in fitness tracker ownership by sex, household income, and cancer site that are largely consistent with the established literature. Higher ownership among female participants corresponds with previous population-based research indicating a greater adoption of self-monitoring technologies by women. Similarly, higher ownership is also observed among individuals with higher incomes [17]. In this cohort, cancer-specific ownership patterns were also identified; however, such patterns have been less frequently explored in previous survivorship research, and the findings of this study should be viewed as preliminary observations to be tested in studies with detailed treatment and clinical data.

Population-based analyses of mHealth tracking have identified lower engagement among older cancer survivors. In a large HINTS-SEER analysis, Hong [18] reported that mHealth and behavior tracking is independently associated with younger age, although this outcome combines app- and wearable-based tracking rather than wearable ownership alone. In this study, age was significantly associated with fitness tracker ownership in unadjusted analyses and was marginally associated in adjusted models, suggesting that the age gradient in this cohort is partially but not fully explained by correlated sociodemographic factors. Differences in cohort composition, outcome definitions, and baseline levels of technology engagement may also contribute to variation across studies. Importantly, these findings should not be interpreted as evidence that age-related disparities have resolved, particularly in geriatric oncology populations, where continued evaluation remains warranted.

Prior oncology research has primarily evaluated the feasibility of fitness trackers under supervised conditions. In chemotherapy populations, Nilanon et al [19] demonstrated that continuous monitoring is feasible, with adherence above 68% and lower activity predicting unplanned clinical encounters. Ng et al [20] similarly reported very high adherence (94%-96%) and high perceived usefulness among inpatient and outpatient rehabilitation patients provided with trackers. These intervention-based findings confirm feasibility when devices are supported within structured settings, but reveal little about voluntary, sustained use. This study extended prior research by examining behavioral, motivational, and usability factors linked to real-world fitness tracker engagement.

Interpretation of Key Findings

The observed associations between fitness tracker ownership and demographic and socioeconomic characteristics suggest that access to these technologies remains patterned by structural factors within cancer survivor populations. In this study, ownership was associated with household income, sex, and cancer site, indicating that socioeconomic resources and survivorship context may play a central role in shaping initial adoption. The income gradient is consistent with broader patterns of unequal access to consumer digital technologies documented in population-based studies [17]. Consumer-grade fitness trackers and smartwatches typically range from US $100 to US $400, a cost that may present a meaningful barrier for lower-income individuals, particularly given the financial burden associated with cancer treatment. Notably, cost was the most frequently endorsed barrier among non-owners in this study (42.7%), further supporting the interpretation that financial accessibility contributes to the observed income gradient.

Lower odds of fitness tracker ownership among male participants is consistent with sex-based differences in health self-monitoring behaviors observed in general population studies [17]. Prior research has documented that women are more likely than men to engage in preventive health behavior, patterns attributed, in part, to gender socialization around health vigilance and self-care [21].

Differences in fitness tracker ownership by cancer site have not been widely examined in prior survivorship research, and the cross-sectional design and absence of clinical covariates (treatment status, disease stage, recurrence, symptom burden, and functional limitations) preclude evaluation of the mechanisms underlying these patterns. The interpretations offered below are therefore hypothesis generating rather than explanatory. Breast cancer survivors, who served as the reference group, demonstrated the highest ownership prevalence, which may reflect both the demographic composition of this group (predominantly female, higher income) and the well-established culture of survivorship engagement, peer support, and wellness-oriented self-management in breast cancer communities. The lower ownership observed among survivors of oral cavity/pharynx cancers (aOR 0.39), bone marrow malignancies (aOR 0.49), and other sites may reflect differences in the treatment burden, physical sequelae that limit device wear (eg, peripheral neuropathy, dermatologic toxicity), or lower exposure to fitness-oriented survivorship messaging.

In contrast to ownership, frequent use among device owners was more closely associated with experiential and motivational factors within this cohort. Device discomfort, technical malfunctions, limited app integration, and information overload emerged as the most strongly associated barriers with lower frequent use. However, these associations warrant different levels of interpretive confidence. Device-level factors, such as discomfort, malfunction, and limited integration, are conceptually distinct from use frequency and are therefore more defensible as independent correlates. In contrast, self-reported motivational and cognitive burden items (low motivation, information overload) overlap conceptually with the outcome itself: reporting that one is “unmotivated” to use a tracker is difficult to disentangle from the behavior of not using it. These items may therefore function, in part, as retrospective explanations for disengagement. This distinction does not invalidate the observed associations but does constrain their

XSL•FO
RenderX

interpretation; none should be treated as established causal barriers, given the cross-sectional design. With this caveat, our findings align with prior digital oncology research emphasizing simplicity, interoperability, and actionable feedback as key correlates of sustained engagement [10,11].

Fitness trackers play a key role in PA monitoring, a central behavioral target in survivorship. PA is safe and beneficial before, during, and after treatment, improving fatigue, functional capacity, and quality of life and reducing mortality risk [6,7,22]. In a systematic review, Singh et al [9] found moderate-to-large improvements in PA and related outcomes from wearable-supported interventions. Similarly, oncology rehabilitation frameworks emphasize early identification of functional decline and referral to exercise oncology services [7]. Together, these bodies of work highlight why understanding real-world fitness tracker engagement is clinically significant.

Implications for Digital Health and Survivorship Care

Although this study did not evaluate implementation strategies, clinical integration models, or interventions targeting usability, the observed patterns of fitness tracker ownership and use may inform future research and program development. As survivorship care increasingly incorporates remote monitoring, the relatively high prevalence of fitness tracker ownership in this cohort (506/893, 56.7%) and frequent use among owners (416/506, 82.2%) suggest that for digitally engaged subgroups, wearables could meaningfully extend assessment beyond the clinic. These rates should not, however, be taken as population-level estimates of what scalable tracker-enabled care can rely on. Wearables may support the identification of emerging functional vulnerability by enabling longitudinal monitoring of mobility-related activity patterns, including sustained reductions in daily step counts and overall activity levels, which have been associated with postoperative recovery trajectories, readmissions, and frailty-related outcomes in cancer and older adult populations [23-25]. In parallel, wearable-derived measures of PA, sedentary behavior, heart rate, and sleep have been proposed as scalable digital indicators of cardiovascular and cardiometabolic risk in cancer survivorship, complementing clinic-based evaluations rather than replacing them [26].

The significant associations between sociodemographic factors and device ownership indicate that equitable access may be the first and most critical step in facilitating fitness tracker adoption. The strong income gradient observed in this study, combined with cost being the most frequently endorsed barrier among non-owners, suggests that financial accessibility remains a meaningful obstacle for lower-income participants. Strategies to address this barrier may include device subsidy or loaner programs, integration of fitness tracker provision into survivorship care plans, or partnerships with device manufacturers to reduce cost. Similarly, the lower ownership observed among male participants and among survivors of certain cancer types suggests that outreach and engagement strategies may need to be tailored to reach populations that are less likely to adopt these technologies independently.

However, inconsistent engagement among device owners highlights that access alone may not be sufficient to support sustained use. Future studies should evaluate implementation approaches that directly address behavioral and usability barriers, including device comfort, cognitive burden, and platform interoperability. Evidence from patients receiving care at federally qualified health centers indicates that baseline wearable ownership is low (21%), despite substantial interest among those without a device (58%), highlighting that effective implementation may require education, onboarding, and user support to address commonly reported usability and knowledge barriers beyond device provision alone, even when economic barriers are addressed [12]. It is important to note, however, that the ownership prevalence in this study cohort was significantly higher, at 56.7%, and the cohort itself showed limited sociodemographic diversity. This suggests that further research involving more diverse study populations is necessary to determine if the sociodemographic patterns observed within this study generalize to populations with greater racial, ethnic, and socioeconomic diversity.

Finally, because most owners in this cohort were already using consumer devices, such as the Apple iWatch and Fitbit, future digital lifestyle interventions may benefit from supporting multidevice integration and bring-your-own-device approaches, allowing cancer survivors to continue using familiar technologies and reducing friction to participation. Evaluating these strategies in more diverse populations and health care settings will be essential to determine how tracker-enabled survivorship approaches can be implemented effectively and equitably.

Strengths and Limitations

This study has several limitations. The cross-sectional design precludes causal inference, and fitness tracker ownership and use were self-reported rather than objectively verified, which may introduce recall or social desirability bias. Second, the dataset lacked clinical characteristics, such as treatment status and recurrence history, which may limit interpretation. In addition, in the use-frequency analysis (Table 3), both the barrier items and the outcome were ascertained simultaneously through the same self-report instrument. Although device-level barriers (eg, discomfort, malfunction) are conceptually separable from use frequency, motivational and cognitive burden items (eg, low motivation, information overload) are difficult to distinguish from the outcome they are intended to predict, raising the possibility of criterion contamination. Future longitudinal designs that assess barriers at baseline and track use patterns prospectively would help clarify whether these factors precede, accompany, or result from changes in engagement.

The sample primarily consisted of non-Hispanic White individuals with a college educated and higher household incomes, which limits the generalizability to more diverse survivor populations. Although several sociodemographic correlates of ownership were detected, including sex, income, and cancer site, the limited racial and ethnic diversity of the sample constrains the ability to identify disparities along these dimensions, and the observed associations may not generalize to populations with greater sociodemographic variability. We were also unable to evaluate device-level characteristics, such as accuracy, brand-specific differences, or wear time

XSL•FO
RenderX

verification, which constrain the interpretation of how technical performance may influence engagement. In particular, several of the barriers identified in the frequent-use model, such as battery life, device malfunction, and limited app or device integration, are plausibly device or brand specific and may vary by manufacturer, device generation, or operating-system ecosystem; because brand-level subgroup analyses were not powered in the sample, these barriers should be interpreted as aggregate user experiences rather than properties of fitness trackers as a class. Finally, although the survey captured additional digital health constructs, the analysis focuses specifically on fitness tracker behaviors to maintain conceptual and analytic clarity.

Despite these limitations, this study has important strengths. The sample included survivors across a broad range of cancer types, and the survey simultaneously assessed demographic, psychosocial, and experiential correlates of fitness tracker use, domains that are not often examined together. By characterizing real-world engagement with consumer-owned devices rather than study-provided fitness trackers, this work provides insight into naturalistic patterns of adoption and sustained use. These findings help fill a critical gap in the literature, which has largely been divided between population-level estimates of device ownership and feasibility studies conducted under structured clinical or research conditions. By integrating both perspectives, this study advances understanding of how survivors adopt and maintain fitness tracker use in everyday contexts, providing a behavioral foundation for developing scalable, equitable digital health interventions to support long-term self-management.

Conclusion

In summary, we conducted one of the largest studies of real-world fitness tracker engagement among cancer survivors and found that within this single-center, demographically homogeneous cohort, although ownership was common, initial ownership was primarily correlated with sociodemographic factors, while sustained use within this sample varied widely and was concurrently associated with self-reported psychosocial and usability factors, which should be interpreted as correlates rather than established antecedent barriers. Because the sample likely reflects a digitally engaged subgroup, these patterns should be interpreted as upper-bound estimates rather than broadly representative figures. These findings highlight the importance of integrating behavioral support, human-centered design, and seamless system-level interoperability into digital survivorship programs. Future research should examine longitudinal associations between fitness tracker–derived activity and clinical outcomes, test interventions targeting motivational and emotional barriers, and develop implementation models that facilitate integration of fitness tracker data into oncology workflows. By strengthening both access and sustained engagement, fitness trackers may help support scalable, personalized approaches to survivorship care.

Acknowledgments

This study was supported by discretionary research funds provided to RMB through the Ohio State University Comprehensive Cancer Center. We also thank the participants who shared their time and insights. The authors wish to acknowledge the late Dr Brian Focht for his invaluable guidance, mentorship, and support throughout the study. This work was made possible, in part, through the use of his laboratory facilities at the Ohio State University. His dedication to the field and to his students continues to inspire this work. The authors thank the Total Cancer Care program staff for facilitating participant recruitment, and Drs Michelle Holko and Tamara Litwin for permission to adapt and reuse their survey instrument originally developed for *npj Digital Medicine*. The authors are grateful to the Lifestyles Lab team members for their contributions to data management and administrative coordination.

Funding

This study was supported by discretionary research funds provided to RMB through the Ohio State University Comprehensive Cancer Center.

Data Availability

The statistical code used to analyze the data in this study is available from the corresponding author upon reasonable request. The datasets generated or analyzed during this study are not publicly available as they contain confidential patient information but can be made available from the corresponding author upon reasonable request.

Authors' Contributions

RMB conceived and designed the study, oversaw data collection and data analysis, conducted the primary interpretation of findings, and led manuscript writing. JLF, RS, and MKT coordinated data management, performed preliminary analyses, and assisted with manuscript writing. ABO contributed to survey adaptation and development of graphical materials. RS, PW, AK, JK-S, BSN, SMSG, MFR, FJP, SK, and EDP contributed to study design, interpretation of results, and critical revision of the manuscript. All authors have read and approved the final manuscript. During the preparation of this manuscript, the authors used ChatGPT (OpenAI) as a writing support tool to proofread for grammar. All content was written, critically reviewed, and edited by the authors, who take full responsibility for the final manuscript.

XSL•FO
RenderX

Conflicts of Interest

Not relevant to the work described in this paper, EDP received grant funding from Pfizer, Merck Foundation, Genentech, Guardant Health, and Astra Zeneca. They served on Advisory Boards for Merck, Novartis, and Glaxo Smith Kline, also unrelated to this work.

Multimedia Appendix 1
Survey instruments.
[PDF File (Adobe PDF File), 121 KB - cancer_v12i1e92876_app1.pdf]

Multimedia Appendix 2
Regression sensitivity analysis.
[DOCX File , 31 KB - cancer_v12i1e92876_app2.docx]

References

1. Gresham G, Schrack J, Gresham LM, Shinde AM, Hendifar AE, Tuli R, et al. Wearable activity monitors in oncology trials: current use of an emerging technology. Contemp Clin Trials 2018 Jan;64:13-21 [FREE Full text] [doi: 10.1016/j.cct.2017.11.002] [Medline: 29129704]
2. Benzo R, Tetrick M, Krok-Schoen J, Brasky T, Washington P, Paskett E, et al. Digital information-seeking behaviors among cancer survivors: associations with sociodemographic determinants, cancer history, and perceived health. J Cancer Surviv 2025 Nov 06:1-12 [FREE Full text] [doi: 10.1007/s11764-025-01927-w] [Medline: 41196508]
3. Rock CL, Thomson CA, Sullivan KR, Howe CL, Kushi LH, Caan BJ, et al. American Cancer Society nutrition and physical activity guideline for cancer survivors. CA Cancer J Clin 2022 May;72(3):230-262 [FREE Full text] [doi: 10.3322/caac.21719] [Medline: 35294043]
4. Campbell KL, Winters-Stone KM, Wiskemann J, May AM, Schwartz AL, Courneya KS, et al. Exercise guidelines for cancer survivors: consensus statement from International Multidisciplinary Roundtable. Med Sci Sports Exerc 2019 Nov;51(11):2375-2390 [FREE Full text] [doi: 10.1249/MSS.0000000000002116] [Medline: 31626055]
5. McTiernan A, Friedenreich CM, Katzmarzyk PT, Powell KE, Macko R, Buchner D, et al. Physical activity in cancer prevention and survival: a systematic review. Med Sci Sports Exerc 2019 Jun;51(6):1252-1261 [FREE Full text] [doi: 10.1249/MSS.0000000000001937] [Medline: 31095082]
6. Schmitz KH, Campbell AM, Stuiver MM, Pinto BM, Schwartz AL, Morris GS, et al. Exercise is medicine in oncology: engaging clinicians to help patients move through cancer. CA Cancer J Clin 2019 Nov;69(6):468-484 [FREE Full text] [doi: 10.3322/caac.21579] [Medline: 31617590]
7. Stout NL, Brown JC, Schwartz AL, Marshall TF, Campbell AM, Nekhlyudov L, et al. An exercise oncology clinical pathway: screening and referral for personalized interventions. Cancer 2020 Jun 15;126(12):2750-2758 [FREE Full text] [doi: 10.1002/cncr.32860] [Medline: 32212338]
8. Chow R, Drkulec H, Im JHB, Tsai J, Nafees A, Kumar S, et al. The use of wearable devices in oncology patients: a systematic review. Oncologist 2024 Apr 04;29(4):e419-e430 [FREE Full text] [doi: 10.1093/oncolo/oyad305] [Medline: 37971410]
9. Singh B, Zopf EM, Howden EJ. Effect and feasibility of wearable physical activity trackers and pedometers for increasing physical activity and improving health outcomes in cancer survivors: a systematic review and meta-analysis. J Sport Health Sci 2022 Mar;11(2):184-193 [FREE Full text] [doi: 10.1016/j.jshs.2021.07.008] [Medline: 34314878]
10. Keats M, Yu X, Sweeney Magee M, Forbes C, Grandy S, Sweeney E, et al. Use of wearable activity-monitoring technologies to promote physical activity in cancer survivors: challenges and opportunities for improved cancer care. Int J Environ Res Public Health 2023 Mar 08;20(6):4784 [FREE Full text] [doi: 10.3390/ijerph20064784] [Medline: 36981693]
11. Kos M, Brouwer CG, van Laarhoven HW, Hopman MT, van Oijen MG, Buffart LM. The association between wearable device metrics and clinical outcomes in oncology: a systematic review with evidence synthesis and meta-analysis. Crit Rev Oncol Hematol 2023 May;185:103979 [FREE Full text] [doi: 10.1016/j.critrevonc.2023.103979] [Medline: 37001837]
12. Holko M, Litwin TR, Munoz F, Theisz KI, Salgin L, Jenks NP, et al. Wearable fitness tracker use in federally qualified health center patients: strategies to improve the health of all of us using digital health devices. NPJ Digit Med 2022 Apr 25;5(1):53 [FREE Full text] [doi: 10.1038/s41746-022-00593-x] [Medline: 35469045]
13. Wan W, Li V, Chin MH, Faldmo DN, Hoefling E, Proser M, et al. Development of PRAPARE social determinants of health clusters and correlation with diabetes and hypertension outcomes. J Am Board Fam Med 2022;35(4):668-679 [FREE Full text] [doi: 10.3122/jabfm.2022.04.200462] [Medline: 35896473]
14. Miyaji T, Kawaguchi T, Azuma K, Suzuki S, Sano Y, Akatsu M, et al. Patient-generated health data collection using a wearable activity tracker in cancer patients-a feasibility study. Support Care Cancer 2020 Dec 12;28(12):5953-5961 [FREE Full text] [doi: 10.1007/s00520-020-05395-z] [Medline: 32281031]
15. Zhou W, Cho Y, Pu J, Shang S. Trends in wearable device use among cancer survivors in the United States from 2019 to 2022. J Geriatr Oncol 2024 May;15(4):101729 [FREE Full text] [doi: 10.1016/j.jgo.2024.101729] [Medline: 38360455]

XSL•FO
RenderX

16. Faro JM, Yue K, Singh A, Soni A, Ding EY, Shi Q, et al. Wearable device use and technology preferences in cancer survivors with or at risk for atrial fibrillation. Cardiovasc Digit Health J 2022 Dec;3(6 Suppl):S23-S27 [FREE Full text] [doi: 10.1016/j.cvdhj.2022.08.002] [Medline: 36589761]
17. Nagappan A, Krasniansky A, Knowles M. Patterns of ownership and usage of wearable devices in the United States, 2020-2022: survey study. J Med Internet Res 2024 Jul 26;26:e56504 [FREE Full text] [doi: 10.2196/56504] [Medline: 39058548]
18. Hong YA. Mobile health and behavior tracking (mHBT) among cancer survivors: results from a large and diverse sample. Mhealth 2024;10:5 [FREE Full text] [doi: 10.21037/mhealth-23-58] [Medline: 38323153]
19. Nilanon T, Nocera LP, Martin AS, Kolatkar A, May M, Hasnain Z, et al. Use of wearable activity tracker in patients with cancer undergoing chemotherapy: toward evaluating risk of unplanned health care encounters. JCO Clin Cancer Inform 2020 Sep;4:839-853 [FREE Full text] [doi: 10.1200/CCI.20.00023] [Medline: 32970482]
20. Ng A, Gupta E, Bansal S, Fontillas RC, Amos CE, Williams JL, et al. Cancer patients' perception of usefulness of wearable exercise trackers. PM R 2021 Aug;13(8):845-851 [FREE Full text] [doi: 10.1002/pmrj.12475] [Medline: 32844592]
21. Deeks A, Lombard C, Michelmore J, Teede H. The effects of gender and age on health related behaviors. BMC Public Health 2009 Jun 30;9(1):213 [FREE Full text] [doi: 10.1186/1471-2458-9-213] [Medline: 19563685]
22. Courneya KS, Booth CM. Exercise as cancer treatment: a clinical oncology framework for exercise oncology research. Front Oncol 2022;12:957135 [FREE Full text] [doi: 10.3389/fonc.2022.957135] [Medline: 36119508]
23. Jonker LT, Lahr MMH, Oonk MHM, de Bock GH, van Leeuwen BL. Post-discharge telemonitoring of physical activity, vital signs, and patient-reported symptoms in older patients undergoing cancer surgery. Ann Surg Oncol 2021 Oct 27;28(11):6512-6522 [FREE Full text] [doi: 10.1245/s10434-021-09707-3] [Medline: 33641013]
24. de Leeuwerk ME, Botjes M, van Vliet V, Geleijn E, de Groot V, van Wegen E, et al. Self-monitoring of physical activity after hospital discharge in patients who have undergone gastrointestinal or lung cancer surgery: mixed methods feasibility study. JMIR Cancer 2022 Jun 24;8(2):e35694 [FREE Full text] [doi: 10.2196/35694] [Medline: 35749165]
25. Friedrich B, Elgert L, Eckhoff D, Bauer JM, Hein A. A system for monitoring the functional status of older adults in daily life. Sci Rep 2023 Jul 31;13(1):12396 [FREE Full text] [doi: 10.1038/s41598-023-39483-x] [Medline: 37524888]
26. Benzo RM, Gogineni A, Tetrick MK, Singh R, Washington P, Fernandez S, et al. mHealth technologies in research studying cardiovascular health in cancer: a systematic review. PLOS Digit Health 2025 Sep;4(9):e0001027 [FREE Full text] [doi: 10.1371/journal.pdig.0001027] [Medline: 40996972]

Abbreviations

aOR: adjusted odds ratio
CONSORT: Consolidated Standards of Reporting Trials
HINTS-SEER: Health Information National Trends Survey – Surveillance, Epidemiology, and End Results
mHealth: mobile health
OR: odds ratio
PA: physical activity
REDCap: Research Electronic Data Capture
TCC: Total Cancer Care
VIF: variance inflation factor

Edited by M Balcarras; submitted 07.Feb.2026; peer-reviewed by WJ He; comments to author 23.Mar.2026; revised version received 03.May.2026; accepted 05.May.2026; published 26.May.2026.

Please cite as:
Benzo RM, Fisher JL, Tetrick MK, Singh R, Osei AB, Krok-Schoen J, Nickerson BS, Paskett ED, Kargarandehkordi A, Kumar S, St George SM, Rayo MF, Penedo FJ, Washington P
Correlates of Fitness Tracker Ownership and Use in Cancer Survivors: Cross-Sectional Survey
JMIR Cancer 2026;12:e92876
URL: https://cancer.jmir.org/2026/1/e92876
doi:10.2196/92876
PMID:

©Roberto M Benzo, James L Fisher, Macy K Tetrick, Rujul Singh, Alex B Osei, Jessica Krok-Schoen, Brett S Nickerson, Electra D Paskett, Ali Kargarandehkordi, Sachin Kumar, Sara M St George, Michael F Rayo, Frank J Penedo, Peter Washington. Originally published in JMIR Cancer (https://cancer.jmir.org), 26.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use,

XSL•FO
RenderX

distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Symptom Management Following Radioactive Iodine Treatment in Adults With Thyroid Cancer: Usability Evaluation and Proof-of-Concept Study

Alaina Carr[1], PhD; Angela M Jenkins[1], BS; Valeria Gómez[1], BS; Janine Alexandria Lorenzo[1]; Ryan Moclair[1,2], BS; James Pocchia[1,2], BS; Devansh Gandhi[1,2], BS; Kristen Miller[3], CPPS, MSPH, MSL, DrPH; Kristi D Graves[1], PhD; Jacqueline Jonklaas[4], MPH, MD, PhD

[1]Department of Oncology, Georgetown University Medical Center, Washington, DC, United States

[2]Department of Biology, College of Arts and Sciences, Georgetown University, Washington, DC, United States

[3]National Center for Human Factors in Healthcare, MedStar Health Research Institute, Washington, DC, United States

[4]Division of Endocrinology, Georgetown University, Washington, DC, United States

Corresponding Author:
Alaina Carr, PhD
Department of Oncology
Georgetown University Medical Center
2115 Wisconsin Ave, NW
Suite 300
Washington, DC, 20007-2214
United States
Phone: 1 2026870958
Email: alc343@georgetown.edu

Abstract

Background: Thyroid cancer is often labeled a “good cancer” due to high survival rates. However, many patients diagnosed with differentiated thyroid cancer (DTC) treated with radioactive iodine (RAI) experience short- and long-term symptoms as a result of treatment. These include salivary dysfunction, dry mouth, dry eyes, and nausea. Despite these challenges, patients report limited knowledge and unmet informational needs regarding RAI symptoms and symptom management. Evidence-based, patient-centered education to support symptom management during DTC survivorship is needed. Our team developed Radioactive Iodine Support (RAI Support), a prototype of a 4-week web-based intervention for RAI symptom management among survivors of DTC.

Objective: This study aimed to evaluate the usability of the RAI Support intervention among survivors of DTC who completed RAI treatment and assess preliminary effect sizes of change in RAI symptom burden, self-efficacy for managing chronic symptoms, informational needs, global physical and mental health, and health-related quality of life (HRQOL).

Methods: We conducted a longitudinal usability testing study with 8 participants over a 4-week period. At baseline (T1), participants completed an online survey and a 90-minute virtual usability testing session. During usability testing sessions, study staff used a think-aloud protocol to elicit feedback on the content, language, and visual appeal of the prototype across 5 tasks. Participants also completed the System Usability Scale (SUS). Participants were encouraged to continue using the intervention over the 4-week period, during which use data were recorded. A follow-up after assessment (T2) was completed at the end of the 4-week period of home use. Survey outcomes included RAI symptom burden, self-efficacy for managing chronic symptoms, informational needs, global physical and mental health, and HRQOL.

Results: Among the 8 survivors of DTC who had undergone RAI treatment (age: mean 45, SD 8.2 years), the mean SUS score was 88.1 (SD 8.7), indicating above-average usability. All participants effectively completed each task, with ease of use ranging from 100% to 87.5% and error rates ranging from 0% to 25%. Preliminary *r* effect size signals of change were found for improved self-efficacy, global mental health, and HRQOL (*r* values ranged from 0.31 to 0.45).

Conclusions: Survivors of DTC reported above-average usability of the RAI Support prototype and completed each task successfully. Variability in error rates and ease of use suggests that refinements are needed to further optimize usability and

XSL•FO
RenderX

patient engagement. RAI Support also shows preliminary promise in improving self-efficacy for symptom management, global mental health, and HRQOL among survivors of DTC. Findings support future pilot testing.

(JMIR Cancer 2026;12:e84151) doi:10.2196/84151

KEYWORDS

internet-based intervention; iodine radioisotopes; self-efficacy; survivorship; symptom burden; thyroid neoplasms; user-centered design

Introduction

Thyroid cancer is frequently referred to as a "good cancer" due to its excellent 5-year survival rates and favorable prognosis [1-5]. However, this label often minimizes the lived experiences that many patients undergo. Radioactive iodine (RAI) treatment is not routinely indicated for all patients with thyroid cancer. Those who do receive RAI typically have undergone total or near-total thyroidectomy, after which postoperative risk stratification incorporating histopathology, staging, and imaging determines whether RAI is indicated for remnant ablation, adjuvant therapy, or treatment of known residual or recurrent disease [6]. Because RAI is used selectively based on these characteristics, those who do receive it often represent a subgroup with higher-risk cancer who correspondingly experience a longer course of surveillance and symptom management [6,7]. These patients report significant and persistent side effects such as salivary gland dysfunction, dry mouth, eye irritation, and nasal discomfort [1,8,9]. These symptoms can negatively impact quality of life, especially in the absence of coordinated, ongoing support [10,11]. Survivors of thyroid cancer also report lower health-related quality of life (HRQOL) than other groups of survivors of cancer [11,12], perhaps due, in part, to symptom burden [11]. Following RAI treatment, care coordination for symptom burden is often inconsistent across health care settings and typically involves multiple subspecialists, including endocrinologists, nuclear medicine physicians, oral health specialists, and otolaryngologists, which may contribute to fragmented follow-up care [10-13].

As cancer survivorship care remains fragmented and increasingly extends beyond traditional clinical settings, scalable strategies are needed to support patient engagement and symptom management throughout the continuum of care [14]. Digital health interventions have demonstrated potential for addressing gaps in survivorship care after primary treatment, particularly by enhancing access to timely support and empowering patients to manage chronic symptoms [15,16]. Despite the complexities of thyroid cancer survivorship care, we are unaware of any dedicated digital health tools to support symptom self-management after RAI treatment, presenting a critical gap in the continuum of care for this population.

We developed Radioactive Iodine Support (RAI Support) as a web-based intervention to fill this unmet need by providing tailored guidance and symptom support to survivors of thyroid cancer following RAI treatment. A number of systematic reviews have highlighted the facilitators and barriers to behavioral intervention adherence due to digital technology–related factors, including website usability ratings, ease of use, plain language, and connectivity [15-17]. One systematic review identified that low ease of use and connectivity issues were associated with 8% to 14% of intervention nonadherence [17]. The literature also highlights that the successful implementation of sustainable digital tools depends heavily on their usability, accessibility, and alignment with patient needs [18,19]. Yet, as with all digital interventions, optimal usability is a necessary component of sustained user engagement, which in turn relates to the efficacy of the digital tool. Human factors, such as interface design, ease of navigation, and relevance of content, play central roles in determining whether users adopt, engage with, and benefit from such tools [16]. Without attention to these factors, even well-designed interventions may fail to translate into real-world impact. Therefore, we conducted a usability and proof-of-concept study to (1) optimize the usability of the web-based intervention and (2) provide an initial exploration of whether RAI Support improves clinically significant end points, with the aim of addressing the complex needs of survivors of thyroid cancer in preparation for future pilot testing.

Methods

Study Design

From September to November 2024, we conducted usability testing in a single-arm longitudinal proof-of-concept study at an academic medical center in the mid-Atlantic United States. A sample size of 8 participants was selected for the usability evaluation, consistent with the study's objective of identifying usability issues [20] with RAI Support and obtaining user feedback to inform intervention optimization rather than testing intervention efficacy.

Recruitment and Setting

Adult patients were identified from medical record review and through a thyroid cancer nonprofit organization's publicly available Facebook (Meta Platforms Inc) page, on which the study was advertised. The 2 sources for recruitment were intended to understand the spectrum of experiences of survivors of differentiated thyroid cancer (DTC), RAI symptom burden profiles, and website feedback from individuals across the United States who were treated across a variety of hospital settings.

Eligible patients met the following inclusion criteria: (1) age of ≥18 years; (2) able to provide informed consent in English; (3) diagnosed with DTC (papillary or follicular carcinoma) in the past 3 years; (4) received RAI treatment within the past 3 years; (5) access to a computer, tablet, or smartphone with internet connectivity; and (6) completion of the Radioiodine-6 screener [21] with a total score of ≥7.6 or a score of ≥3 on at least 1 out

XSL•FO
RenderX

of the 3 RAI symptom subitems (mouth, eyes, or nasal). Exclusion criteria included non–English-speaking patients and individuals with cognitive, psychiatric, or hearing impairments that could limit their ability to consent or participate in the study.

Following institutional review board approval (#00005285), potential participants were contacted through introductory emails or mailed study flyers, followed by a Health Insurance Portability and Accountability Act (HIPAA)–compliant videoconference with research staff to assess eligibility through the completion of a screening questionnaire. Informed consent was obtained through a 15-minute telephone call or a HIPAA-compliant videoconference. Participants reviewed and signed the informed consent form electronically. A total of 8 patients diagnosed with DTC and treated with RAI within the past 3 years were enrolled.

Imposter Participants

Our Facebook recruitment approach also led to 29 "imposter participants," defined as individuals who provide false identities and information for the purpose of participating in research studies [22-24]. Imposter participants were identified by a combination of the following indicators: nonworking telephone numbers or telephone numbers for businesses, refusal to turn on their cameras (a requirement for usability participation), use of similar (yet unusual) email domains, and inconsistent internet connectivity.

To increase rigor and guard against imposter participants, study staff modified the screening procedure to include open-ended questions about the name of the hospital or clinic setting in which participants received treatment. We then ended Facebook recruitment to reduce the likelihood of additional imposter participants. Of note, imposter participants were not included in the final data analysis.

Data Collection

Participants completed baseline and postintervention surveys via REDCap (Vanderbilt University) [25] (for details, refer to the Measures section).

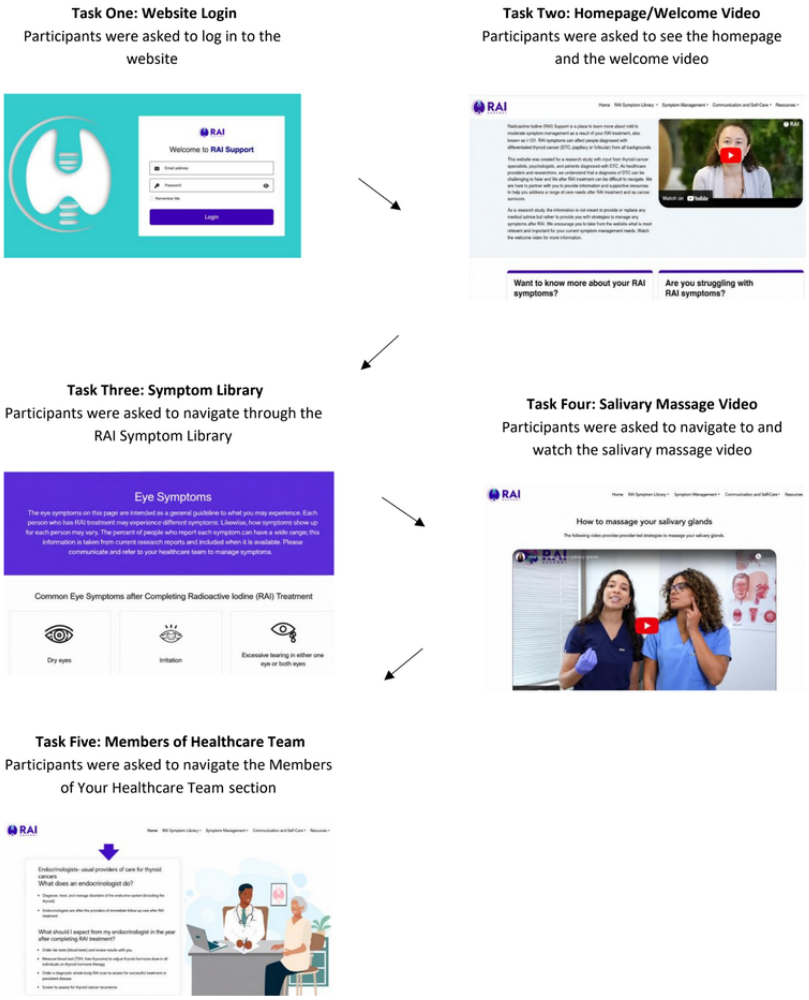
Study staff trained in qualitative research methods developed the semistructured usability testing interview guides (including author AC, a female psychologist, author KM, a female human factors scientist, and a nonauthor, female research and human factors scientist). Interviews lasted approximately 90 minutes.

Study staff (authors) trained in qualitative research conducted the interviews: AMJ (a female research assistant), JP (a male research assistant), and RM (a male research assistant). A second member of the research team (JAL, a female research assistant, and JP), trained in qualitative data research, took detailed notes during the interviews to track usability task outcomes.

Study staff used a think-aloud method to prompt participants to provide feedback as they navigated key website elements, including the log-in process, home page, welcome video, symptoms page, salivary symptom management page, and resource page. Feedback focused on content clarity, visual appeal, and ease of navigation (Figure 1). All recorded usability testing interviews were transcribed verbatim and deidentified prior to data analysis. At the end of the usability testing interviews, participants completed a System Usability Scale (SUS) questionnaire [26] via REDCap. Participants were then asked to continue using RAI Support for 4 weeks. At the conclusion of the 4 weeks, they completed a follow-up questionnaire.

XSL•FO
RenderX

Figure 1. Screenshots of web-based Radioactive Iodine Support (RAI Support) prototype.



Intervention

RAI Support has been previously described [9]. RAI Support is a theoretically driven web-based intervention to support self-management of RAI symptoms among survivors of DTC. RAI Support is informed by social cognitive theory [27] and was developed following person-based design principles, with survivors of DTC and health care providers (endocrinologists, endocrine surgeons, nuclear medicine providers, registered oncology dietitians, social workers, and psycho-oncologists) as integral members of the intervention development process [28]. The web-based intervention is asynchronous and consists of content related to education on RAI symptoms; brief provider-led videos on symptom management strategies for mild to moderate symptoms; communication and self-care, including mindfulness exercises; and DTC survivorship resources. Website programming was produced by DreamSoft4u, Pvt. Ltd [29].

Measures

Usability Evaluation Outcome Measures: SUS

The SUS is a 10-item questionnaire that assesses patients' perceived usability of a system [26]. Items are rated on a 5-point Likert scale (1="Strongly disagree" to 5="Strongly agree"), with composite scores ranging from 0 to 100. A SUS score >68 represents above-average usability [26].

Effectiveness is defined as the success rate or the degree to which a participant can successfully achieve their goal in using the system. For the present study, effectiveness is assessed dichotomously (yes or no) to indicate the ability to complete a task.

Ease of use is the degree to which a participant can complete a task on their very first attempt without training. For the present study, ease of use is assessed dichotomously (yes or no) and refers to the ability to complete a task the first time.

Error rate is the degree to which the interaction between the participant and the system is smooth and/or the type of mistakes a participant makes while completing each task. For the present study, error rate is assessed on a 4-point Likert scale (1="cosmetic or visual error," 2="minor error or user recovers quickly," 3="major error or user recovers with difficulty," and 4="catastrophic error or system crashes"), with higher scores indicating greater severity of error.

Intervention Outcome Measures

Salivary, Lacrimal, and Nasal Measure

The salivary, lacrimal, and nasal (SALANS) [8] measure consists of 33-items and assesses SALANS symptoms among

XSL•FO
RenderX

patients with DTC after RAI treatment on a 5-point Likert scale (0="Not at all" to 4="Very much"). The scale has 6 subscales (sialadenitis, taste, xerostomia, dry eyes, epiphora, and nasal). Higher scores indicate greater symptom burden. The internal consistency α coefficient for the current sample is 0.98.

Functional Assessment of Cancer Therapy-General-7 Item

The Functional Assessment of Cancer Therapy-General-7 Item (FACT-G7) [30] is a 7-item questionnaire that assesses physical, emotional, and functional well-being of adult oncology patients. Each item is rated on a 5-point Likert scale (0="Not at all" to 4="Very much"). Scores range from 0 to 28, with higher total scores indicating greater HRQOL. The α coefficient for the current sample is 0.88.

Patient-Reported Outcomes Measurement Information System Scale

The Patient-Reported Outcomes Measurement Information System (PROMIS) Scale version 1.2 Global Health is a 4-item assessment of physical and mental health measured on a 5-point Likert scale [31]. Raw summed scores range from 4 to 20 and are converted to T-scores, with a mean of 50 and an SD of 10. A higher T-score indicates greater physical and mental health. The α coefficient for the current sample is 0.92.

PROMIS Self-Efficacy for Managing Chronic Conditions–Managing Symptoms

The PROMIS Self-efficacy for Managing Chronic Conditions–Managing Symptoms is a 28-item questionnaire that evaluates physical, mental, and social health in individuals with chronic conditions, with the T-score (mean 50, SD 10) representing the final score [32]. A higher score indicates higher self-efficacy. The α coefficient for the current sample is 0.96.

Informational Needs Questionnaire for DTC

The Informational Needs Questionnaire for DTC (INQ-DTC) contains 33 items and assesses informational needs of patients after RAI treatment [33]. Items are measured on a 5-point Likert scale (1="Not at all" to 5="Very"). Higher total scores indicate greater informational needs. The α coefficient for the current sample is 0.93.

Exit Questionnaire

Participants completed an exit questionnaire to assess their perceptions of the helpfulness of components of the RAI Support intervention, likelihood of recommending the program to survivors of DTC, satisfaction with intervention length, and overall satisfaction with the program. Responses were assessed using 10-point Likert scales for the majority of items.

Intervention Back-End Use Data

Intervention back-end use data were captured through the RAI Support back-end system. Metrics captured included frequency of log-ins and number of website pages visited. Use data were exported in aggregate form for analysis to complement self-report measures of intervention engagement.

Data Analysis

A mixed methods approach was used, incorporating quantitative data from questionnaires, usability logs, and the web-based intervention back-end system, as well as qualitative data from usability interviews for analysis. Descriptive statistics were used to summarize the demographic and clinical characteristics and web-based intervention use of participants. Signals of change in outcomes of interest were evaluated using within-group effect sizes calculated with the Rosenthal effect size formula ($r=Z/\sqrt{N}$), where Z is the standardized Wilcoxon signed-rank test statistic and N is the number of paired observations [34,35]. Cohen guidelines for r effect sizes were used and defined as "small" for $r=0.1$; "medium" for $r=0.3$; "large" for $r=0.5$ [35]. Signed effect sizes were reported to indicate the direction of change, with directionality defined based on follow-up minus baseline. Positive values indicate increases from baseline, whereas negative values indicate decreases.

Content analysis was used to analyze the usability testing interviews [36,37]. To gain an initial understanding of the content, 3 authors (AC, AMJ, and JAL) independently reviewed the interview transcripts and met biweekly to identify and refine preliminary codes and examine data discrepancies until team resolution. Coding was conducted using ATLAS.ti Scientific Software Development GmbH (version 8.15.3; Lumivero LLC). Three authors (AC, AMJ, and JAL) collaboratively developed a codebook, consolidating initial codes into broader categories through iterative discussion. These categories were further synthesized into major themes and subthemes to capture key patterns in participant feedback regarding the context underlying usability issues and potential methods for optimization. Thematic saturation [38] was operationally defined as the point at which review of interviews did not identify new themes or subthemes relevant to usability. Using this approach, no new themes or subthemes emerged from the final interview data, indicating thematic saturation within the scope of the usability evaluation.

Ethical Considerations

Ethics approval was obtained in accordance with the principles of the Declaration of Helsinki. The protocol was reviewed and approved by the Georgetown University Institutional Review Board (#00005285). All participants provided signed informed consent. Study data were deidentified prior to data analysis. Participants were compensated with a US $50 gift card for study participation.

Results

Recruitment

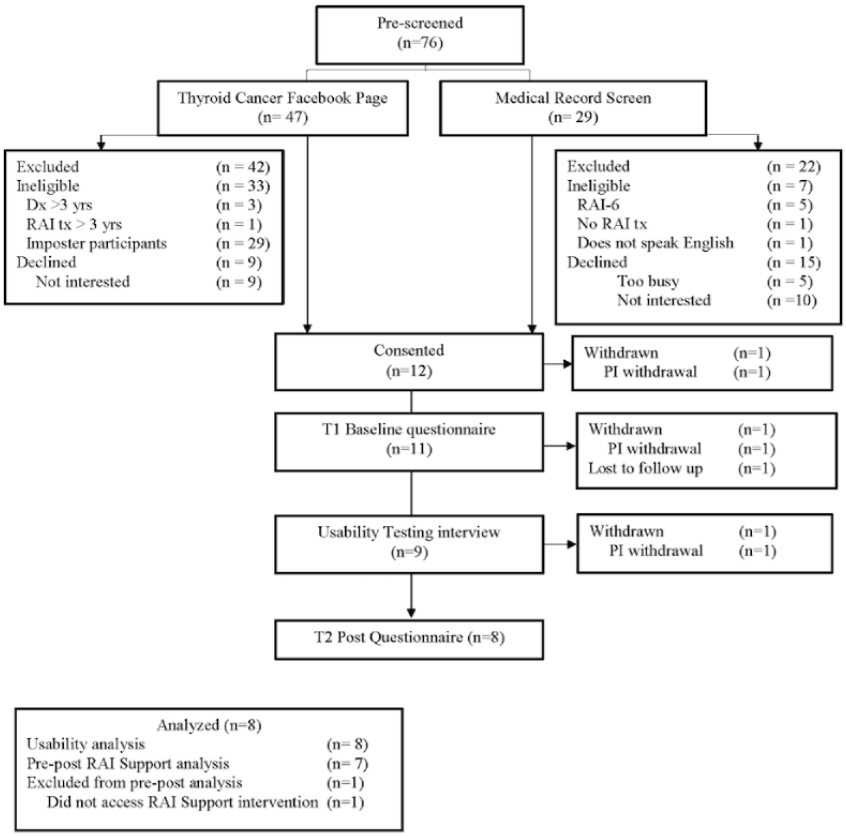
A total of 76 individuals were assessed for eligibility (Figure 2). A total of 47 (61.8%) individuals contacted study staff in response to Thyroid Cancer Survivors' Association Inc (ThyCa) Facebook advertisements, and 29 (38.2%) individuals were screened for eligibility from medical records. In total, 24 participants declined study participation: 19 were not interested and 5 were too busy. Overall, 64 individuals were excluded, including 24 who declined and 40 who were ineligible for the following reasons: 3 were diagnosed with DTC >3 years ago, 1 received RAI >3 years ago, 1 did not receive RAI, 1 was not an English speaker, 29 were imposter participants, and 5 did

XSL•FO
RenderX

not meet screening criteria. Among those who were eligible, 12 participants consented to the study.

Of the total sample of consented participants (n=12), 1 was lost to follow-up and 3 withdrew, leaving a final sample of 8 consented participants who completed baseline, usability testing, and postassessment procedures. In the final study sample (N=8), the average age of participants was 45 (SD 8.2; range 35-61) years. In terms of clinical characteristics, the average times since diagnosis and RAI treatment were 1.2 (SD 1.1) and 0.6 (SD 0.4) years, respectively. Most participants were diagnosed with papillary thyroid cancer (7/8, 87.5%), and most received an RAI dose of 100 to 150 mCi (n=6, 75%). See Table 1 for demographics and clinical characteristics.

Figure 2. Study flow diagram. PI: principal investigator; RAI: radioactive iodine.



XSL•FO
RenderX

Table 1. Demographics and clinical characteristics of participants (N=8).

Variables	Participants
Age (years), mean (SD; range)	45 (8.2; 35-61)
Sex, n (%)	
Male	1 (12.5)
Female	7 (87.5)
Race or ethnicity, n (%)	
Asian	2 (25)
Black or African American	1 (12.5)
White	5 (62.5)
Relationship status, n (%)	
Married	4 (50)
Divorced	1 (12.5)
Never married	3 (37.5)
Education, n (%)	
Associate degree or some college	1 (12.5)
College degree or higher	7 (87.5)
Employment status, n (%)	
Employed (full-time)	5 (62.5)
Employed (part-time)	1 (12.5)
Self-employed	1 (12.5)
Unemployed	1 (12.5)
Health insurance (private), n (%)	8 (100)
Household income (US $), n (%)	
No income	1 (12.5)
100,000-149,999	2 (25)
150,000-199,999	1 (12.5)
≥200,000	4 (50)
Disability, n (%)	
Yes	1 (12.5)
No	7 (87.5)
Time since diagnosis (years), mean (SD; range)	1.2 (1.1; 0.5-3.0)
Time since RAI[a] treatment (years), mean (SD; range)	0.64 (0.4; 0.13-1.26)
Cancer histology, n (%)	
Papillary	7 (87.5)
Do not know	1 (12.5)
RAI dose (millicuries), n (%)	
30-75	1 (12.5)
76-100	1 (12.5)
101-150	6 (75)

[a]RAI: radioactive iodine.

XSL•FO
RenderX

Usability Evaluation

Eight participants completed usability testing sessions. The mean SUS score was 88.1 (SD 8.7), indicating above-average usability. All participants effectively completed each task. Ease of use, as measured by completion of a task on the first attempt, ranged from 100% to 87.5%, and error rates ranged from 0% to 25%. Participants rated the clarity of the welcome page (mean 4.3, SD 0.8) and the welcome video (mean 5.0, SD 0; task 1), indicating that the page was perceived as clear and the video as very clear. Mobile device screen sharing blocked the visuals needed to complete tasks 3 (RAI symptom library) and 4 (health care provider–led salivary gland massage video) for 1 participant. When the screen sharing issue was resolved with prompting from study staff, ease of use for this participant increased to 100% for task 5. Refer to Table 2 for an overview of the usability outcome metrics, including error rates by task.

Table 2. Usability evaluation outcomes (N=8).

Tasks	Participants
Task 1: website log in	
Effectiveness	8 (100)
Ease of use	8 (100)
Error: none	8 (100)
Task 2: home page and welcome video	
Effectiveness	8 (100)
Ease of use	8 (100)
Error: none	8 (100)
Task 3: radioactive iodine symptom library	
Effectiveness	8 (100)
Ease of use	7 (87.5)
Error	
None	6 (75)
Cosmetic	1 (12.5)
Minor	1 (12.5)
Task 4: health care provider–led salivary massage video	
Effectiveness	8 (100)
Ease of use	7 (87.5)
Error	
None	6 (75)
Cosmetic	1 (12.5)
Minor	1 (12.5)
Task 5: members of your health care team	
Effectiveness	8 (100)
Ease of use	8 (100)
Error	
None	7 (87.5)
Cosmetic	0 (0)
Minor	1 (12.5)

User Testing Feedback

Participants reported that they found the website content “informative,” “helpful,” and that they would recommend the website to someone else experiencing RAI symptoms. User testing feedback provided insight into barriers to user engagement, contextual information about the root causes of barriers to user engagement with the prototype, and areas for improvement. Overall, participants highlighted navigation across website pages as a primary area for improvement. Participants noted challenges in locating specific content and suggested adding hyperlinks to facilitate navigation. Participants also reported being timed out of the website and suggested increasing the session idle time. Example suggestions for website improvement and refinement are illustrated in Table 3.

XSL•FO
RenderX

Table 3. User testing feedback.

Domains	Exemplar quotes	Refinements
Navigation (all elements related to moving between website pages and additional functions to enhance user engagement with the website)		
Hyperlinks	• "There should be like more of a connection between and leading you where to go. Or if you want to learn more about this, 'Go here.' I wanted to be guided to something more, instead of having to go look for it." [PID[a] 16] • "I think, clicking to the solution. A direct click might have been beneficial." [PID 58]	Incorporation of hyperlinks to facilitate user navigation within the system
Session time-out	• "Let's see. I have to sign back in...Now, that might be something want to look into." [PID 59] • "Oops got logged out...Look at it logged out again." [PID 16]	Increase in the session idle time
Design (reference to the visual appeal, length, formatting, or other design elements of the web-based platform, regardless of task or section of website)	• "I think the fonts size has to be a little bit more uniform." [PID 49] • "Making it a less medicalized and more down to earth like you're talking more to an actual human being instead of somebody specializes in it [thyroid cancer]." [PID 52] • "It's just three paragraphs. When you use bullets where it's quick that makes for a better, cleaner, read." [PID 59]	Inclusion of bullet points to ease readability and understanding of complex information
Features: welcome page or video (area for refinement to acclimate the user to key features of the intervention)	• "I didn't know if the [welcome] video was going to lead into more interactive things or showing what maybe clicking around or how to use the site." [PID 58] • "I don't know if you want to include the navigation here as an additional video." [PID 65]	Refinement of the welcome video to include website navigation instructions
Features: interactive (areas for refinement to reduce the static nature of the intervention)	• "Graphic visuals, pop ups or things that move across the screen. Just so, it's not like quite as static and just information sitting there. Make a little more flash to it, or activity." [PID 16] • "This is not static content that you're presenting... And when I hover on top of this card, perhaps you can then pop that information out, or pop it as a mobile dialogue, or expand on the key concept, on this one box." [PID 65]	Incorporation of interactive features to enhance user engagement

[a]PID: participant ID.

Exit Questionnaire and Intervention Back-End Use Outcomes

Participant scores on the exit questionnaire (n=7) indicate moderate to high ratings of helpfulness for the website videos: managing dry mouth (mean 7.14, SD 2.11), changes in taste (mean 6.5, SD 2.65), and relaxation and mindfulness (mean 6.83, SD 1.94). Back-end use data indicate that, on average, 6 (85.7%) of the 7 website users engaged with 3 out of the 4 content sections during the 4-week home intervention period. Refer to Multimedia Appendices 1 and 2 for the participant exit questionnaire and back-end metrics.

Preliminary Intervention Outcomes

Preliminary effect size signals of change were calculated for participants who completed baseline and postassessment measures (n=7). The distributions of scores on outcomes of interest are presented in Table 4. Overall, preliminary effect size signals of change were found in self-efficacy (r=0.32), global mental health (r=0.31), and HRQOL (r=0.45), with positive effect sizes indicating increases from baseline. A small negative effect size was observed for RAI informational needs (r=−0.15), indicating a favorable decrease in informational needs from baseline. RAI symptom burden and psychosocial informational needs remained relatively stable throughout the study, and global physical health decreased across study time points.

XSL•FO
RenderX

Table 4. Signals of change from baseline to postintervention assessment with RAI Support (N=7).

Measures	Baseline assessment	Postintervention assessment	Changes in mean from baseline	*r* effect sizes[a]
RAI[b] symptom burden[c], mean (SD; range)	62.3 (38.7; 35-148)	63.0 (36.6; 38-144)	+0.7	+0.03
Self-efficacy, mean (SD; range)	40.6 (5.8; 31.4-50.3)	44.2 (12.2; 30.6-69.5)	+3.6	+0.32
Informational needs , mean (SD; range)				
RAI	28.0 (2.2; 24-30)	25.7 (5.8; 14-30)	−2.3	−0.15
Psychosocial	19.4 (5.7; 10-25)	19.1 (5.2; 10-25)	−0.3	0
Global mental health, mean T-score (SE; range)	40.5 (3.45; 26.1-56.7)	42.5 (3.68; 21.3-48.8)	+1.9	+0.31
Global physical health, mean T-score (SE; range)	43.9 (4.11; 19.7-50.7)	42.7 (3.94; 19.7-50.7)	−1.2	−0.55
HRQOL, mean (SD; range)	15.0 (7.3; 0-22)	15.9 (7.5; 1-23)	+0.9	+0.45

[a]Cohen *r* effect size: 0.1="small," 0.3="medium," and 0.5="large."

[b]RAI: radioactive iodine.

[c]RAI symptom burden is measured by the salivary, lacrimal, and nasal (SALANS) measure.

Discussion

Principal Results

Findings from this study add to the literature by showing high usability and preliminary proof-of-concept data for a website designed to support symptom management following RAI treatment among patients with DTC. In addition, this is among the first studies to evaluate the usability of a web-based intervention for symptom management in an understudied population of survivors of DTC. Overall, participants found the RAI Support intervention to be informative and the health care provider–led videos to be helpful. Participant feedback identified user-centered areas to enhance the intervention's ease of use, including navigation across webpages and interactive features to foster user engagement. Future adequately powered studies should further examine temporal patterns and predictors of intervention engagement, including whether sustained or repeated use varies by baseline RAI symptom burden, clinical characteristics, or recruitment source, to better understand for whom and under what circumstances the RAI Support intervention is most engaging. These focal points are consistent with broader gaps identified in the implementation literature on digital health interventions in thyroid cancer care. A recent scoping review of digital health interventions in thyroid cancer care synthesized intervention characteristics, effectiveness, and uptake, as well as reported barriers and facilitators, and identified few interventions focused on symptom self-management following primary treatment [39], highlighting the limited availability of digital approaches addressing this need. Within this context, among participants with complete pre-post RAI intervention data, we found evidence of small to medium clinically meaningful changes in self-efficacy for managing chronic symptoms, RAI informational needs, global mental health, and HRQOL. These findings support the potential of digital interventions such as RAI Support as a valuable and feasible approach to address symptom management needs, with further research and refinement likely to enhance their utility.

Results did not support RAI Support's preliminary effect size signals of change on RAI symptom burden, informational needs for personal growth, and global physical health. One explanation for these findings is that the RAI Support intervention is intended to provide noncurative symptom management strategies, which may not eliminate the persistent nature of the frequency and severity of RAI symptoms [40-42]. Additionally, these data did not capture information on common medical comorbidities among survivors of DTC such as myocardial infarction, osteoporosis, hyperthyroidism, and hypothyroidism [10,11], which may explain the pre-post differences in PROMIS Global Physical Health scale scores in the sample. Future studies with larger samples that capture common medical comorbidities are warranted to further evaluate these outcomes. Another explanation for these findings is that the static informational content of the website may not fully address the range of patient needs. Future research could explore more integrative and interactive features, such as direct messaging with health care professionals or cancer-specific conversational agents to provide tailored guidance and deeper patient connection [43,44]. Recent interventions using a range of web-based elements yielded a reduction in chemotherapy-related side effects and improved effectiveness in self-care behaviors in patients with cancer [43,44]. These findings also contribute to a broader shift toward risk-adapted, digitally enabled thyroid cancer care. Advances in molecular testing and AI-assisted cytopathology are refining the diagnostic evaluation of indeterminate nodules, while biomarker approaches are being examined to refine selection for RAI [45-47]. Together, these changes further underscore the potential value of scalable digital survivorship interventions, such as RAI Support, for addressing persistent symptoms and supporting patients throughout posttreatment care.

XSL•FO
RenderX

Limitations

There are several limitations of this study. Despite recruitment efforts through a publicly available thyroid cancer nonprofit Facebook page, most participants were recruited based on purposeful sampling from electronic medical record review. This led to participants in the study being relatively similar, with the majority being female, well educated, and from higher-income backgrounds (>US $200,000 per year). Although DTC does predominantly affect female individuals, having the preponderance of study participants come from such analogous backgrounds limits the generalizability of the study to larger and more diverse DTC populations. Usability and acceptability need to be confirmed in patients with broader socioeconomic, educational, and digital literacy backgrounds before implementation.

The small sample size of the study was appropriate to address the aims of the usability evaluation and ORBIT (Obesity-Related Behavioral Intervention Trials) model proof-of-concept testing, which focused on identifying preliminary effect size signals of change [20,48]. However, this limited sample size and the absence of a comparison condition further limit the generalizability of the findings and do not allow analyses of statistical differences or preliminary intervention effects. The study also covered a short time frame for participants, with user engagement and clinical outcomes tracked only over a 4-week period. This limited interval of analysis prevents a full understanding of the long-term benefits of the intervention for patients after RAI treatment.

An additional limitation is that the study encountered 29 imposter participants. Although those individuals were excluded, this recruitment issue is relevant to feasibility, recruitment, and generalizability of remote usability testing procedures and should be interpreted with caution [20,48]. The potential influence of outlier scores is another limiting factor, such as for PROMIS Global Physical Health, particularly given the small sample size. Future pilot feasibility trials should incorporate larger participant samples from different geographic areas, participants with varying intervals since RAI treatment, and a control condition to allow more rigorous evaluation of the RAI Support intervention.

Conclusions

This study demonstrated a positive usability testing evaluation of the web-based RAI Support prototype for RAI symptom management among survivors of DTC, a population that has received limited survivorship-focused attention. The study also identified preliminary effect size signals of change in self-efficacy in survivors of DTC, global mental health, and HRQOL. Ongoing research has incorporated the patient-recommended refinements and will examine the feasibility and acceptability of the RAI Support intervention through a pilot randomized controlled trial.

Acknowledgments

The authors thank Gary Bloom, Jana Schwartz, Kelsie Hitesman, Rachel Wong, Cynthia McDonald, Amanda Rhodes, Elyse Jackson, Tatiana Ferraro, Lauren Bello, and Joseph Voight for their contributions to the intervention content development. The authors also thank Robin A Littlejohn, for their contribution to the development of the semistructured moderator guide. The authors declare the use of generative AI (GenAI) in the research and writing process. According to the Generative AI Delegation Taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: proofreading, editing, and reformatting. The GenAI tool used was ChatGPT (GPT-5.6 Sol; OpenAI).The authors used ChatGPT 5.6 Sol to provide a higher-resolution version of Figure 1. Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This project was funded by the National Cancer Institute (K99CA277588). AMJ received research funding support for trainees provided by the Lombardi Comprehensive Cancer Center CCSG (P30CA051008). This work was supported in part through funding provided by the Lombardi Comprehensive Cancer Center (LCCC) Cancer Research Training Education Coordination (CRTEC) Office, awarded to JAL. This project has been supported with services from Federal funds (UL1TR000101 previously UL1RR031975) from the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, through the Clinical and Translational Science Awards Program (CTSA), a trademark of the US Department of Health and Human Services (DHHS), part of the Roadmap Initiative, “Re-Engineering the Clinical Research Enterprise.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical considerations related to the protection of participant confidentiality, given the small sample size in this study, and in accordance with the guidance of the Health Insurance Portability and Accountability Act (HIPAA) and the consent process but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Exit questionnaire summary after 4 weeks of at-home Radioactive Iodine Support (RAI Support) use.
[DOCX File , 16 KB - cancer_v12i1e84151_app1.docx]

Multimedia Appendix 2
Back-end log-in data from 4 weeks of at-home Radioactive Iodine Support RAI (RAI Support) intervention use.
[DOCX File , 378 KB - cancer_v12i1e84151_app2.docx]

References

1. Diamond-Rossi SA, Jonklaas J, Jensen RE, Kuo C, Stearns S, Esposito G, et al. Looking under the hood of "the cadillac of cancers:" radioactive iodine-related craniofacial side effects among patients with thyroid cancer. J Cancer Surviv 2020 Dec;14(6):847-857 [FREE Full text] [doi: 10.1007/s11764-020-00897-5] [Medline: 32506220]
2. Alexander K, Lee SY, Georgiades S, Constantinou C. The "not so good" thyroid cancer: a scoping review on risk factors associated with anxiety, depression and quality of life. J Med Life 2023 Mar;16(3):348-371 [FREE Full text] [doi: 10.25122/jml-2022-0204] [Medline: 37168306]
3. Randle RW, Bushman NM, Orne J, Balentine CJ, Wendt E, Saucke M, et al. Papillary thyroid cancer: the good and bad of the "good cancer". Thyroid 2017 Jul;27(7):902-907 [FREE Full text] [doi: 10.1089/thy.2016.0632] [Medline: 28510505]
4. Kitahara CM, Schneider AB. Epidemiology of thyroid cancer. Cancer Epidemiol Biomarkers Prev 2022 Jul 01;31(7):1284-1297. [doi: 10.1158/1055-9965.EPI-21-1440] [Medline: 35775227]
5. Patel J, Klopper J, Cottrill EE. Molecular diagnostics in the evaluation of thyroid nodules: current use and prospective opportunities. Front Endocrinol (Lausanne) 2023 Feb 24;14:1101410 [FREE Full text] [doi: 10.3389/fendo.2023.1101410] [Medline: 36909304]
6. Ringel MD, Sosa JA, Baloch Z, Bischoff L, Bloom G, Brent GA, et al. 2025 American Thyroid Association management guidelines for adult patients with differentiated thyroid cancer. Thyroid 2025 Aug;35(8):841-985. [doi: 10.1177/10507256251363120] [Medline: 40844370]
7. Pasqual E, Sosa JA, Chen Y, Schonfeld SJ, Berrington de González A, Kitahara CM. Trends in the management of localized papillary thyroid carcinoma in the United States (2000-2018). Thyroid 2022 Apr;32(4):397-410 [FREE Full text] [doi: 10.1089/thy.2021.0557] [Medline: 35078347]
8. Jonklaas J, Carr AL, Luta G, Yu C, Jensen RE, Reasner E, et al. SAlivary, LAcrimal NaSal (SALANS) measure to assess side effects following radioactive iodine treatment: development, psychometric properties, and factor structure. Qual Life Res 2024 Jul;33(7):2011-2023. [doi: 10.1007/s11136-024-03684-2] [Medline: 38769210]
9. Carr AL, Jenkins AM, Jonklaas J, Gabriel K, Miller KE, Graves KD. Patient and provider perspectives of a web-based intervention to support symptom management after radioactive iodine treatment for differentiated thyroid cancer: qualitative study. JMIR Form Res 2025 Mar 19;9:e60588 [FREE Full text] [doi: 10.2196/60588] [Medline: 40106313]
10. Kotwal A, Fingeret A, Knape A, Patel A, Bradford Bell E, Goldner W. Thyroid cancer survivorship: challenges and opportunities. Endocr Pract 2024 Nov;30(11):1097-1102. [doi: 10.1016/j.eprac.2024.08.003] [Medline: 39209023]
11. McCabe JK, Keller HR, Goldfarb M. Survivorship and quality of life in thyroid cancer survivors. Endocrinol Metab Clin North Am 2025 Sep;54(3):469-481. [doi: 10.1016/j.ecl.2025.03.015] [Medline: 40716900]
12. Radhakrishnan A, Reyes-Gastelum D, Gay B, Hawley ST, Hamilton AS, Ward KC, et al. Primary care provider involvement in thyroid cancer survivorship care. J Clin Endocrinol Metab 2020 Sep 01;105(9):e3300-e3306 [FREE Full text] [doi: 10.1210/clinem/dgaa437] [Medline: 32639557]
13. Das A, Rodriguez Garcia GP, Rubio EC, Acosta GJ, Bylund CL, Brito JP, et al. Unmet care needs of adult survivors of thyroid cancer: a systematic review. JAMA Otolaryngol Head Neck Surg 2025 Jul 01;151(7):714-724. [doi: 10.1001/jamaoto.2025.0855] [Medline: 40402541]
14. Penedo FJ, Oswald LB, Kronenfeld JP, Garcia SF, Cella D, Yanez B. The increasing value of eHealth in the delivery of patient-centred cancer care. Lancet Oncol 2020 May;21(5):e240-e251 [FREE Full text] [doi: 10.1016/S1470-2045(20)30021-8] [Medline: 32359500]
15. Montalescot L, Baussard L, Charbonnier E. Factors associated with digital intervention engagement and adherence in patients with cancer: systematic review. J Med Internet Res 2024 Dec 11;26:e52542 [FREE Full text] [doi: 10.2196/52542] [Medline: 39661976]
16. Aapro M, Bossi P, Dasari A, Fallowfield L, Gascón P, Geller M, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct;28(10):4589-4612 [FREE Full text] [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
17. Beatty L, Binnion C. A systematic review of predictors of, and reasons for, adherence to online psychological interventions. Int J Behav Med 2016 Dec;23(6):776-794. [doi: 10.1007/s12529-016-9556-9] [Medline: 26957109]
18. Jake-Schoffman DE, Silfee VJ, Waring ME, Boudreaux ED, Sadasivam RS, Mullen SP, et al. Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. JMIR Mhealth Uhealth 2017 Dec 18;5(12):e190 [FREE Full text] [doi: 10.2196/mhealth.8758] [Medline: 29254914]

XSL•FO
RenderX

19. Farao J, Malila B, Conrad N, Mutsvangwa T, Rangaka MX, Douglas TS. A user-centred design framework for mHealth. PLoS One 2020 Aug 19;15(8):e0237910 [FREE Full text] [doi: 10.1371/journal.pone.0237910] [Medline: 32813711]
20. Nielsen J. Estimating the number of subjects needed for a thinking aloud test. Int J Hum Comput Stud 1994 Sep;41(3):385-397. [doi: 10.1006/ijhc.1994.1065]
21. Buchholzer S, Thakachy Subha S, Tchérémissinoff L, Boselie F, Triponez F, Faure F, et al. The RAI-6 Questionnaire: a new screening questionnaire to monitor complications of radioiodine treatment. Front Surg 2021 Sep 1;8:641945 [FREE Full text] [doi: 10.3389/fsurg.2021.641945] [Medline: 34540883]
22. Medero K, Abdi H, Ford C, Gollust S. Detecting and preventing imposter participants: methods and recommendations for qualitative researchers. Qual Health Res (Forthcoming) 2025 Apr 18:10497323251333243 [FREE Full text] [doi: 10.1177/10497323251333243] [Medline: 40249633]
23. Sharma P, McPhail SM, Kularatna S, Senanayake S, Abell B. Navigating the challenges of imposter participants in online qualitative research: lessons learned from a paediatric health services study. BMC Health Serv Res 2024 Jun 12;24(1):724 [FREE Full text] [doi: 10.1186/s12913-024-11166-x] [Medline: 38867177]
24. Roehl JM, Harland DJ. Imposter participants: overcoming methodological challenges related to balancing participant privacy with data quality when using online recruitment and data collection. Qual Rep 2022 Nov 09;27(11):2469-2485. [doi: 10.46743/2160-3715/2022.5475]
25. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
26. Borsci S, Federici S, Lauriola M. On the dimensionality of the System Usability Scale: a test of alternative measurement models. Cogn Process 2009 Aug;10(3):193-197. [doi: 10.1007/s10339-009-0268-9] [Medline: 19565283]
27. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
28. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
29. DreamSoft4u. URL: http://www.dreamsoft4u.com/ [accessed 2025-09-03]
30. Yanez B, Pearman T, Lis CG, Beaumont JL, Cella D. The FACT-G7: a rapid version of the Functional Assessment of Cancer Therapy-General (FACT-G) for monitoring symptoms and concerns in oncology practice and research. Ann Oncol 2013 Apr;24(4):1073-1078 [FREE Full text] [doi: 10.1093/annonc/mds539] [Medline: 23136235]
31. Hays RD, Spritzer KL, Schalet BD, Cella D. PROMIS-29 v2.0 profile physical and mental health summary scores. Qual Life Res 2018 Jul;27(7):1885-1891 [FREE Full text] [doi: 10.1007/s11136-018-1842-3] [Medline: 29569016]
32. Gruber-Baldini AL, Velozo C, Romero S, Shulman LM. Validation of the PROMIS® measures of self-efficacy for managing chronic conditions. Qual Life Res 2017 Jul;26(7):1915-1924 [FREE Full text] [doi: 10.1007/s11136-017-1527-3] [Medline: 28239781]
33. Zhang J, Li L, Liu D, Zhu T, Gao Q, Chen H, et al. Development and validation of the Information Needs Questionnaire for Differentiated Thyroid Cancer patient with radioactive iodine therapy (INQ-DTC). Nurs Open 2022 May;9(3):1794-1804 [FREE Full text] [doi: 10.1002/nop2.925] [Medline: 34003590]
34. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen 2012 Feb;141(1):2-18. [doi: 10.1037/a0024338] [Medline: 21823805]
35. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge; 1988.
36. Krippendorff K. Content Analysis: An Introduction to Its Methodology. Thousand Oaks, CA: SAGE Publications; 2004.
37. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
38. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
39. Zhu M, Wang X, Mei L. Implementation of digital health interventions in thyroid cancer care: a scoping review. Int J Med Inform 2026 Mar 01;207:106221. [doi: 10.1016/j.ijmedinf.2025.106221] [Medline: 41386068]
40. Florenzano P, Guarda FJ, Jaimovich R, Droppelmann N, González H, Domínguez JM. Radioactive iodine administration is associated with persistent related symptoms in patients with differentiated thyroid cancer. Int J Endocrinol 2016;2016:2586512 [FREE Full text] [doi: 10.1155/2016/2586512] [Medline: 27867395]
41. Lee SL. Complications of radioactive iodine treatment of thyroid carcinoma. J Natl Compr Canc Netw 2010 Nov;8(11):1277-86; quiz 1287. [doi: 10.6004/jnccn.2010.0094] [Medline: 21081784]
42. Goswami S, Peipert BJ, Mongelli MN, Kurumety SK, Helenowski IB, Yount SE, et al. Clinical factors associated with worse quality-of-life scores in United States thyroid cancer survivors. Surgery 2019 Jul;166(1):69-74. [doi: 10.1016/j.surg.2019.01.034] [Medline: 30898373]
43. Peerbolte TF, van Diggelen RJ, van den Haak P, Geurts K, Evers LJ, Bloem BR, et al. Conversational agents supporting self-management in people with a chronic disease: systematic review. J Med Internet Res 2025 Aug 26;27:e72309 [FREE Full text] [doi: 10.2196/72309] [Medline: 40857094]

XSL•FO
RenderX

44. Fridriksdottir N, Gunnarsdottir S, Zoëga S, Ingadottir B, Hafsteinsdottir EJ. Effects of web-based interventions on cancer patients' symptoms: review of randomized trials. Support Care Cancer 2018 Feb;26(2):337-351. [doi: 10.1007/s00520-017-3882-6] [Medline: 28921391]
45. Dell'Aquila M, Fiorentino V, Martini M, Capodimonti S, Cenci T, Lombardi C, et al. How limited molecular testing can also offer diagnostic and prognostic evaluation of thyroid nodules processed with liquid-based cytology: role of TERT promoter and BRAF V600E mutation analysis. Cancer Cytopathol 2021 Oct;129(10):819-829 [FREE Full text] [doi: 10.1002/cncy.22454] [Medline: 34076961]
46. Fiorentino V, Pizzimenti C, Franchina M, Micali M, Russotto F, Pepe L, et al. The minefield of indeterminate thyroid nodules: could artificial intelligence be a suitable diagnostic tool? Diagn Histopathol 2023 Aug;29(8):396-401. [doi: 10.1016/j.mpdhp.2023.06.013]
47. Pizzimenti C, Fiorentino V, Ieni A, Rossi E, Germanà E, Giovanella L, et al. BRAF-AXL-PD-L1 signaling axis as a possible biological marker for RAI treatment in the thyroid cancer ATA intermediate risk category. Int J Mol Sci 2023 Jun 12;24(12):10024 [FREE Full text] [doi: 10.3390/ijms241210024] [Medline: 37373171]
48. Czajkowski SM, Powell LH, Adler N, Naar-King S, Reynolds KD, Hunter CM, et al. From ideas to efficacy: the ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol 2015 Oct;34(10):971-982 [FREE Full text] [doi: 10.1037/hea0000161] [Medline: 25642841]

Abbreviations

DTC: differentiated thyroid cancer
FACT-G7: Functional Assessment of Cancer Therapy-General-7 Item
HIPAA: Health Insurance Portability and Accountability Act
HRQOL: health-related quality of life
INQ-DTC: Informational Needs Questionnaire for Differentiated Thyroid Cancer
ORBIT: Obesity-Related Behavioral Intervention Trials
PROMIS: Patient-Reported Outcomes Measurement Information System
RAI Support: Radioactive Iodine Support
RAI: radioactive iodine
SALANS: salivary, lacrimal, and nasal
SUS: System Usability Scale
ThyCa: Thyroid Cancer Survivors' Association Inc

Edited by M Balcarras; submitted 15.Sep.2025; peer-reviewed by N Neuendorf, V Fiorentino, YW Lai; comments to author 24.Mar.2026; revised version received 18.Sep.2026; accepted 18.Sep.2026; published 13.Oct.2026.

Please cite as:
Carr A, Jenkins AM, Gómez V, Lorenzo JA, Moclair R, Pocchia J, Gandhi D, Miller K, Graves KD, Jonklaas J
Symptom Management Following Radioactive Iodine Treatment in Adults With Thyroid Cancer: Usability Evaluation and Proof-of-Concept Study
JMIR Cancer 2026;12:e84151
URL: https://cancer.jmir.org/2026/1/e84151
doi:10.2196/84151
PMID:

©Alaina Carr, Angela M Jenkins, Valeria Gómez, Janine Alexandria Lorenzo, Ryan Moclair, James Pocchia, Devansh Gandhi, Kristen Miller, Kristi D Graves, Jacqueline Jonklaas. Originally published in JMIR Cancer (https://cancer.jmir.org), 13.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinical Feasibility of Six Inflammatory Markers for Predicting the Mortality of Patients With Cancer: Longitudinal Study

Kangwei Wang[1], MM; Ce Shi[2], MSc; Dingchao Xia[3,4,5], MM; Ya Lin[6], MD

[1]Department of Cardiology, Affiliated Yueqing Hospital of Wenzhou Medical University, Wenzhou, China

[2]School of Digital Economics, Wenzhou Vocational College of Science and Technology, Wenzhou, China

[3]Department of Infectious Diseases, Wenzhou Central Hospital, Wenzhou, China

[4]Department of Infectious Diseases, Wenzhou Sixth People's Hospital, Wenzhou, China

[5]Wenzhou Key Laboratory of Diagnosis and Treatment of Emerging and Recurrent Infectious Diseases, Wenzhou, China

[6]MAFLD Research Center, Department of Hepatology, The First Affiliated Hospital of Wenzhou Medical University, Nanbaixiang Street, Ouhai District, Wenzhou, China

Corresponding Author:
Ya Lin, MD
MAFLD Research Center, Department of Hepatology, The First Affiliated Hospital of Wenzhou Medical University, Nanbaixiang Street, Ouhai District, Wenzhou, China

Abstract

Background: Emerging evidence indicates that inflammation plays a crucial role in cancer prognosis. Inflammatory response biomarkers are recognized as promising prognostic factors for mortality in patients with cancer.

Objective: This study aims to evaluate the prognostic significance of the systemic inflammatory response index (SIRI), systemic immune-inflammation index (SII), platelet-to-lymphocyte ratio (PLR), neutrophil-to-lymphocyte ratio (NLR), inflammatory prognostic index (IPI), and C-reactive protein-albumin-lymphocyte (CALLY) index.

Methods: Weighted Cox regression analyses, restricted cubic spline models, Kaplan-Meier survival curves, and receiver operating characteristic analyses were performed to assess the predictive value of the 6 inflammatory markers for mortality. Subgroup analyses and sensitivity analyses were conducted to examine associations within specific subpopulations.

Results: Cox regression models demonstrated that SIRI, NLR, IPI, and CALLY were significant predictors of all-cause mortality (tertile 3 vs tertile 1; hazard ratio [HR]: SIRI: 1.72, 95% CI 1.29 - 2.27; NLR: 1.33, 95% CI 1.02 - 1.74; IPI: 1.48, 95% CI 1.14 - 1.92; CALLY: 0.66, 95% CI 0.51 - 0.85). IPI (HR 1.91, 95% CI 1.11 - 3.27) and CALLY (HR 0.53, 95% CI 0.31 - 0.90) were significantly associated with cancer-specific mortality, whereas only SIRI was able to predict cardiovascular mortality (*P* value for trend=.04). Dose-response relationships were observed between the 6 inflammatory markers and mortality outcomes. Kaplan-Meier survival curves further illustrated significant differences between tertile groups (log-rank test, *P*<.001). The 6 inflammatory indices exhibited moderate predictive ability for all-cause mortality. IPI yielded the highest area under the curve (AUC) for cancer-specific mortality (AUC=0.6338), and SIRI was the most efficient predictor of cardiovascular mortality (AUC=0.687). No significant interactions were observed between the 6 inflammatory markers and most subgroup variables.

Conclusions: SIRI, NLR, IPI, and CALLY represent convenient and cost-effective prognostic tools for predicting mortality in patients with cancer. In contrast, SII and PLR may not be reliable prognostic biomarkers.

(JMIR Cancer 2026;12:e99410) doi:10.2196/99410

KEYWORDS

inflammatory markers; all-cause mortality; cancer mortality; cardiovascular mortality; prognosis

Introduction

Cancer remains a major societal, public health, and economic burden, ranking as the third leading cause of death worldwide [1-3]. Although advances in diagnosis and treatment have substantially reduced cancer-related mortality rates, the number of cancer survivors is projected to exceed 26 million by 2040 [4,5]. Nevertheless, cancer treatments—particularly anticancer agents—may induce cardiotoxicity and elevate the risk of cardiovascular disease (CVD) [6]. Moreover, cancer and CVD may share common pathophysiological pathways, given the overlapping risk factors, inflammatory processes, metabolic dysregulation, and neurohormonal activation associated with both conditions [7,8]. Consequently, there is an urgent clinical need for early detection and routine screening of all-cause, cancer-specific, and cardiovascular mortality among patients with cancer, using simple, cost-effective, and readily accessible tools.

XSL•FO
RenderX

Accumulating evidence has established inflammation as a common pathogenic driver of various diseases. Inflammation is a hallmark of tumorigenesis and plays a critical role in cancer initiation, progression, and metastasis [9-11]. Similarly, CVD is recognized as a chronic inflammatory condition, with inflammation being central to the onset and progression of atherosclerosis [11-13]. Inflammatory status can be assessed using various immune cell parameters and biochemical markers. Furthermore, systemic inflammatory responses contribute to decreased serum albumin levels. Wu et al [14] demonstrated that albumin levels were associated with cause-specific mortality in older adults [14]. In recent years, novel composite inflammatory markers—including the systemic inflammatory response index (SIRI), systemic immune-inflammation index (SII), platelet-to-lymphocyte ratio (PLR), neutrophil-to-lymphocyte ratio (NLR), inflammatory prognostic index (IPI), and C-reactive protein-albumin-lymphocyte (CALLY) index—which integrate complete blood count (CBC) with albumin and C-reactive protein (CRP)—have emerged as promising prognostic indicators for predicting disease outcomes [15-19].

The primary objective of this study was to evaluate the predictive value of SIRI, SII, PLR, NLR, IPI, and CALLY for all-cause, cancer-specific, and cardiovascular mortality in patients with cancer, and to further explore the potential use of these markers in the early identification of mortality risk.

Methods

Study Population and Exclusion Criteria

Data for this study were obtained from the National Health and Nutrition Examination Survey (NHANES) [20], a program designed to generate vital and health statistics for the US population. Given that CRP data were available only for 5 consecutive survey cycles (2001 - 2002, 2003 - 2004, 2005 - 2006, 2007 - 2008, and 2009 - 2010), we initially collected data from 52,195 participants across these cycles. Participants were excluded if they met any of the following criteria: (1) missing blood test results for albumin, CBC, or CRP; (2) absence of a cancer diagnosis or missing information on cancer history; (3) missing data on relevant covariates; or (4) pregnancy.

Evaluation of 6 Inflammatory Markers

For all participants, CBC was performed on blood specimens using a Beckman Coulter MAXM instrument at the Mobile Examination Centers. CRP levels were quantified using latex-enhanced nephelometry, and albumin concentrations were measured using the DcX800 method.

The inflammatory markers were calculated using the following formulas:

SIRI=neutrophil count × monocyte countlymphocyte count

SII=neutrophil count × platelet countlymphocyte count

PLR=platelet countlymphocyte count

NLR=neutrophil countlymphocyte count

IPI=CRP × NLRalbumin

CALLY=albumin × lymphocyte countCRP × 10

Outcome and Follow-Up

Cancer diagnosis was based on self-reported medical history. Participants were identified as having cancer if they responded affirmatively to the question: "Have you ever been told by a doctor or other health professional that you had cancer or a malignancy of any kind?" Mortality status for the NHANES follow-up population was obtained from the publicly available mortality files [21]. The primary outcome was all-cause mortality, while secondary outcomes included cancer-specific and cardiovascular mortality, coded according to the *International Classification of Diseases, 10th Revision* (*ICD-10*). Follow-up time was calculated in person-months from the date of the interview, with the observation period ending on December 31, 2019.

Assessment of Covariates

Demographic characteristics included gender, age, race, educational level, marital status, smoking status, alcohol consumption, and the poverty-to-income ratio. Race was categorized as non-Hispanic White, non-Hispanic Black, other Hispanic, Mexican American, or other races. Educational level was classified as: less than 9th grade, 9th to 11th grade (including 12th grade with no diploma), high school graduate or General Educational Development diploma, some college or associate degree, and college graduate or above. Marital status included the following categories: married, widowed, divorced, separated, never married, and living with a partner. Participants who consumed at least 12 drinks of any type of alcoholic beverage within 1 year were defined as drinkers. Those who reported smoking at least 100 cigarettes during their lifetime were categorized as smokers. Medical history—including hypertension, diabetes, hyperlipidemia, myocardial infarction (MI), congestive heart failure (CHF), coronary heart disease (CHD), and stroke—was based on prior medical records provided by health care professionals or physicians.

Statistical Analysis

The distribution of continuous variables was assessed using the Shapiro-Wilk test. Given the nonnormal distribution of the variables, they were summarized as medians with IQRs. Categorical data were expressed as frequencies and percentages. Differences among groups stratified by the 6 inflammatory markers were evaluated using the Kruskal-Wallis test or Fisher exact test, as appropriate.

To evaluate the prognostic value of the 6 inflammatory markers for mortality, weighted Cox regression analyses were performed to estimate hazard ratios (HRs) with 95% CIs. Three adjustment models were constructed: model I was unadjusted; model II was adjusted for gender, age, race, educational level, marital status, and poverty-to-income ratio; and model III was additionally adjusted for smoking, alcohol consumption, hypertension, diabetes, hyperlipidemia, CHF, CHD, MI, and stroke, in addition to the covariates included in model II. Restricted cubic spline (RCS) models were further used to explore potential dose-response relationships between the inflammatory markers and mortality outcomes, with adjustment for demographic characteristics, smoking, alcohol consumption,

XSL•FO
RenderX

and disease history. Kaplan-Meier survival curves with log-rank tests were used to estimate survival distributions across tertiles of the inflammatory markers. Receiver operating characteristic curves were generated to assess the sensitivity and specificity of the 6 inflammatory markers, and the area under the curve (AUC) was calculated to evaluate their predictive performance for all-cause, cancer-specific, and cardiovascular mortality. Subgroup analyses were conducted concurrently to elucidate potential interactions between the 6 inflammatory markers and covariates. To assess the robustness of the results, a sensitivity analysis was performed excluding patients with asthma and arthritis.

All statistical analyses were performed using R software (version 4.3.0; R Core Team 2023), and a *P* value of less than .05 was deemed statistically significant.

Ethical Considerations

The protocol for NHANES was approved by the National Center for Health Statistics Ethics Review Board (protocols: 98‐12 and 2005‐06 [22]) in accordance with the Declaration of Helsinki. All participants signed written informed consent.

Results

Baseline Characteristics

A total of 946 participants were enrolled in this study (Figure S1 in Multimedia Appendix 1). Based on all-cause mortality status, participants were classified into survival and nonsurvival groups. Table 1 summarizes the weighted baseline characteristics comparing the 2 groups. The survival group had a higher proportion of women, was younger, and exhibited higher educational levels, higher income, and more favorable marital status compared with the nonsurvival group. The nonsurvival group reported a higher prevalence of hypertension, diabetes, CHF, CHD, MI, stroke, and weak or failing kidneys. Furthermore, the values of SIRI, SII, PLR, NLR, and IPI were higher in the nonsurvival group than in the survival group, whereas the CALLY index was lower in the nonsurvival group (Table 1; Figure S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Survey-weighted baseline characteristics of individuals from the National Health and Nutrition Examination Survey.

Characteristics	Overall (weighted N=7,374,727.2)	Survival (weighted n=4,573,558.9)	Nonsurvival (weighted n=2,801,168.3)	*P* value
Gender, n (%)				<.001
Man	3,047,010.5 (41.3)	1,641,241.2 (35.9)	1,405,769.3 (50.2)	
Woman	4,327,716.7 (58.7)	2,932,317.7 (64.1)	1,395,399.0 (49.8)	
Age (y), median (IQR)	63.00 (52.00-74.00)	58.00 (48.00-67.00)	74.00 (65.00-80.00)	<.001
Race, n (%)				.12
Mexican American	132,737.2 (1.8)	99,944.5 (2.2)	32,792.8 (1.2)	
Other Hispanic	150,294.8 (2.0)	52,640.9 (1.2)	97,653.9 (3.5)	
Non-Hispanic White	6,633,329.9 (89.9)	4,147,593.8 (90.7)	2,485,736.1 (88.7)	
Non-Hispanic Black	328,374.8 (4.5)	166,095.2 (3.6)	162,279.5 (5.8)	
Other race, including multiracial	129,990.5 (1.8)	107,284.5 (2.3)	22,706.1 (0.8)	
Educational level, n (%)				<.001
Less than 9th grade	487,769.1 (6.6)	154,587.5 (3.4)	333,181.6 (11.9)	
9-11th grade (includes 12th grade with no diploma)	845,599.6 (11.5)	403,359.8 (8.8)	442,239.8 (15.8)	
High school graduate or general equivalent diploma	1,882,405.5 (25.5)	1,110,510.7 (24.3)	771,894.8 (27.6)	
Some college or associate's degree	2,026,301.5 (27.5)	1,304,406.2 (28.5)	721,895.3 (25.8)	
College graduate or above	2,132,651.5 (28.9)	1,600,694.8 (35.0)	531,956.7 (19.0)	
Marital status, n (%)				<.001
Married	4,831,440.5 (65.5)	3,250,681.2 (71.1)	1,580,759.3 (56.4)	
Widowed	999,693.0 (13.6)	324,868.2 (7.1)	674,824.7 (24.1)	
Divorced	853,298.7 (11.6)	567,372.8 (12.4)	285,925.9 (10.2)	
Separated	137,531.6 (1.9)	91,109.7 (2.0)	46,421.9 (1.7)	
Never married	348,198.2 (4.7)	220243.0 (4.8)	127,955.2 (4.6)	
Living with partner	204,565.3 (2.8)	119,284.0 (2.6)	85,281.3 (3.0)	
PIR[a], median (IQR)	3.32 (1.76-5.00)	4.03 (2.18-5.00)	2.24 (1.33-3.79)	<.001
Smoking, n (%)				.12
Yes	4,333,336.8 (58.8)	2,596,368.8 (56.8)	1,736,968.1 (62.0)	
No	3,041,390.4 (41.2)	1,977,190.2 (43.2)	1,064,200.2 (38.0)	
Drinking, n (%)				.06
Yes	5,063,639.4 (68.7)	3,260,109.2 (71.3)	1,803,530.3 (64.4)	
No	2,311,087.8 (31.3)	1,313,449.7 (28.7)	997,638.0 (35.6)	
BMI (kg/m^2), median (IQR)	27.47 (23.95-31.96)	27.49 (23.98-32.07)	27.24 (23.84-31.64)	.77
Waist circumference (cm), median (IQR)	99.53 (89.00-110.55)	98.30 (88.20-109.00)	101.12 (91.03-111.30)	.02
SBP[b] (mm Hg), median (IQR)	124.00 (112.67-136.67)	120.83 (110.00-131.33)	131.33 (117.33-146.00)	<.001
DBP[c] (mm Hg), median (IQR)	70.00 (62.09-76.67)	71.33 (65.33-78.00)	66.00 (58.00-74.67)	<.001
Complications and comorbidities, n (%)				
Hypertension				<.001

XSL•FO
RenderX

Characteristics	Overall (weighted N=7,374,727.2)	Survival (weighted n=4,573,558.9)	Nonsurvival (weighted n=2,801,168.3)	*P* value
Yes	3,592,871.3 (48.8)	1,846,042.6 (40.4)	1,746,828.7 (62.5)	
No	3,775,509.2 (51.2)	2,727,516.3 (59.6)	1,047,992.9 (37.5)	
Diabetes				<.001
Yes	1,013,308.4 (13.7)	419,496.7 (9.2)	593,811.8 (21.2)	
No	6,145,988.4 (83.3)	4,033,495.1 (88.2)	2,112,493.3 (75.4)	
Borderline	215,430.4 (2.9)	120,567.2 (2.6)	94,863.3 (3.4)	
Hyperlipidemia				.07
Yes	3,197,057.8 (49.1)	2,076,882.2 (52.3)	1,120,175.6 (44.1)	
No	3,316,009.8 (50.9)	1,895,939.6 (47.7)	1,420,070.1 (55.9)	
Myocardial infarction				<.001
Yes	628,163.1 (8.5)	206,479.8 (4.5)	421,683.3 (15.1)	
No	6,746,564.1 (91.5)	4,367,079.1 (95.5)	2,379,485.0 (84.9)	
Congestive heart failure				<.001
Yes	478,238.4 (6.5)	93,396.5 (2.0)	384,841.9 (13.9)	
No	6,871,854.8 (93.5)	4,480,162.4 (98.0)	2,391,692.4 (86.1)	
Coronary heart disease				<.001
Yes	613,276.9 (8.3)	210,189.8 (4.6)	403,087.2 (14.5)	
No	6,733,110.4 (91.7)	4,358,771.7 (95.4)	2,374,338.8 (85.5)	
Stroke				<.001
Yes	474,510.7 (6.4)	111,899.8 (2.5)	362,610.9 (12.9)	
No	6,891,332.2 (93.6)	4,452,774.8 (97.5)	2,438,557.4 (87.1)	
Weak or failing kidney				<.001
Yes	301,610.5 (4.1)	79,138.2 (1.7)	222,472.3 (7.9)	
No	7,060,022.9 (95.9)	4,481,326.8 (98.3)	2,578,696.0 (92.1)	
Laboratory parameters, median (IQR)				
WBC[d] (1000 cells/μL)	6.40 (5.30-7.90)	6.20 (5.10-7.80)	6.60 (5.50-8.10)	.01
Neutrophil (1000 cells/μL)	3.80 (3.00-4.90)	3.70 (2.90-4.70)	4.00 (3.20-5.15)	.001
Lymphocyte (1000 cells/μL)	1.70 (1.40-2.20)	1.80 (1.50-2.20)	1.70 (1.20-2.20)	.01
Monocyte (1000 cells/μL)	0.50 (0.40-0.70)	0.50 (0.40-0.60)	0.60 (0.49-0.70)	<.001
RBC[e] (1,000,000 cells/μL)	4.59 (4.26-4.94)	4.61 (4.32-4.95)	4.56 (4.15-4.89)	.03
Hemoglobin (g/dL)	14.30 (13.40-15.20)	14.30 (13.50-15.20)	14.20 (12.96-15.10)	.07
Platelet (1000 cells/μL)	244.00 (204.00-287.42)	243.00 (207.99-285.46)	245.00 (196.93-293.25)	.55
Albumin (g/dL)	4.20 (4.00-4.40)	4.20 (4.00-4.40)	4.10 (3.90-4.30)	<.001
CRP[f] (mg/dL)	0.23 (0.09-0.58)	0.18 (0.08-0.47)	0.31 (0.14-0.79)	<.001
SIRI[g]	1.12 (0.76-1.69)	1.01 (0.71-1.48)	1.33 (0.93-2.05)	<.001
SII[h]	528.62 (377.94-745.98)	506.92 (363.83-697.67)	582.50 (403.96-870.95)	<.001
PLR[i]	138.84 (109.59-178.65)	137.88 (109.08-171.82)	141.15 (110.00-189.04)	.10
NLR[j]	2.14 (1.64-3.00)	2.05 (1.59-2.69)	2.36 (1.71-3.44)	<.001
IPI[k]	0.11 (0.04-0.31)	0.09 (0.03-0.23)	0.19 (0.07-0.51)	<.001

XSL•FO
RenderX

Characteristics	Overall (weighted N=7,374,727.2)	Survival (weighted n=4,573,558.9)	Nonsurvival (weighted n=2,801,168.3)	*P* value
CALLY[l]	3.27 (1.22-7.79)	3.97 (1.69-9.83)	2.20 (0.81-5.25)	<.001
Follow-up period (mo), median (IQR)	136.00 (109.00-177.00)	158.00 (130.00-192.50)	89.00 (42.00-131.38)	<.001

[a]PIR: poverty-to-income ratio.

[b]SBP: systolic blood pressure.

[c]DBP: diastolic blood pressure.

[d]WBC: white blood cell.

[e]RBC: red blood cell.

[f]CRP: C-reactive protein.

[g]SIRI: systemic inflammation response index.

[h]SII: systemic immune-inflammation index.

[i]PLR: platelet-to-lymphocyte ratio.

[j]NLR: neutrophil-to-lymphocyte ratio.

[k]IPI: inflammatory prognosis index.

[l]CALLY: C-reactive protein-albumin-lymphocyte index.

Associations of 6 Inflammatory Markers With Study Outcomes

In Cox regression models, the levels of the 6 inflammatory markers were stratified by tertiles to evaluate their associations with all-cause, cancer-specific, and cardiovascular mortality (Table 2 and Tables S1 and S2 in Multimedia Appendix 1). A 1-SD increase in SIRI was associated with an increased risk of all-cause, cancer-specific, and cardiovascular mortality (all *P*<.001). Similar associations were observed for SII, NLR, and IPI, with elevations in these markers corresponding to higher mortality risks for all 3 outcomes. Associations of PLR and CALLY with all-cause and cardiovascular mortality were observed only in unadjusted or partially adjusted models, with no significant associations with cancer-specific mortality.

XSL•FO
RenderX

Table . Association between 6 inflammatory markers and all-cause mortality.

	Model I[a], HR[b] (95% CI)	*P* value	Model II[c], HR (95% CI)	*P* value	Model III[d], HR (95% CI)	*P* value
SIRI[e]						
Per-SD increase	1.40 (1.32-1.49)	<.001	1.28 (1.18-1.39)	<.001	1.32 (1.21-1.45)	<.001
Tertile 1 (≤0.887)	Reference	—[f]	Reference	—	Reference	—
Tertile 2 (>0.887, ≤1.52)	1.97 (1.53-2.54)	<.001	1.60 (1.24-2.06)	<.001	1.65 (1.25-2.17)	<.001
Tertile 3 (>1.52)	2.67 (2.09-3.41)	<.001	1.74 (1.35-2.25)	<.001	1.72 (1.29-2.27)	<.001
P value for trend		<.001		<.001		<.001
SII[g]						
Per-SD increase	1.34 (1.24-1.45)	<.001	1.29 (1.18-1.41)	<.001	1.32 (1.19-1.46)	<.001
Tertile 1 (≤429)	Reference	—	Reference	—	Reference	—
Tertile 2 (>429, ≤689)	0.99 (0.78-1.26)	.94	0.85 (0.67-1.09)	.19	0.79 (0.61-1.03)	.09
Tertile 3 (>689)	1.67 (1.34-2.09)	<.001	1.29 (1.02-1.63)	.03	1.16 (0.90-1.49)	.24
P value for trend		<.001		.02		.15
PLR[h]						
Per-SD increase	1.23 (1.13-1.35)	<.001	1.14 (1.04-1.26)	.005	1.15 (1.03-1.27)	.01
Tertile 1 (≤119)	Reference	—	Reference	—	Reference	—
Tertile 2 (>119, ≤167)	0.89 (0.71-1.13)	.36	0.80 (0.63-1.01)	.06	0.74 (0.57-0.96)	.02
Tertile 3 (>167)	1.36 (1.09-1.70)	.006	1.09 (0.87-1.36)	.48	1.04 (0.81-1.32)	.78
P value for trend		.005		.39		.64
NLR[i]						
Per-SD increase	1.46 (1.35-1.58)	<.001	1.27 (1.16-1.38)	<.001	1.29 (1.17-1.41)	<.001
Tertile 1 (≤1.8)	Reference	—	Reference	—	Reference	—
Tertile 2 (>1.8, ≤2.75)	1.25 (0.98-1.59)	.07	1.05 (0.81-1.35)	.71	1.05 (0.80-1.38)	.72
Tertile 3 (>2.75)	2.05 (1.63-2.59)	<.001	1.44 (1.13-1.83)	.003	1.33 (1.02-1.74)	.03
P value for trend		<.001		.002		.02
IPI[j]						
Per-SD increase	1.31 (1.23-1.40)	<.001	1.30 (1.22-1.38)	<.001	1.30 (1.21-1.40)	<.001
Tertile 1 (≤0.072)	Reference	—	Reference	—	Reference	—
Tertile 2 (>0.072, ≤0.237)	1.28 (1.00-1.63)	.046	1.11 (0.87-1.42)	.39	1.18 (0.91-1.54)	.21
Tertile 3 (>0.237)	1.92 (1.53-2.41)	<.001	1.75 (1.38-2.21)	<.001	1.48 (1.14-1.92)	.003
P value for trend		<.001		<.001		.003
CALLY[k]						
Per-SD increase	0.79 (0.68-0.92)	.003	0.87 (0.77-0.99)	.04	0.89 (0.78-1.01)	.06
Tertile 1 (≤1.68)	Reference	—	Reference	—	Reference	—
Tertile 2 (>1.68, ≤4.99)	0.72 (0.58-0.89)	.002	0.67 (0.54-0.83)	<.001	0.76 (0.60-0.97)	.03

XSL•FO
RenderX

	Model I[a], HR[b] (95% CI)	*P* value	Model II[c], HR (95% CI)	*P* value	Model III[d], HR (95% CI)	*P* value
Tertile 3 (>4.99)	0.55 (0.43-0.69)	<.001	0.59 (0.47-0.75)	<.001	0.66 (0.51-0.85)	.001
P value for trend		<.001		<.001		<.001

[a]Model I was not adjusted for any covariates.

[b]HR: hazard ratio.

[c]Model II was adjusted for gender, age, race, educational level, marital status, and poverty-to-income ratio.

[d]Model III was adjusted for gender, age, race, educational level, marital status, poverty-to-income ratio, smoking, drinking, hypertension, diabetes, hyperlipidemia, congestive heart failure, coronary heart disease, myocardial infarction, and stroke.

[e]SIRI: systemic inflammation response index.

[f]Not applicable.

[g]SII: systemic immune-inflammation index.

[h]PLR: platelet-to-lymphocyte ratio.

[i]NLR: neutrophil-to-lymphocyte ratio.

[j]IPI: inflammatory prognosis index.

[k]CALLY: C-reactive protein-albumin-lymphocyte index.

With regard to the primary outcome of all-cause mortality, the higher SIRI tertile was associated with an increased risk (tertile 3 vs tertile 1; hazard ratio [HR]: model I: 2.67, 95% CI 2.09 - 3.41; model II: 1.74, 95% CI 1.35 - 2.25; model III: 1.72, 95% CI 1.29 - 2.27). The HRs for NLR and IPI in the second and third tertiles were significantly higher than those in the first (reference) tertile, with *P* value for trend less than .05. CALLY was the only protective factor, with HR values progressively decreasing from tertile 2 to tertile 3 compared with tertile 1 (*P* value for trend<.001).

Regarding cancer-specific mortality, IPI and CALLY appeared to be the most sensitive indicators. A higher IPI level was associated with an increased risk of cancer-specific mortality (tertile 3 vs tertile 1; HR: model I: 2.74, 95% CI 1.74 - 4.33; model II: 2.55, 95% CI 1.59 - 4.09; model III: 1.91, 95% CI 1.11 - 3.27). In contrast, a higher CALLY level was associated with a decreased risk of cancer-specific mortality (tertile 3 vs tertile 1; HR, 95% CI: model I: 0.38, 0.24 - 0.60; model II: 0.41, 0.25 - 0.65; model III: 0.53, 0.31 - 0.90). The other 4 markers were not significantly associated with cancer-specific mortality.

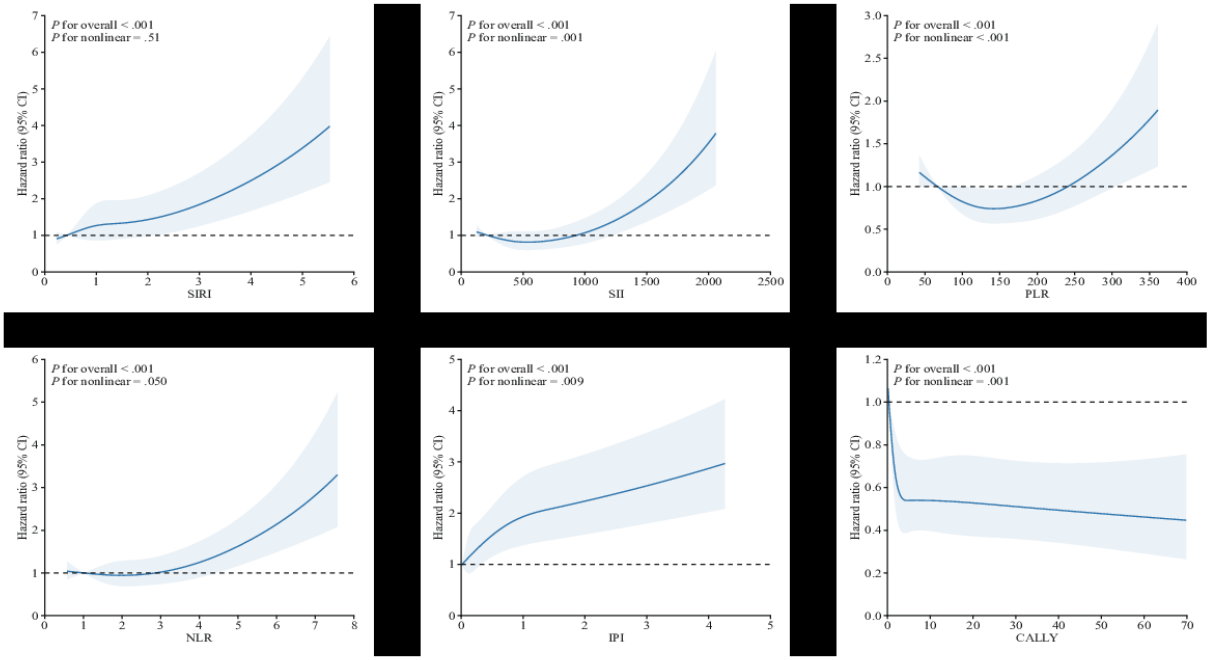
As shown in Table S2 in Multimedia Appendix 1, SIRI may be considered a predictor of cardiovascular mortality. Compared with tertile 1, the HR for tertile 3 was 3.50 (95% CI 1.51 - 3.66) in model I, 1.86 (1.06 - 3.27) in model II, and 1.83 (0.99 - 3.39) in model III, with *P* value for trend of less than .001, .02, and .04, respectively. The other markers were not significantly associated with cardiovascular mortality.

Dose-Response Relationship of 6 Inflammatory Markers With Study Outcomes

Multivariable-adjusted RCS analyses revealed significant dose-response relationships between the 6 inflammatory markers and all-cause mortality (Figure 1). SIRI (*P* value for nonlinear=.51) and NLR (*P* value for nonlinear=.05) presented linear correlations, while SII (*P* value for nonlinear=.001), PLR (*P* value for nonlinear<.001), IPI (*P* value for nonlinear=.009), and CALLY (*P* value for nonlinear=.001) exhibited nonlinear associations.

XSL•FO
RenderX

Figure 1. The restricted cubic spline (RCS) curve of the association between SIRI, SII, PLR, NLR, IPI, CALLY, and all-cause mortality. RCS regressions were adjusted for gender, age, race, educational level, marital status, PIR, smoking, drinking, hypertension, diabetes, hyperlipidemia, congestive heart failure, coronary heart disease, myocardial infarction, and stroke. CALLY: C-reactive protein-albumin-lymphocyte index; IPI: inflammatory prognosis index; NLR: neutrophil-to-lymphocyte ratio; PIR: poverty-to-income ratio; PLR: platelet-to-lymphocyte ratio; SII: systemic immune-inflammation index; SIRI: systemic inflammation response index.



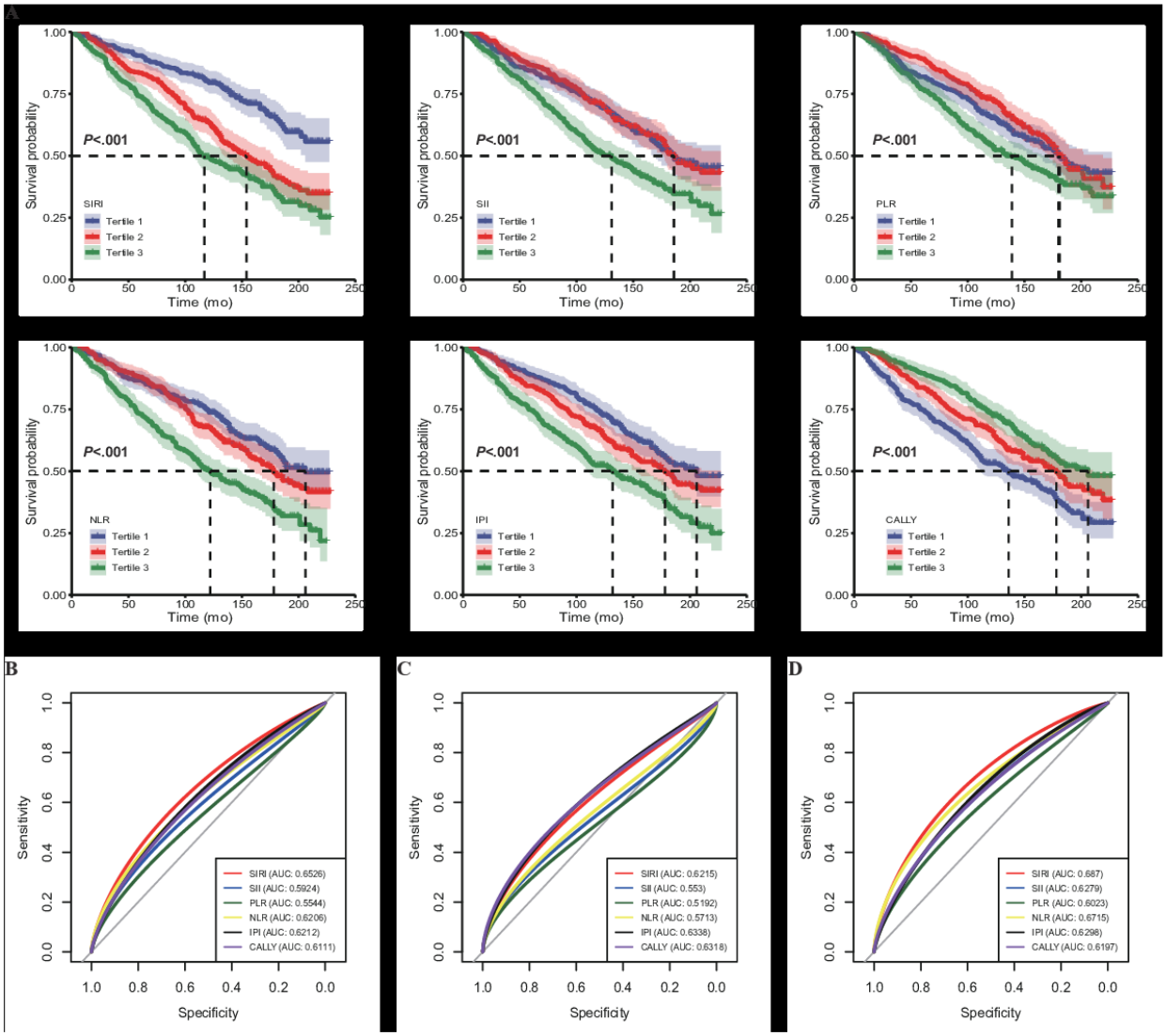
As shown in Figure S3 in Multimedia Appendix 1, SII (*P* value for nonlinear<.001), PLR (*P* value for nonlinear<.001), NLR (*P* value for nonlinear<.001), IPI (*P* value for nonlinear<.001), and CALLY (*P* value for nonlinear=.008) exhibited nonlinear dose-response relationships with cancer-specific mortality, whereas SIRI did not (*P* value for nonlinear=.16). Regarding cardiovascular mortality, SII (*P* value for nonlinear=.005) and CALLY (*P* value for nonlinear=.02) demonstrated nonlinear dose-response relationships, whereas SIRI, NLR, and IPI did not.

Six Inflammatory Markers as Predictors of the Clinical End Points

Survival curves for all-cause mortality stratified by tertiles of SIRI, SII, PLR, NLR, IPI, and CALLY are presented in Figure 2A (log-rank test; *P*<.001). Higher levels of SIRI, SII, PLR, NLR, and IPI were associated with worse survival probability, whereas CALLY exhibited the opposite trend.

XSL•FO
RenderX

Figure 2. (A) Kaplan-Meier survival curve for all-cause mortality. In the Kaplan-Meier curves, the population is stratified into tertile groups, and statistical analysis is conducted using the log-rank test. (B) Receiver operating characteristic (ROC) curve analysis of SIRI, SII, PLR, NLR, IPI, CALLY, and all-cause mortality. (C) ROC curve analysis of SIRI, SII, PLR, NLR, IPI, CALLY, and cancer mortality. (D) ROC curve analysis of SIRI, SII, PLR, NLR, IPI, CALLY, and cardiovascular mortality. AUC: area under the curve; CALLY: C-reactive protein-albumin-lymphocyte index; IPI: inflammatory prognosis index; NLR: neutrophil-to-lymphocyte ratio; PLR: platelet-to-lymphocyte ratio; SII: systemic immune-inflammation index; SIRI: systemic inflammation response index.



Notably, the survival trends for cancer-specific and cardiovascular mortality across the 6 inflammatory markers were similar to those observed for all-cause mortality. However, no significant difference in survival probability was observed among PLR tertile groups (Figure S4 in Multimedia Appendix 1).

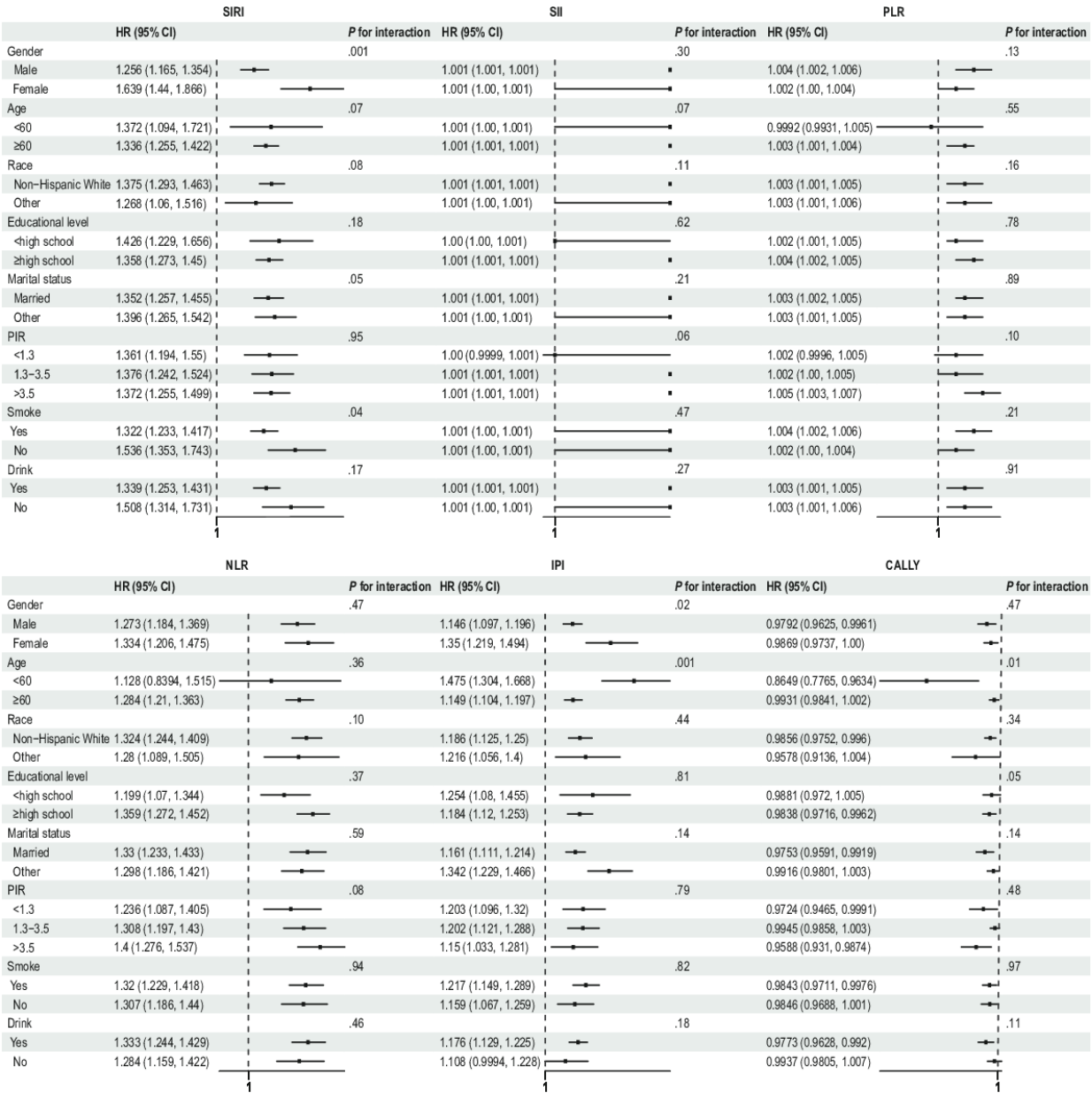
Figure 2B-D presents the predictive performance of the 6 inflammatory indices for all-cause, cancer-specific, and cardiovascular mortality. For all-cause mortality, the following AUC values were observed: SIRI: 0.6526, SII: 0.5924, PLR: 0.5544, NLR: 0.6206, IPI: 0.6212, and CALLY: 0.6111. Additionally, IPI yielded the highest AUC for cancer-specific mortality, while SIRI was the most efficient predictor of cardiovascular mortality.

Subgroup Analysis

No significant interactions were observed between SII, PLR, or NLR and all-cause mortality (all $P>.05$). However, gender, age, and smoking were found to potentially interact with SIRI, IPI, and CALLY (Figure 3). Regarding cancer-specific mortality, race, educational level, and marital status exhibited prominent interactions with SIRI, PLR, and IPI (Figure S5 in Multimedia Appendix 1). Furthermore, higher SIRI, SII, and NLR levels were more strongly associated with a higher prevalence of cardiovascular mortality in the married subgroup, whereas a higher cardiovascular mortality risk associated with IPI was observed in the other group (Figure S6 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 3. Subgroup analysis for the association between SIRI, SII, PLR, NLR, IPI, CALLY, and all-cause mortality. CALLY: C-reactive protein-albumin-lymphocyte index; HR: hazard ratio; IPI: inflammatory prognosis index; NLR: neutrophil-to-lymphocyte ratio; PIR: poverty-to-income ratio; PLR: platelet-to-lymphocyte ratio; SII: systemic immune-inflammation index; SIRI: systemic inflammation response index.



Sensitivity Analysis

After excluding patients with asthma, the risk of all-cause mortality showed a progressively increasing trend across tertiles of SIRI, NLR, and IPI, whereas CALLY showed a progressively decreasing trend (*P* value for trend<.01; Table S3 in Multimedia Appendix 1). After excluding patients with arthritis, the risk of all-cause mortality showed a progressively increasing trend across tertiles of SIRI and IPI, whereas CALLY showed a progressively decreasing trend (*P* value for trend<.05; Table S4 in Multimedia Appendix 1).

Discussion

To investigate the predictive role of inflammatory biomarkers in patients with cancer, we conducted a cross-sectional study involving 946 participants from the NHANES database. Our findings indicate that SIRI, SII, PLR, NLR, IPI, and CALLY were, to some extent, associated with mortality outcomes in patients with cancer, even after adjustment for selected confounders. Dose-response relationships were observed between most of these inflammatory indicators and mortality outcomes, with the exception of PLR in relation to cardiovascular mortality. Based on receiver operating characteristic curve analyses, the 6 inflammatory biomarkers demonstrated moderate predictive value for all-cause, cancer-specific, and cardiovascular mortality in patients with cancer. Furthermore, subgroup and sensitivity analyses validated the robustness of our models.

It is now well established that inflammation may contribute to the initiation and maintenance of unregulated cellular

XSL•FO
RenderX

proliferation, which can lead to tumor formation and growth [9]. Accumulating evidence supports the role of both local immune responses and systemic inflammation in tumor progression and patient survival [23,24]. Systemic inflammatory markers have been linked to increased cancer risk and mortality in numerous studies [25].

SIRI and SII are commonly used to assess the balance between inflammatory response and immune status. Previous studies have found that elevated levels of these markers are associated with poor prognosis in various cancers, including colorectal cancer [26,27], pancreatic cancer [28], breast cancer [29], chordoma [30], and lung cancer [27]. Our study demonstrated that SIRI and SII were elevated in patients who died of cancer, with SIRI showing strong prognostic value across all 3 mortality outcomes. However, the predictive power of SII was less robust in our analysis.

NLR and PLR are major leukocyte-based scores [25] and have been shown to be superior to traditional indicators in renal cell carcinoma [31], multiple sclerosis [32], and acute pancreatitis [33]. In predicting mortality in patients with cancer, NLR may be considered a significant predictor to some extent. Unfortunately, PLR did not demonstrate strong predictive capability in our study.

IPI and CALLY are novel composite indicators that evaluate prognosis through a comprehensive analysis of inflammatory levels, nutritional status, and immune function [17,18]. In our study, IPI was identified as a risk factor for mortality, whereas CALLY was negatively associated with mortality. These findings are consistent with previous cancer-related studies [17,18,34,35]. Both markers are highly promising indicators for predicting prognosis in patients with cancer and warrant further exploration and promotion in future clinical practice.

Collectively, these 6 inflammatory biomarkers—SIRI, SII, PLR, NLR, IPI, and CALLY—represent a simple and economical tool for evaluating systemic inflammation and predicting mortality across multiple cancer types. There is a genuine possibility that these biomarkers may have clinical use in the future. They may enable early stratification of patients for treatment and could potentially be used in clinical settings to determine the optimal timing for treatment modification.

Several limitations of this study should be acknowledged. First, as a single-center, retrospective study, selection bias was inevitable. Therefore, large-scale, multi-institutional, prospective studies are warranted to validate the effectiveness of the examined inflammatory biomarkers. Second, owing to data collection constraints, participants from other survey cycles were not included. In addition, several participants were excluded for failing to meet the inclusion criteria, which may have biased the results. Third, information on cancer history and other diseases was self-reported by participants, rendering it susceptible to recall bias and potential inaccuracies. Moreover, the inability to obtain detailed information on cancer type, stage, or treatment precluded stratified analyses by these variables. This limitation restricts our understanding of the predictive value of inflammatory markers across different malignant tumor types or disease severities. Fourth, this study could not accurately capture the long-term dynamic variations in CBC, CRP, and albumin. Inflammatory markers can vary considerably over time due to infections, treatments, or disease progression. Consequently, future longitudinal studies with repeated biomarker assessments are needed to confirm our findings. Finally, the set of covariates included in this study was incomplete, and some unmeasured confounding factors were not addressed. Specifically, data on the use of anti-inflammatory medications—which are known to lower inflammatory marker levels—were not available. The absence of this information likely introduces nondifferential misclassification of exposure, potentially leading to an underestimation of the true association. Future prospective studies are needed to adjust for these potential confounders.

In conclusion, our preliminary findings indicate that SIRI, NLR, IPI, and CALLY are economical, convenient, and readily available predictors of all-cause, cancer-specific, and cardiovascular mortality in patients with cancer. However, insufficient evidence was found to support the use of SII and PLR as reliable prognostic biomarkers. These inflammatory biomarkers appear to be potential supplementary tools for predicting mortality in patients with cancer. Future prospective trials with long-term follow-up are required to accurately determine the optimal timing and cutoff values that may be applied in clinical practice.

Acknowledgments

The authors acknowledge all the study staff and participants who participated in the National Health and Nutrition Examination Survey program. The authors used ChatGPT-5 to assist in the language revision of this manuscript.

Funding

This work was supported by the Research Start-up Funds for Postdoctoral Fellows at Wenzhou Medical University (2025BH011).

Data Availability

This study used data from the National Health and Nutrition Examination Survey [20]. Mortality and follow-up data were obtained from the National Death Index [21].

XSL•FO
RenderX

Authors' Contributions

KW and YL conceived and designed the study. CS and KW acquired the data. KW, CS, and YL drafted the manuscript. KW and DX performed the statistical analysis. YL supervised the study. All authors contributed important intellectual content to the manuscript and approved the final submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental figures and tables for inflammatory markers and mortality outcomes.
[DOCX File, 2161 KB - cancer_v12i1e99410_app1.docx]

References

1. Global Burden of Disease 2019 Cancer Collaboration, Kocarnik JM, Compton K, et al. Cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life years for 29 cancer groups from 2010 to 2019: a systematic analysis for the Global Burden of Disease Study 2019. JAMA Oncol 2022 Mar 1;8(3):420-444. [doi: 10.1001/jamaoncol.2021.6987] [Medline: 34967848]
2. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024;74(1):12-49. [doi: 10.3322/caac.21820] [Medline: 38230766]
3. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
4. Bluethmann SM, Mariotto AB, Rowland JH. Anticipating the “silver tsunami”: prevalence trajectories and comorbidity burden among older cancer survivors in the United States. Cancer Epidemiol Biomarkers Prev 2016 Jul;25(7):1029-1036. [doi: 10.1158/1055-9965.EPI-16-0133] [Medline: 27371756]
5. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
6. de Boer RA, Meijers WC, van der Meer P, van Veldhuisen DJ. Cancer and heart disease: associations and relations. Eur J Heart Fail 2019 Dec;21(12):1515-1525. [doi: 10.1002/ejhf.1539] [Medline: 31321851]
7. Makram OM, Okwuosa T, Addison D, et al. Cardiovascular diseases increase cancer mortality in adults: NHANES-continuous study. J Am Heart Assoc 2024 Aug 6;13(15):e035500. [doi: 10.1161/JAHA.124.035500] [Medline: 39056333]
8. Bertero E, Canepa M, Maack C, Ameri P. Linking heart failure to cancer: background evidence and research perspectives. Circulation 2018 Aug 14;138(7):735-742. [doi: 10.1161/CIRCULATIONAHA.118.033603] [Medline: 30359132]
9. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. Cell 2010 Mar 19;140(6):883-899. [doi: 10.1016/j.cell.2010.01.025] [Medline: 20303878]
10. Greten FR, Grivennikov SI. Inflammation and cancer: triggers, mechanisms, and consequences. Immunity 2019 Jul 16;51(1):27-41. [doi: 10.1016/j.immuni.2019.06.025] [Medline: 31315034]
11. Libby P, Kobold S. Inflammation: a common contributor to cancer, aging, and cardiovascular diseases-expanding the concept of cardio-oncology. Cardiovasc Res 2019 Apr 15;115(5):824-829. [doi: 10.1093/cvr/cvz058] [Medline: 30830168]
12. Frostegård J. Immunity, atherosclerosis and cardiovascular disease. BMC Med 2013 May 1;11:117. [doi: 10.1186/1741-7015-11-117] [Medline: 23635324]
13. Golia E, Limongelli G, Natale F, et al. Inflammation and cardiovascular disease: from pathogenesis to therapeutic target. Curr Atheroscler Rep 2014 Sep;16(9):435. [doi: 10.1007/s11883-014-0435-z] [Medline: 25037581]
14. Wu CY, Hu HY, Huang N, Chou YC, Li CP, Chou YJ. Albumin levels and cause-specific mortality in community-dwelling older adults. Prev Med 2018 Jul;112:145-151. [doi: 10.1016/j.ypmed.2018.04.015] [Medline: 29649489]
15. Wang RH, Wen WX, Jiang ZP, et al. The clinical value of neutrophil-to-lymphocyte ratio (NLR), systemic immune-inflammation index (SII), platelet-to-lymphocyte ratio (PLR) and systemic inflammation response index (SIRI) for predicting the occurrence and severity of pneumonia in patients with intracerebral hemorrhage. Front Immunol 2023;14:1115031. [doi: 10.3389/fimmu.2023.1115031] [Medline: 36860868]
16. Gu L, Xia Z, Qing B, et al. Systemic Inflammatory Response Index (SIRI) is associated with all-cause mortality and cardiovascular mortality in population with chronic kidney disease: evidence from NHANES (2001-2018). Front Immunol 2024;15:1338025. [doi: 10.3389/fimmu.2024.1338025] [Medline: 38558798]
17. Ma F, Li L, Xu L, et al. The relationship between systemic inflammation index, systemic immune-inflammatory index, and inflammatory prognostic index and 90-day outcomes in acute ischemic stroke patients treated with intravenous thrombolysis. J Neuroinflammation 2023 Sep 30;20(1):220. [doi: 10.1186/s12974-023-02890-y] [Medline: 37777768]

XSL•FO
RenderX

18. Yang M, Lin SQ, Liu XY, et al. Association between C-reactive protein-albumin-lymphocyte (CALLY) index and overall survival in patients with colorectal cancer: from the investigation on nutrition status and clinical outcome of common cancers study. Front Immunol 2023;14:1131496. [doi: 10.3389/fimmu.2023.1131496] [Medline: 37063910]
19. Bilen MA, Martini DJ, Liu Y, et al. The prognostic and predictive impact of inflammatory biomarkers in patients who have advanced-stage cancer treated with immunotherapy. Cancer 2019 Jan 1;125(1):127-134. [doi: 10.1002/cncr.31778] [Medline: 30329148]
20. NHANES questionnaires, datasets, and related documentation. Centers for Disease Control and Prevention (CDC) – National Center for Health Statistics. URL: https://wwwn.cdc.gov/nchs/nhanes/Default.aspx [accessed 2026-07-03]
21. National Death Index. Centers for Disease Control and Prevention (CDC) – National Center for Health Statistics. URL: https://www.cdc.gov/nchs/linked-data/mortality-files/?CDC_AAref_Val=https://www.cdc.gov/nchs/data-linkage/mortality-public.htm [accessed 2026-07-03]
22. Ethics review board approval. Centers for Disease Control and Prevention (CDC) – National Center for Health Statistics. URL: https://www.cdc.gov/nchs/nhanes/about/erb.html?CDC_AAref_Val=https://www.cdc.gov/nchs/nhanes/irba98.htm [accessed 2026-07-03]
23. Diakos CI, Charles KA, McMillan DC, Clarke SJ. Cancer-related inflammation and treatment effectiveness. Lancet Oncol 2014 Oct;15(11):e493-e503. [doi: 10.1016/S1470-2045(14)70263-3] [Medline: 25281468]
24. Dolan RD, Lim J, McSorley ST, Horgan PG, McMillan DC. The role of the systemic inflammatory response in predicting outcomes in patients with operable cancer: systematic review and meta-analysis. Sci Rep 2017 Dec 1;7(1):16717. [doi: 10.1038/s41598-017-16955-5] [Medline: 29196718]
25. Dupré A, Malik HZ. Inflammation and cancer: what a surgical oncologist should know. Eur J Surg Oncol 2018 May;44(5):566-570. [doi: 10.1016/j.ejso.2018.02.209] [Medline: 29530345]
26. Menyhart O, Fekete JT, Győrffy B. Inflammation and colorectal cancer: a meta-analysis of the prognostic significance of the systemic immune-inflammation index (SII) and the systemic inflammation response index (SIRI). Int J Mol Sci 2024 Aug 2;25(15):15. [doi: 10.3390/ijms25158441] [Medline: 39126008]
27. Nøst TH, Alcala K, Urbarova I, et al. Systemic inflammation markers and cancer incidence in the UK Biobank. Eur J Epidemiol 2021 Aug;36(8):841-848. [doi: 10.1007/s10654-021-00752-6] [Medline: 34036468]
28. Qi Q, Zhuang L, Shen Y, et al. A novel systemic inflammation response index (SIRI) for predicting the survival of patients with pancreatic cancer after chemotherapy. Cancer 2016 Jul 15;122(14):2158-2167. [doi: 10.1002/cncr.30057] [Medline: 27152949]
29. Zhang S, Cheng T. Prognostic and clinicopathological value of systemic inflammation response index (SIRI) in patients with breast cancer: a meta-analysis. Ann Med 2024 Dec;56(1):2337729. [doi: 10.1080/07853890.2024.2337729] [Medline: 38569199]
30. Hoppe M, Gersey ZC, Muthiah N, et al. The utility of inflammatory biomarkers in predicting overall survival and recurrence in skull base chordoma. Neurosurg Focus 2024 May;56(5):E16. [doi: 10.3171/2024.2.FOCUS2421] [Medline: 38691858]
31. Hu H, Yao X, Xie X, et al. Prognostic value of preoperative NLR, dNLR, PLR and CRP in surgical renal cell carcinoma patients. World J Urol 2017 Feb;35(2):261-270. [doi: 10.1007/s00345-016-1864-9] [Medline: 27255479]
32. Olsson A, Gustavsen S, Gisselø Lauridsen K, et al. Neutrophil-to-lymphocyte ratio and CRP as biomarkers in multiple sclerosis: a systematic review. Acta Neurol Scand 2021 Jun;143(6):577-586. [doi: 10.1111/ane.13401] [Medline: 33591593]
33. Karabuga B, Gemcioglu E, Konca Karabuga E, Baser S, Ersoy O. Comparison of the predictive values of CRP, CRP/albumin, RDW, neutrophil/lymphocyte, and platelet/lymphocyte levels in determining the severity of acute pancreatitis in patients with acute pancreatitis according to the BISAP score. Bratisl Lek Listy 2022;123(2):129-135. [doi: 10.4149/BLL_2022_020] [Medline: 35065589]
34. Feng J, Wang L, Yang X, Chen Q. Clinical significance of preoperative CALLY index for prognostication in patients with esophageal squamous cell carcinoma undergoing surgery. Sci Rep 2024;14(1):713. [doi: 10.1038/s41598-023-51109-w] [Medline: 38184747]
35. Zhu D, Lin YD, Yao YZ, Qi XJ, Qian K, Lin LZ. Negative association of C-reactive protein-albumin-lymphocyte index (CALLY index) with all-cause and cause-specific mortality in patients with cancer: results from NHANES 1999-2018. BMC Cancer 2024 Dec 5;24(1):1499. [doi: 10.1186/s12885-024-13261-y] [Medline: 39639229]

Abbreviations

AUC: area under the curve
CALLY: C-reactive protein-albumin-lymphocyte
CBC: complete blood count
CHD: coronary heart disease
CHF: congestive heart failure
CRP: C-reactive protein
HR: hazard ratio
***ICD-10*:** *International Classification of Diseases, 10th Revision*

XSL•FO
RenderX

IPI: inflammatory prognostic index
MI: myocardial infarction
NHANES: National Health and Nutrition Examination Survey
NLR: neutrophil-to-lymphocyte ratio
PLR: platelet-to-lymphocyte ratio
RCS: restricted cubic spline
SII: systemic immune-inflammation index
SIRI: systemic inflammatory response index

Edited by J Lau; submitted 24.Apr.2026; peer-reviewed by ES Bora, I Bilgetekin; revised version received 26.May.2026; accepted 28.May.2026; published 22.Jul.2026.

Please cite as:
Wang K, Shi C, Xia D, Lin Y
Clinical Feasibility of Six Inflammatory Markers for Predicting the Mortality of Patients With Cancer: Longitudinal Study
JMIR Cancer 2026;12:e99410
URL: https://cancer.jmir.org/2026/1/e99410
doi:10.2196/99410

© Kangwei Wang, Ce Shi, Dingchao Xia, Ya Lin. Originally published in JMIR Cancer (https://cancer.jmir.org), 22.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Immunotherapy Safety in Thymic Epithelial Tumors: Disproportionality Analysis of the Food and Drug Administration Adverse Event Reporting System

Ruilian Chen[1,2,3], MD; Hanrui Chen[1,3], PhD; Lingling Sun[1,3], PhD; Yang Cao[1,3], PhD; Lizhu Lin[1,3,4], PhD

[1]Oncology Center, The First Affiliated Hospital of Guangzhou University of Chinese Medicine, 16 Jichang Road, Baiyun District, Guangzhou, China

[2]The First Clinical Medical School of Guangzhou University of Chinese Medicine, Guangzhou, China

[3]Guangdong Clinical Research Academy of Chinese Medicine, Guangzhou, China

[4]Lingnan Medical Research Center, Guangzhou University of Chinese Medicine, Guangzhou, China

Corresponding Author:
Lizhu Lin, PhD
Oncology Center, The First Affiliated Hospital of Guangzhou University of Chinese Medicine, 16 Jichang Road, Baiyun District, Guangzhou, China

Abstract

Background: Immune checkpoint inhibitors (ICIs) have revolutionized cancer treatment, but their safety profile in patients with thymic epithelial tumors (TETs) remains poorly characterized due to the rarity of these malignancies.

Objective: This study aims to comprehensively analyze immune-related adverse events (irAEs) profiles in patients with TETs using real-world pharmacovigilance data.

Methods: We conducted a retrospective analysis of the US Food and Drug Administration Adverse Event Reporting System database from the first quarter of 2016 through the fourth quarter of 2024. Cases of TETs with ICI-related adverse events were identified and deduplicated following Food and Drug Administration recommendations. Disproportionality analysis was performed by calculating odds ratios, using the entire US Food and Drug Administration Adverse Event Reporting System database as the reference cohort. Signals were defined as significant with at least 3 cases and a lower 95% CI exceeded 1. Time-to-onset analysis and Weibull Shape Parameter testing were used to characterize the temporal pattern of irAEs. Clinical characteristics between fatal and nonfatal cases and cardiotoxicity specifics were analyzed descriptively.

Results: Among 152 eligible TET cases with irAEs, males slightly predominated (80/152, 52.6%), with a median age of 58.5 years. Reports originated predominantly from the United States (51/152, 33.6%) and Japan (16/152, 10.5%). PD-1 inhibitors were implicated in 66.4% (101/152) of cases. Disproportionality analysis identified 14 significant irAE signals across 11 System Organ Classes. Myositis (reporting odds ratio [ROR] 113.15, 95% CI 26.65 - 480.34), myocarditis (ROR 10.96, 95% CI 5.70 - 21.07), myasthenia gravis (ROR 3.86, 95% CI 1.86 - 7.85), and febrile neutropenia (ROR 18.33, 95% CI 4.57 - 73.55) demonstrated the strongest associations. The median time to irAEs onset was 21.0 days, with 73.2% (41/56) occurring within 2 months of treatment initiation. Fatal outcomes were reported in 23.7% (36/152) of cases and were significantly associated with gender distribution (P=.04) and different treatment strategies (P=.01). The utilization of PD-1 inhibitors was higher in the fatal group (29/36, 80.06%) than in the nonfatal group (72/116, 62.1%). Myocarditis was the most frequent cardiotoxicity, accounting for 51.3% (20/39) of cardiac events.

Conclusions: This large-scale pharmacovigilance study delineates a distinct and severe irAEs profile for ICIs in patients with TETs, characterized by robust disproportionality signals for myositis, cardiotoxicity, and myasthenia gravis. As a hypothesis-generating analysis, these findings underscore the need for vigilant monitoring and early detection strategies to mitigate irAE risks, particularly in high-risk subgroups such as patients with thymoma. The results provide clinically relevant evidence to guide risk-benefit evaluation and inform tailored surveillance protocols during ICI therapy in this population.

(JMIR Cancer 2026;12:e76908) doi:10.2196/76908

KEYWORDS

immunotherapy; immune-related adverse events; thymic epithelial tumors; myositis; myocarditis

XSL•FO
RenderX

Introduction

Thymic epithelial tumors (TETs), originated from thymic epithelial cells, are categorized as thymomas, thymic carcinomas, and thymic neuroendocrine tumors according to the World Health Organization [1]. TETs encompass a group of rare malignancies but constitute a significant proportion of anterior mediastinal tumors [2]. For early-stage TETs, curative surgical resection is the standard treatment strategy, yielding promising outcomes for a subset of patients. For those with advanced TETs, platinum-based chemotherapy has been considered as the first-line treatment [2,3]. However, there is no standard therapy for patients who experience disease progression after the first-line treatment of platinum-based regimens. Antiangiogenic drugs, mammalian target of rapamycin inhibitors, or other chemotherapy may be the potential option, although their effectiveness in this context remains uncertain [4,5].

Immunotherapy has gained great promise in the treatment landscape of various cancers, which sparked a significant shift in the paradigm of treating oncological malignancies [6,7]. The efficacy of immune checkpoint inhibitors (ICIs) was considered with a close relationship to the expression of programmed death-ligand 1 (PD-L1) [8,9]. Notably, TETs have been reported to exhibit high PD-L1 expression, with positivity rates ranging from 23% to 92% [10]. Encouraging clinical benefits have been observed with ICIs such as pembrolizumab, nivolumab, avelumab, and atezolizumab in patients with advanced thymic carcinoma or thymoma [11-14]. A phase II clinical trial demonstrated that the objective response rate (ORR) of pembrolizumab in 41 patients with thymic carcinoma was 23%, with a median progression-free survival (PFS) of 4.2 months and a median duration of response of approximately 3 years [11,12]. In the PRIMER study, the median PFS for nivolumab was 3.8 months, and the 12-month overall survival rate was 60.0% in a cohort of 15 patients with unresectable or recurrent thymic carcinoma [13]. In a study of 7 patients with thymoma treated with avelumab, an ORR of 57% was reported [14]. Despite these advancements, the unique immunologic environment of the thymus, which plays an essential role in T-cell maturation, raises concerns regarding immune-related adverse events (irAEs) [15].

Although the spectrum of irAEs associated with ICIs has been reported in various malignancies, the safety profile in TETs remains poorly defined [16]. This represents a critical knowledge gap, as the thymus is a primary lymphoid organ central to immune tolerance. The administration of ICIs in this immunologically unique context may disrupt central tolerance mechanisms, potentially leading to a distinct and more severe irAEs profile. Evidence from small-scale studies had shown a high frequency of severe toxicities in patients with TETs receiving ICI therapy [17,18]. Therefore, a TET-specific pharmacovigilance analysis is essential to systematically delineate these unique safety signals and inform clinical management. The US Food and Drug Administration Adverse Event Reporting System (FAERS) database provides real-world data on a global scale, while disproportionality analysis offers a validated method for detecting drug-event associations, especially for rare cancers such as TETs, where clinical trial data are limited. This approach effectively addresses evidence gaps by identifying clinically relevant toxicity signals and generating actionable hypotheses for risk management.

To gain a deeper understanding of the adverse events (AEs) associated with the therapy of ICIs in patients with TETs in clinical practice, we conducted a systematic analysis using the FAERS database to statistically evaluate irAEs in this patient population. This research endeavor aims to provide valuable insights into optimizing immunotherapy strategies for TETs while mitigating potential adverse effects.

Methods

Study Design and Data Source

This study was designed as a disproportionality analysis of individual case safety reports extracted from the FAERS database [19], covering the period from the first quarter of 2016 through the fourth quarter of 2024. The study was conducted and reported in accordance with the REporting of A Disproportionality analysis for drUg Safety signal detection using individual case safety reports (READUS) guidelines for studies using spontaneous reporting systems [20,21]. The primary analysis used disproportionality methods to detect significant associations between ICIs and irAEs in patients with TETs. Sensitivity analyses included stratification by ICI subclass and TETs histologic subtype. Additional methodological components comprised case-by-case clinical review of cardiotoxicity reports and contextualization of findings within the existing scientific literature. All AEs are coded using the preferred term (PT) according to the international MedDRA (Medical Dictionary for Regulatory Activities) (version 26.1). Detailed descriptions of the FAERS database and MedDRA structure are provided in Multimedia Appendix 1.

Identification of ICIs Reports With Adverse Events

Our comprehensive analysis was conducted in 4 steps to identify AEs linked to ICIs in the patients diagnosed with TETs. Initially, we performed a deduplication process according to the Food and Drug Administration's (FDA's) recommendations in the FAERS database. For reports sharing the same Case Identifier (CASEID), the latest FDA Date (FDA_DT) is selected, which contained the most comprehensive information. If the CASEID and FDA_DT are the same, we retained the higher Primary Identifier (PRIMARYID). In addition, an extra deduplication step was performed by excluding cases that had identical gender, age, reporting country, event date, AEs, and prescribed drugs. Second, we extracted data for patients diagnosed with TETs in the INDI files. Our focus was on ICIs reported in the DRUG files, including programmed death-1 (PD-1) inhibitors (nivolumab, pembrolizumab, sintilimab, penpulimab, tislelizumab, toripalimab, camrelizumab, and cemiplimab), PD-L1 inhibitors (atezolizumab, avelumab, and durvalumab), and cytotoxic T-lymphocyte–associated antigen 4 inhibitors (ipilimumab and tremelimumab). Reported drug names were mapped to their corresponding active pharmaceutical ingredient names using the FDA's National Drug Code directory and manual verification. For ICIs, both generic names and common brand names were unified under their respective active

XSL•FO
RenderX

pharmaceutical ingredient terms. Age, gender, and country fields were normalized to FAERS conventions. Only reports in which ICIs were classified as the primary suspect were included, as the AEs in these cases were deemed most likely to be associated with the ICIs compared with other medications. Third, we extracted all reports containing AEs from the REAC files. Finally, we identified ICIs reports diagnosed with TETs with AEs by intersecting the PRIMARYID.

Signal Mining

A disproportionality analysis is frequently conducted in pharmacovigilance research to identify safety signals and develop hypotheses concerning potential causal links between drugs and AEs. The reporting odds ratio (ROR) is used to compare the odds of reporting a specific event related to a particular drug against all other events while considering the reporting odds for other drugs within the FAERS database. The ROR was selected as the primary metric for this analysis due to its computational simplicity, transparency, and widespread adoption in the pharmacovigilance literature for initial signal detection. In this study, the ROR was used to identify signals of AEs in reports associated with ICIs in the patients experiencing TETs. By using AE reports from the complete FAERS database as a reference, we conducted a disproportionality analysis to evaluate the potential association between AEs and ICIs through the calculation of the ROR. The ROR and the 95% CI were calculated using the following formula:

$$ROR=a/cb/d$$

$$95\%CI=eln\ ROR\pm1.961a+1b+1c+1d$$

a: Number of reports with both the target ICIs and the target events; b: number of reports with the target ICIs but without the target events; c: number of reports with the target events but without the target ICIs; and d: number of reports with neither the target ICIs nor the target events.

Following established pharmacovigilance practice aimed at balancing sensitivity and specificity, we defined a significant signal as one with at least 3 case reports and a lower limit of the 95% CI for the ROR exceeding 1.

Data Analysis

A descriptive analysis of the clinical characteristics of reports with ICI-related AEs was conducted after screening. The outcome of death reported was included in the fatal group, while others were identified in the nonfatal group. The analysis of time to onset used cases that had a completed START_DT (the date when therapy commenced) recorded in the THER files along with EVENT_DT (the date on which AEs occurred) documented in the DEMO files. To appropriately account for right truncation in estimating the time to onset of AEs from spontaneous reporting databases, we defined an analysis period of 365 days following the commencement of drug administration. The statistical analysis of time to onset was performed with the Weibull Shape Parameter (WSP) test, which assesses the likelihood of changes in AE occurrence over time [22-24]. Specifically, when the shape parameter β is less than 1 and its 95% CI is also less than 1, this indicates a decreasing hazard of AEs over time. Conversely, if β is approximately equal to 1 with a 95% CI that includes 1, this suggests that the hazard remains constant. Furthermore, when β exceeds 1 along with a 95% CI greater than 1, this signifies an escalating hazard over time. Cumulative distribution curves were performed to illustrate the event-to-onset data for ICI-related AEs in different groups.

Statistical Analysis

Descriptive statistics were applied to summarize the demographic and administrative features of cases involving ICI-related AEs. Categorical variables are presented as frequencies and percentages, compared using chi-square or Fisher exact test. The human anatomy heatmap was conducted with the MOAHIT web tool [25]. Statistical analysis was conducted using R software (version 4.3.0; R Foundation for Statistical Computing) in this study. In this study, statistical significance was determined at the $P<.05$ level.

Ethical Considerations

This research uses the publicly available and anonymized FAERS database. As this database contains no personally identifiable information and the study does not involve experiments on human participants, it qualifies for exemption from institutional review board assessment and approval.

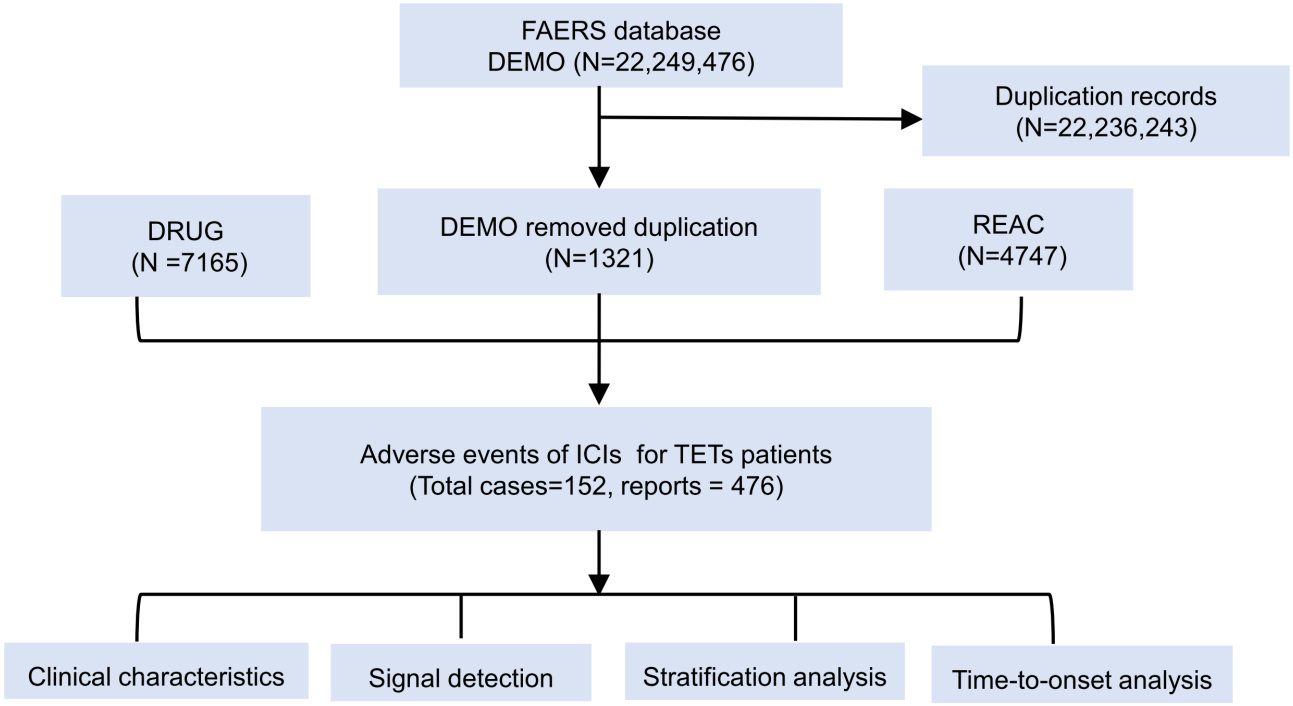
Results

ICI-Related AEs in the Patients With TETs in the FAERS, 2016 Q1 to 2024 Q4

We first investigated the amount of irAEs among patients diagnosed with TETs in the FAERS database. The comprehensive data processing procedure of this study is shown in Figure 1. From the first quarter of 2016 to the fourth quarter of 2024, the FAERS database cataloged a total of 22,249,476 AE reports. The number of ICI-related reports has experienced an increase since 2016 and the proportion of irAEs also varied among different ICI treatment strategies (Figure S1 in Multimedia Appendix 2).

XSL•FO
RenderX

Figure 1. Flowchart showing the analysis process of the study. A detailed description of the selection process of adverse events for ICIs in the FAERS. FAERS: US Food and Drug Administration Adverse Event Reporting System; ICIs: Immune checkpoint inhibitors; TETs: thymic epithelial tumors.



After excluding duplicate cases, 152 cases with TETs who experienced irAEs were identified. The clinical characteristics of patients are summarized in Table 1. Among the total study population with TETs, males accounted for 52.6% (80/152) of cases experiencing irAEs, compared with 40.8% (62/152) in females. Our results indicated that patients aged 18-64 years demonstrated a higher prevalence, accounting for 47.4% (72/152), with the median age of 58.5 years. The majority of cases were sourced from the United States (51/152, 33.6%), followed by Japan (16/152, 10.5%), France (14/152, 9.2%), and China (14/152, 9.2%), with the remaining reports (51/152, 37.5%) coming from 22 other countries. About 77.6% (118/152) of the cases were submitted by health care professionals, while 22.4% (34/152) were reported by consumers (Table 1). We found that 66.4% (101/152) of patients were associated with PD-1 inhibitor treatment strategies, 25.7% (38/152) with PD-L1 inhibitors, and 7.9% (12/152) in the group of the combination therapy of PD-1 inhibitors plus cytotoxic T-lymphocyte–associated antigen 4 inhibitors. Notably, a total of 85.5% (130/152) cases with irAEs were reported in patients receiving ICIs therapy, while 14.5% (22/152) were related to the use of immunotherapy in combination with chemotherapy or targeted agents.

XSL•FO
RenderX

Table . Characteristics of reports with immune checkpoint inhibitor–related adverse events sourced from the US Food and Drug Administration Adverse Event Reporting System database (January 1, 2011, to December 31, 2024).

Clinical characteristics	Fatal (n=36)	Nonfatal (n=116)	Total (N=152)	*P* value
Sex, n (%)				.038
Male	16 (44.4)	64 (55.2)	80 (52.6)	
Female	20 (55.6)	42 (36.2)	62 (40.8)	
Missing	0 (0.0)	10 (8.6)	10 (6.6)	
Age (years), n (%)				.135
18 - 64	18 (50.0)	54 (46.6)	72 (47.4)	
65 - 74	3 (8.3)	28 (24.1)	31 (20.4)	
≥75	3 (8.3)	6 (5.2)	9 (5.9)	
Missing	12 (33.3)	28 (24.1)	40 (26.3)	
Age (years), median	53	61	58.5	
Weight, n (%)				.100
<70	8 (22.2)	32 (27.6)	37 (25.2)	
≥70	3 (8.3)	25 (21.6)	28 (19.0)	
Missing	25 (69.4)	59 (51.8)	82 (55.8)	
Time to onset (days)				
Mean (SD)	57.0 (35.6)	73.2 (26.5)	71.2 (23.6)	.880
Median (minimum-maximum)	18.0 (8-266)	21.0 (2-1250)	21.0 (2-1250)	
Missing, n (%)	29 (80.1)	66 (56.9)	92 (62.6)	
Disease, n (%)				.489
Thymomas	15 (41.7)	46 (39.7)	61 (40.1)	
Thymic carcinomas	21 (58.3)	70 (60.3)	91 (59.9)	
Thymic neuroendocrine tumors	0	0	0	
Country, n (%)				.003
United States	12 (33.3)	39 (33.6)	51 (33.6)	
Japan	0 (0)	16 (13.8)	16 (10.5)	
China	1 (2.8)	13 (11.2)	14 (9.2)	
France	2 (5.6)	12 (10.3)	14 (9.2)	
Spain	1 (2.8)	6 (5.2)	7 (4.6)	
Korea, South	5 (13.9)	2 (1.7)	7 (4.6)	
Other country	15 (41.7)	28 (24.1)	43 (28.3)	
Treatment strategy, n (%)				.012
Anti-PD-1[a]	29 (80.6)	72 (62.1)	101 (66.4)	
Anti-PD-L1[b]	3 (8.3)	36 (31.0)	39 (25.7)	
Anti-CTLA4[c]	0 (0)	0 (0)	0 (0)	
Combination therapy	4 (11.1)	8 (6.9)	12 (7.9)	
Reported type, n (%)				.250
Consumer	10 (27.8)	24 (20.7)	34 (22.4)	

XSL•FO
RenderX

Clinical characteristics	Fatal (n=36)	Nonfatal (n=116)	Total (N=152)	*P* value
Health care professional	26 (72.2)	92 (79.3)	118 (77.6)	

[a]PD-1: programmed death-1.

[b]PD-L1: programmed death-ligand 1.

[c]CTLA4: cytotoxic T-lymphocyte–associated antigen 4.

Descriptive Analysis of Cases With ICI-Related AEs

We analyzed the categories and number of irAEs reported in patients with TETs treated with ICIs, and a total of 476 AEs were reported in 152 cases of TETs. To further investigate these irAEs, the disproportionality analysis was conducted using the full FAERS database as a reference by calculating ROR of PTs with at least 3 reported cases and the lower limit of the 95% CI for the ROR more than 1. About 14 specific PTs were considered as valid and strongly associated with ICI therapies, covering 11 System Organ Classes. Respiratory, thoracic, and mediastinal disorders were present in 9.03% (43/476) of reports, while cardiac disorders and musculoskeletal and connective tissue disorders were present in 8.19% (39/476) and 7.77% (37/476) of reports, respectively (Table S1 in Multimedia Appendix 2). In addition, myositis (24/476, 5.04%), myocarditis (20/476, 4.20%), myasthenia gravis (11/476, 2.31%), febrile neutropenia (9/476, 1.89%), and pneumonitis were among the top 20 of ICI-related AEs (Table 2). The frequent irAEs of TETs comprised myositis (n=24; ROR 113.15, 95% CI 26.65 - 480.34), myocarditis (n=20; ROR 10.96, 95% CI 5.70 - 21.07), myasthenia gravis (n=11; ROR 3.86, 95% CI 1.86 - 7.85), and febrile neutropenia (n=9; ROR 18.33, 95% CI 4.57 - 73.55) (Figure 2A). Among the ICI treatment strategies, anti-PD-1 accounts for the highest proportion at 66.4% (101/152), while anti-PD-L1 and combination therapies account for 25.7% (39/152) and 7.9% (12/152), respectively (Table S2 in Multimedia Appendix 2). The analysis revealed that different irAEs were associated with different ICIs treatment strategies (Figure 2B). The most common AEs of anti-PD-1 agents included myositis (n=18; ROR 56.68, 95% CI 19.08 - 168.42), myocarditis (n=17; ROR 6.37, 95% CI 6.37 - 25.38), and myasthenia gravis (n=9; ROR 4.23, 95% CI 1.97 - 9.09). The most common AEs of anti-PD-L1 agents included febrile neutropenia (n=9; ROR 15.09, 95% CI 6.99 - 32.57), rash maculopapular (n=4; ROR 29.12, 95% CI 8.38 - 101.24), and pneumonitis (n=3; ROR 10.79, 95% CI 3.05 - 38.18). The ICI-related AEs occurrence sites and the number of cases in patients with TETs are shown in the anatomical diagram of Figure 2C.

XSL•FO
RenderX

Table . The top 20 preferred terms of immune checkpoint inhibitor–related adverse events in patients with thymic epithelial tumors.

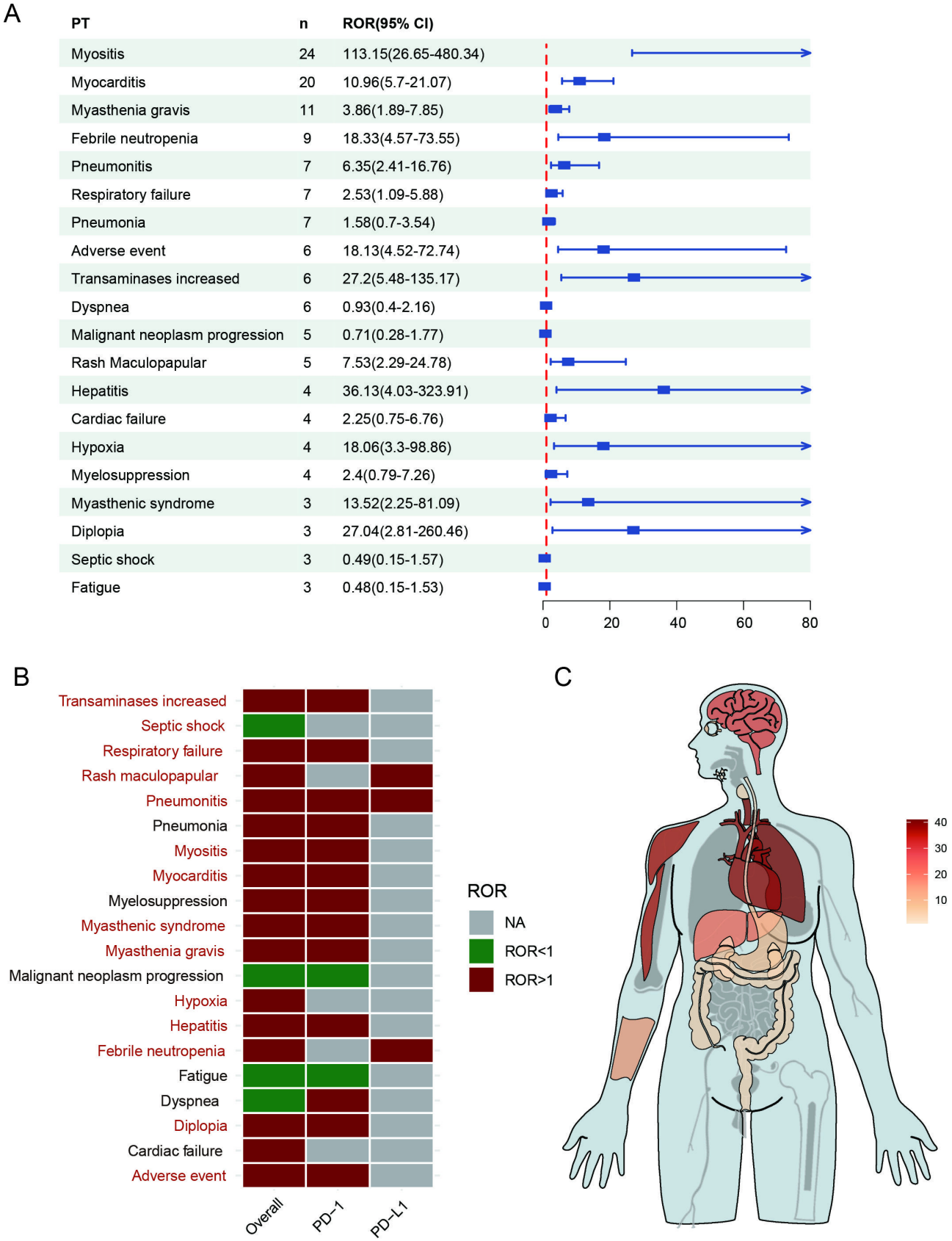
PT[a]	n (%)	SOC[b]
Myositis	24 (5.04)	Musculoskeletal and connective tissue disorders
Myocarditis	20 (4.20)	Cardiac disorders
Myasthenia gravis	11 (2.31)	Nervous system disorders
Febrile neutropenia	9 (1.89)	Blood and lymphatic system disorders
Pneumonitis	7 (1.47)	Respiratory, thoracic, and mediastinal disorders
Respiratory failure	7 (1.47)	Respiratory, thoracic, and mediastinal disorders
Pneumonia	7 (1.47)	Infections and infestations
Adverse event	6 (1.26)	General disorders and administration-site conditions
Transaminases increased	6 (1.26)	Investigations
Dyspnea	6 (1.26)	Respiratory, thoracic, and mediastinal disorders
Malignant neoplasm progression	5 (1.05)	Neoplasms benign, malignant, and unspecified (including cysts and polyps)
Rash maculopapular	5 (1.05)	Skin and subcutaneous tissue disorders
Hepatitis	4 (0.84)	Hepatobiliary disorders
Cardiac failure	4 (0.84)	Cardiac disorders
Hypoxia	4 (0.84)	Respiratory, thoracic, and mediastinal disorders
Myelosuppression	4 (0.84)	Blood and lymphatic system disorders
Myasthenic syndrome	3 (0.63)	Nervous system disorders
Diplopia	3 (0.63)	Eye disorders
Septic shock	3 (0.63)	Infections and infestations
Fatigue	3 (0.63)	General disorders and administration-site conditions

[a]PT: preferred term.

[b]SOC: System Organ Class.

XSL•FO
RenderX

Figure 2. Descriptive analysis of reports with immune-related adverse events (irAEs). (A) A forest plot shows the ROR of the top 20 irAEs. (B) The heatmap shows the ROR for top 20 irAEs (with cases no less than 3) in the US Food and Drug Administration Adverse Event Reporting System database under different immune checkpoint inhibitor treatment strategies (including overall situation, anti-PD-1, and anti-PD-L1). Red indicates ROR values greater than 1, green indicates ROR values less than 1, and gray indicates ROR values that could not be calculated. A significant association signal was labeled in dark red when the lower limit of the 95% CI for the ROR was greater than 1 and the number of cases was at least 3. (C) The anatomical diagram of irAEs occurrence site and the number of cases in patients with thymic epithelial tumors. The darker the color, the more cases with irAEs form this site. NA: not available; PD-1: programmed death-1; PD-L1: programmed death-ligand 1; PT: preferred term; ROR: reporting odds ratio.

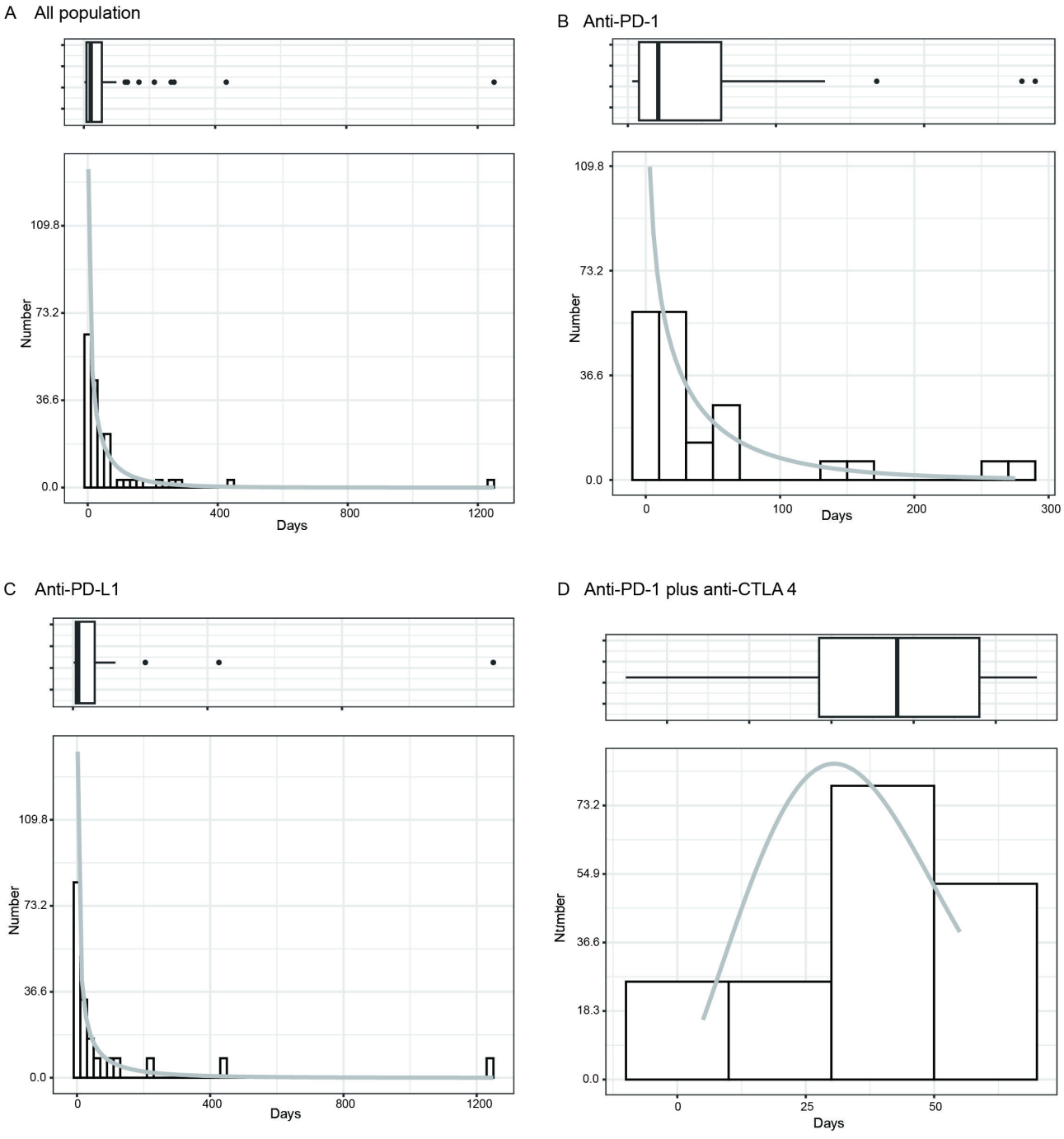


XSL•FO
RenderX

Time-to-Onset Analysis

Among the 152 ICI-related cases, 57 patients contained time-to-onset data. The cumulative distribution curves for the onset time of irAEs in different subgroups are present in Figure S2 in Multimedia Appendix 2. The median onset time of irAEs was significantly shorter in patients treated with ICIs plus chemotherapy than in those treated with ICIs alone (9.5 days vs 27.0 days; *P*=.005). No significant difference was present in the subgroup of age, gender, or the specific ICI treatment regimens on the onset time of irAEs in patients with TETs. We found that more than 70% of irAEs occurred within the first 2 months of initiating ICI therapies, with a median onset time of 21.0 days. Results of the WSP test for irAEs are summarized in Table S3 in Multimedia Appendix 2. The distribution of irAEs over the course of ICI treatment is shown in Figure 3. In the WSP test for irAEs, the upper limit of 95% CI of the shape parameter of other drugs was less than 1, indicating an early failure-type profile. Analyses of the time-to-onset profiles of irAEs indicated that early failure–type profiles were found in all populations with ICIs (Figure 3A) or subgroup of anti-PD-1 (Figure 3B) or anti-PD-L1 therapy (Figure 3C). The combination therapy group (anti-PD-1 plus anti-CTLA-4) exhibited a random failure pattern (Figure 3D).

Figure 3. Density plot of the onset time of cases with immune checkpoint inhibitor–related adverse events in different populations with thymic epithelial tumor. (A) Density plot of all populations. (B) Density plot of patients with anti-PD-1 treatments. (**C**) Density plot of patients with anti-PD-L1 treatments. (D) Density plot of patients with anti-PD-1 plus anti-CTLA4 treatments. CTLA4: cytotoxic T-lymphocyte–associated antigen 4; PD-1: programmed death-1; PD-L1: programmed death-ligand 1.



XSL•FO
RenderX

Comparison Between the Fatal and Nonfatal Groups

In cases of thymoma or thymic carcinoma with irAEs, about 23.7% (36/152) of patients experienced fatal outcomes. Although the presence of irAEs may not directly lead to death, analyzing the differences between the fatal and nonfatal groups remained crucial, as it may provide clues to reduce mortality rates. The gender distribution differed significantly between the fatal and nonfatal groups (P=.04). The proportion of female patients was higher in the fatal group (20/36, 55.6%) than in the nonfatal group (42/116, 36.2%). A significant difference was found in the application of ICIs treatment strategies between the fatal and nonfatal groups (P=.01). In total, 80.6% (29/36) of cases were treated with PD-1 inhibitors in the fatal group, while 62.1% (72/116) in the nonfatal group.

Cardiotoxicities Associated With ICIs in Patients With Thymomas and Thymic Carcinomas

A total of 29 patients experienced cardiac toxicities among the 152 cases of TETs with irAEs. More than 1 type of ICI-related cardiotoxicities was found in 7 patients. Among 39 cardiotoxicity-related AEs, myocarditis was the most frequent (20/39, 51.3%), followed by cardiac failure (3/39, 7.9%). ICI treatments showed a significant association with myocarditis (ROR 10.96, 95% CI 5.70-21.07), but no significant association was observed with cardiac failure (ROR 1.82, 95% CI 0.52-6.30). Among the 20 patients diagnosed with TETs who experienced immune-related myocarditis, 11 were female, representing a significant proportion of this subgroup. The median age of these patients was 58 years, and 95.0% (19/20) of patients received PD-1 inhibitor treatment. ICI-related myocarditis was more predominantly observed in patients with thymoma, accounting for 75% (15/20) of cases, and occurs in 25% (5/20) of those with thymic carcinoma (Table S4 in Multimedia Appendix 2). Among the 20 immune-related myocarditis cases, only 6 reports contained time-to-onset data, with the median time of 7 days.

Discussion

Principal Findings

This large-scale pharmacovigilance study, using the FAERS database from 2016 to 2024, systematically characterizes the safety profile of ICIs in the patients with TETs. Among 152 identified cases, we observed a distinct and severe spectrum of irAEs. Disproportionality analysis revealed strong signals for myositis (ROR 113.15, 95% CI 26.65 - 480.34), myocarditis (ROR 10.96, 95% CI 5.70 - 21.07), febrile neutropenia (ROR 18.33, 95% CI 4.57 - 73.55), and myasthenia gravis (ROR 3.86, 95% CI 1.86 - 7.85). The irAEs profile differed between ICI subclasses, with PD-1 inhibitors predominantly associated with myotoxic and cardiac events, while PD-L1 inhibitors showed stronger signals for febrile neutropenia and rash. Notably, 23.7% (36/152) of reported cases had fatal outcomes, and the median time to irAEs onset was 21 days, with more than 70% of events occurring within the first 2 months of therapy, significantly earlier when ICIs were combined with chemotherapy. These findings underscore a unique and concerning toxicity pattern for ICIs in TETs, dominated by early, severe neuromuscular, cardiac, and hematological complications. The findings will provide crucial insights for the clinical management of ICI therapy in patients with TETs, guiding safer and more effective treatment strategies.

Comparison With Prior Work

ICIs are now widely used in the treatment of various tumors [6,7]. In long-term follow-up of 40 patients with thymic carcinoma treated with pembrolizumab, an ORR of 22.5%, a median duration of response of 35.9 months, and a 5-year survival rate of 18% were observed [12]. In an open-label phase II trial, pembrolizumab showed encouraging antitumor activity in patients with advanced TETs, with the median PFS of 6 months [26]. Several clinical trials indicated that patients with TETs experienced varying degrees of clinical efficacy from the treatment with PD-1 and PD-L1 inhibitors [11,13,14]. As we all know, tumors escape the surveillance and elimination of the host immune system by activating immune checkpoints, which in turn suppress the cytotoxic killing effect of T cells. By blocking the interaction between immune checkpoints and their ligands, immune checkpoint drugs relieve the inhibition imposed on immune cells, thereby reactivating them to fight against cancer. However, the use of ICIs can also adversely affect normal cells, leading to the disruption of self-tolerance, inducing the development of autoimmune diseases in the body. In a systematic study of toxicities associated with cancer immunotherapy, about 81.15% (6959/8576) of patients had at least 1 toxicity event of varying degrees, and 29.27% (2510/8576) of cases had at least 1 toxicity event at grades 3 - 4 [27]. Hematologic toxicities, the most common types of grades 3 - 4 toxicities, were present in 9.41% (807/8576) of cases, followed by gastrointestinal toxicity in 5.34% (458/8576) of patients. For toxicities at any grade, patients experiencing respiratory toxicity were at a higher rate of treatment discontinuation than those with gastrointestinal toxicity or hematologic toxicity. However, compared with other cancers, there were certain differences in the terms of irAEs in patients with TETs. Although a meta-analysis reported that ICIs may bring clinical advantages for patients with thymomas, the advantages are limited by the excessively high risk of toxicity [28]. Grades 3 - 5 irAEs were reported in 26.4% for patients with TETs, with proportions of 17.1% for thymic carcinoma and 58.3% for thymoma [28]. The most common grade of ≥3 irAEs in TETs included myocarditis, hepatitis, myositis, and myasthenia gravis [12,26,28]. Consistent with reported studies, within our analysis of 476 AEs reports associated with ICIs, myositis and myocarditis were the common AEs, identified in 5.04% (24/476) and 4.20% (20/476) of the cases, respectively. Interestingly, there are differences between thymoma and thymic carcinoma in terms of irAEs in patients with TETs treated with ICIs [29]. Patients with thymoma tend to have a higher incidence of irAEs than those with thymic carcinoma [30]. In patients treated with ICIs, grade 3 or 4 irAEs were reported in 15.4% (4/26) of thymic carcinoma cases compared with 71.4% (5/7) of thymoma cases [26]. They suggested that ICIs should be avoided in patients with thymomas and should be considered with careful monitoring in those with thymic carcinoma. Patients with thymoma who receive ICIs are at considerable risk of experiencing severe immune-mediated toxicity, often presenting

XSL•FO
RenderX

with multiple irAEs simultaneously, which is not frequently seen in patients with other cancers receiving ICIs [31].

However, there are several researchers who have been dedicated to unraveling the mechanisms behind the distinct irAEs observed in TETs following the treatment of ICIs. The thymus, a crucial immune organ, is integral to the development and maturation of T cells. It is fundamental in maintaining immune tolerance by generating self-tolerant T cells through positive and negative selection processes [32]. In the context of tumorigenesis, the expression of autoimmune regulator and major histocompatibility complex in thymic epithelial cells declines, resulting in structural alterations within the thymus and impairing immune tolerance mechanisms [33]. This disruption can lead to an increase in immature CD8+ T cells and a decrease in CD4+ T cells and regulatory T cells, consequently promoting the development of autoimmune diseases [34,35]. TETs can promote the release of autoreactive T cells, disturbing peripheral homeostasis and predisposing patients to autoimmunity or causing autoimmune diseases. Some cytokines and chemokines produced by tumor cells induce cross-reactions between tumor antigens and tissue-specific antigens, promoting autoantibody production [29,36]. Furthermore, the administration of ICIs impacts central immune tolerance and disrupts the homeostasis of the peripheral T-cell receptor repertoire, leading to the activation of autoreactive T cells. This disruption can trigger autoimmune diseases in susceptible individuals [37].

ICI-induced myocarditis is a rare immune-related AE in various cancers, occurring in approximately 1% of patients undergoing treatment. However, it is noteworthy for its exceptionally high fatality rate, ranging between 30% and 50% [38-40]. There is limited knowledge on risk factors for ICI myocarditis, with only the combination therapy involving ICIs identified as a reliable risk factor of this condition [41,42]. In addition, the frequency of ICI myotoxicity has been reported 10 - to 30-fold higher in patients with TETs than in those with other cancers [43]. The seriousness of several life-threatening ICI-related myotoxicity was demonstrated in young patients with thymoma, who experienced a fatal triad of fulminant myocarditis, myositis, and a myasthenia gravis–like syndrome, emerging within a few days following a single dose of ICI treatment [44]. In line with our findings, the median time to onset of myocarditis was 7 days after the administration of immunotherapy. ICI-related myocarditis in patients with TETs tends to occur earlier and more severely after the initiation of ICI treatment than other cancers, resulting in life-threatening arrhythmias. Our findings revealed that the ICI-related myocarditis in patients with TETs was observed more in females and those with the treatment of anti-PD-1 inhibitors or the presence of thymomas. Emerging evidence indicated that autoreactive T cells recognizing the cardiac antigen α-actinin played a crucial role in the development of ICI-related myocarditis, unveiling the key mechanisms behind ICI-related myocarditis [45]. Increased frequencies and spatial colocalization of cytotoxic T lymphocytes, conventional dendritic cells, and inflammatory fibroblasts have been shown in ICI-related myocarditis heart tissue. In addition, the presence of heart-expanded T-cell receptors in a cycling blood CD8+ T-cell population may indicate whether a patient's myocarditis could potentially lead to death [46]. For patients with TETs who experienced myocarditis, research indicated that thymus-related dysfunction was linked to the incidence and severity of ICI-induced myotoxicities, including myocarditis and myositis, and identifying TETs and assessing acetylcholine receptor through radiobiological workup may help predict and stratify the risk of severe ICI myotoxicities in patients with cancer undergoing ICI treatment [43].

Strengths and Limitations

We conducted a relatively large-scale, disease-specific pharmacovigilance analysis of irAEs in patients with TETs in the real-world setting. This study addressed a significant knowledge gap in the field, as TETs originate from the central immune organ, thereby establishing a distinctive and high-risk immunological context for ICIs toxicity that is often obscured in broader, pan-cancer analyses. The findings provide essential evidence to inform clinical guidelines for ICIs use in TETs, representing a significant contribution to a niche yet high-stakes area of oncology. By systematically delineating the unique toxicity profile, our work has the potential to impact both clinical decision-making and design of future studies in TETs.

There were several limitations in this research. First, the data were sourced from a spontaneously reported database, which is subject to known biases such as underreporting, duplicate reports, and confounding factors, potentially affecting data quality and completeness [47]. Moreover, reports in the FAERS database predominantly originate from North America and Europe, resulting in a relatively low representation of Asian populations. Second, the use of ROR, although valuable for signal detection, may inflate associations for established class-specific AEs, such as myocarditis, myositis, and pneumonitis in the context of ICIs, due to heightened clinical awareness and targeted reporting. Consequently, ROR-based signals for known irAEs should be interpreted cautiously, as they largely reflect anticipated class effects rather than novel risks. Furthermore, the signal strength reflects relative risk and does not allow for the quantification of absolute risk. Third, the comparative analysis between anti-PD-1 and anti-PD-L1 agents should be interpreted with caution. While the differential toxicity signals provide hypothesis-generating insights, the clinical application of PD-L1 inhibitors in TETs remains limited. Consequently, the smaller sample size for PD-L1 cases in this study may reduce the statistical robustness of direct comparisons and limit the generalizability of subclass-specific findings. Finally, it is of critical clinical importance to understand the factors that determine who may develop irAEs. Due to the limited number of cases with time-to-onset data, the study could not conduct univariate or multivariate regression analyses for irAEs, which represents an important gap for clinical risk stratification. In addition, these findings are hypothesis-generating and require validation through well-designed epidemiological studies to clarify incidence rates, underlying mechanisms, and real-world clinical relevance of the identified signals.

Clinical Implications and Future Directions

The findings from this pharmacovigilance analysis carry significant clinical implications for the management of patients

XSL•FO
RenderX

with TETs receiving ICI therapy. The robust and early-onset signals for severe toxicities, particularly myositis and myocarditis, necessitate a paradigm shift in monitoring protocols. We recommend implementing proactive, protocol-driven monitoring within the first 2 months of treatment initiation for this high-risk population. Furthermore, the high fatality rate associated with cardiotoxicity, especially in specific subgroups, underscores the critical need for multidisciplinary management involving cardio-oncology specialists.

Future research should be directed toward translating these hypothesis-generating signals into actionable clinical knowledge. First, prospective cohort studies are needed to establish accurate incidence rates and validate risk stratification models incorporating clinical and molecular biomarkers. Second, mechanistic investigations should explore the unique immunobiology of TETs that predisposes patients to these severe toxicities. Third, there is an urgent need to develop TETs-specific clinical guidelines for irAE prevention, monitoring, and management, which should include standardized baseline assessment protocols and treatment algorithms tailored to this population's distinct risk profile.

Conclusions

This large-scale pharmacovigilance analysis, using real-world data from the FAERS database, delineates irAEs profile associated with ICIs therapy in patients with TETs. The study identified strong disproportionate reporting signals for myositis, myocarditis, myasthenia gravis, and febrile neutropenia, with more than 70% of irAEs occurring within the first 2 months of treatment initiation, with a median onset time of 21.0 days. As a hypothesis-generating investigation, these findings emphasize the need for vigilant early monitoring and support tailored surveillance strategies in clinical practice, particularly in patients with thymoma. Further research is warranted to elucidate underlying mechanisms, identify predictive biomarkers, and develop risk mitigation approaches to optimize the safety and efficacy of immunotherapy in this population.

Acknowledgments

The authors express their sincere gratitude to all colleagues and participants involved in this research, whose expertise and collaborative spirit were essential to its success.No artificial intelligence was used in this work.

Funding

This work was supported by the Guangdong Province General Universities Young Innovative Talents Program (2024KQNCX071) and the Guangzhou Municipal Science and Technology Bureau Basic Research Program (2023A03J0300). The funding source had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

The data used in this study are available on the US Food and Drug Administration's Adverse Events Reporting System (FAERS) from the website [19].

Authors' Contributions

RC participated in conceptualization, funding acquisition, writing—original draft, investigation, project administration, software, supervision, resources, methodology, and writing—review and editing. HC contributed to methodology, writing—original draft, and writing—review and editing. LS contributed to formal analysis, validation, writing—original draft, visualization, and writing—review and editing. YC contributed to formal analysis, validation, data curation, and writing—review and editing. LL participated in conceptualization, funding acquisition, formal analysis, writing—original draft, and writing—review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details on the US Food and Drug Administration Adverse Event Reporting System database and the Medical Dictionary for Regulatory Activities terminology used in the analysis.
[PDF File, 107 KB - cancer_v12i1e76908_app1.pdf]

Multimedia Appendix 2
Supplementary figures and tables supporting the analysis of immune checkpoint inhibitor–related adverse events in patients with thymic epithelial tumors.
[PDF File, 443 KB - cancer_v12i1e76908_app2.pdf]

References

XSL•FO
RenderX

1. Marx A, Chan JKC, Chalabreysse L, et al. The 2021 WHO classification of tumors of the thymus and mediastinum: what is new in thymic epithelial, germ cell, and mesenchymal tumors? J Thorac Oncol 2022 Feb;17(2):200-213. [doi: 10.1016/j.jtho.2021.10.010]
2. Remon J, Bernabé R, Diz P, et al. SEOM-GECP-GETTHI Clinical Guidelines for the treatment of patients with thymic epithelial tumours (2021). Clin Transl Oncol 2022 Apr;24(4):635-645. [doi: 10.1007/s12094-022-02788-w]
3. Multidisciplinary Committee of Oncology, Chinese Physicians Association. Chinese guideline for clinical diagnosis and treatment of thymic epithelial tumors (2021 Edition). Zhonghua Zhong Liu Za Zhi 2021 Apr 23;43(4):395-404. [doi: 10.3760/cma.j.cn112152-20210313-00226] [Medline: 33902202]
4. Conforti F, Pala L, Giaccone G, De Pas T. Thymic epithelial tumors: from biology to treatment. Cancer Treat Rev 2020 Jun;86:102014. [doi: 10.1016/j.ctrv.2020.102014]
5. Girard N, Ruffini E, Marx A, Faivre-Finn C, Peters S. Thymic epithelial tumours: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2015 Sep;26:v40-v55. [doi: 10.1093/annonc/mdv277]
6. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012 Mar 22;12(4):252-264. [doi: 10.1038/nrc3239] [Medline: 22437870]
7. Sanmamed MF, Chen L. A paradigm shift in cancer immunotherapy: from enhancement to normalization. Cell 2018 Oct 4;175(2):313-326. [doi: 10.1016/j.cell.2018.09.035] [Medline: 30290139]
8. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. Lancet Oncol 2016 Dec;17(12):e542-e551. [doi: 10.1016/S1470-2045(16)30406-5] [Medline: 27924752]
9. Zhang Y, Zhang Z. The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. Cell Mol Immunol 2020 Aug;17(8):807-821. [doi: 10.1038/s41423-020-0488-6]
10. Xu C, Zhang Y, Wang W, et al. Chinese expert consensus on the diagnosis and treatment of thymic epithelial tumors. Thorac Cancer 2023 Apr;14(12):1102-1117. [doi: 10.1111/1759-7714.14847] [Medline: 36924056]
11. Giaccone G, Kim C. Durable response in patients with thymic carcinoma treated with pembrolizumab after prolonged follow-up. J Thorac Oncol 2021 Mar;16(3):483-485. [doi: 10.1016/j.jtho.2020.11.003] [Medline: 33248322]
12. Giaccone G, Kim C, Thompson J, et al. Pembrolizumab in patients with thymic carcinoma: a single-arm, single-centre, phase 2 study. Lancet Oncol 2018 Mar;19(3):347-355. [doi: 10.1016/S1470-2045(18)30062-7] [Medline: 29395863]
13. Katsuya Y, Horinouchi H, Seto T, et al. Single-arm, multicentre, phase II trial of nivolumab for unresectable or recurrent thymic carcinoma: PRIMER study. Eur J Cancer 2019 May;113:78-86. [doi: 10.1016/j.ejca.2019.03.012] [Medline: 30991261]
14. Rajan A, Heery CR, Thomas A, et al. Efficacy and tolerability of anti-programmed death-ligand 1 (PD-L1) antibody (Avelumab) treatment in advanced thymoma. J Immunother Cancer 2019 Oct 21;7(1):269. [doi: 10.1186/s40425-019-0723-9] [Medline: 31639039]
15. Thapa P, Farber DL. The role of the thymus in the immune response. Thorac Surg Clin 2019 May;29(2):123-131. [doi: 10.1016/j.thorsurg.2018.12.001] [Medline: 30927993]
16. Jayathilaka B, Mian F, Franchini F, Au-Yeung G, IJzerman M. Cancer and treatment specific incidence rates of immune-related adverse events induced by immune checkpoint inhibitors: a systematic review. Br J Cancer 2025 Jan;132(1):51-57. [doi: 10.1038/s41416-024-02887-1] [Medline: 39489880]
17. Girard N, Ponce Aix S, Cedres S, et al. Efficacy and safety of nivolumab for patients with pre-treated type B3 thymoma and thymic carcinoma: results from the EORTC-ETOP NIVOTHYM phase II trial. ESMO Open 2023 Jun;8(3):101576. [doi: 10.1016/j.esmoop.2023.101576] [Medline: 37285717]
18. Hao Y, Lin G, Xiang J, et al. Analysis of the efficacy and safety of immunotherapy in advanced thymoma patients. Cancer Med 2023 Mar;12(5):5649-5655. [doi: 10.1002/cam4.5357] [Medline: 36394097]
19. FDA Adverse Event Reporting System (FAERS). US Food & Drug Administration. URL: https://fis.fda.gov/extensions/FPD-QDE-FAERS/FPD-QDE-FAERS.html [accessed 2025-02-25]
20. Fusaroli M, Salvo F, Begaud B, et al. The Reporting of a Disproportionality Analysis for Drug Safety Signal Detection Using Individual Case Safety Reports in PharmacoVigilance (READUS-PV): development and statement. Drug Saf 2024 Jun;47(6):575-584. [doi: 10.1007/s40264-024-01421-9]
21. Fusaroli M, Salvo F, Begaud B, et al. The REporting of A Disproportionality Analysis for DrUg Safety Signal Detection Using Individual Case Safety Reports in PharmacoVigilance (READUS-PV): explanation and elaboration. Drug Saf 2024 Jun;47(6):585-599. [doi: 10.1007/s40264-024-01423-7] [Medline: 38713347]
22. Abe J, Umetsu R, Mataki K, et al. Analysis of Stevens-Johnson syndrome and toxic epidermal necrolysis using the Japanese Adverse Drug Event Report database. J Pharm Health Care Sci 2016 Dec;2(1):14. [doi: 10.1186/s40780-016-0048-5]
23. Nakamura M, Umetsu R, Abe J, et al. Analysis of the time-to-onset of osteonecrosis of jaw with bisphosphonate treatment using the data from a spontaneous reporting system of adverse drug events. J Pharm Health Care Sci 2015;1:34. [doi: 10.1186/s40780-015-0035-2] [Medline: 26819745]
24. Sauzet O, Carvajal A, Escudero A, Molokhia M, Cornelius VR. Illustration of the Weibull shape parameter signal detection tool using electronic healthcare record data. Drug Saf 2013 Oct;36(10):995-1006. [doi: 10.1007/s40264-013-0061-7]

XSL•FO
RenderX

25. Zhou C, Lin A, Li M, Qi C, Zhang J, Luo P. MOAHIT: a web tool for visualizing tumor multi-omics data with human anatomy heatmaps. bioRxiv. Preprint posted online on 2023. [doi: 10.1101/2022.09.07.506938]
26. Cho J, Kim HS, Ku BM, et al. Pembrolizumab for patients with refractory or relapsed thymic epithelial tumor: an open-label phase II trial. J Clin Oncol 2019 Aug 20;37(24):2162-2170. [doi: 10.1200/JCO.2017.77.3184] [Medline: 29906252]
27. Kong X, Chen L, Su Z, et al. Toxicities associated with immune checkpoint inhibitors: a systematic study. Int J Surg 2023 Jun 1;109(6):1753-1768. [doi: 10.1097/JS9.0000000000000368] [Medline: 37132038]
28. Remon J, Villacampa G, Facchinetti F, et al. Immune checkpoint blockers in patients with unresectable or metastatic thymic epithelial tumours: a meta-analysis. Eur J Cancer 2023 Feb;180:117-124. [doi: 10.1016/j.ejca.2022.12.005] [Medline: 36592506]
29. Ao YQ, Gao J, Wang S, et al. Immunotherapy of thymic epithelial tumors: molecular understandings and clinical perspectives. Mol Cancer 2023 Apr 13;22(1):70. [doi: 10.1186/s12943-023-01772-4] [Medline: 37055838]
30. Song X, Fan J, Zhu L, Wang Z, He Y, Zhou C. The efficacy and safety of immunotherapy in thymic epithelial tumors: more effective, more risky: a systematic review. J Thorac Dis 2021 Aug;13(8):5093-5103. [doi: 10.21037/jtd-21-290] [Medline: 34527346]
31. Perrino M, Cordua N, De Vincenzo F, et al. Thymic epithelial tumor and immune system: the role of immunotherapy. Cancers (Basel) 2023 Nov 25;15(23):5574. [doi: 10.3390/cancers15235574] [Medline: 38067278]
32. Ashby KM, Hogquist KA. A guide to thymic selection of T cells. Nat Rev Immunol 2024 Feb;24(2):103-117. [doi: 10.1038/s41577-023-00911-8] [Medline: 37464188]
33. Weksler B, Lu B. Alterations of the immune system in thymic malignancies. J Thorac Oncol 2014 Sep;9(9 Suppl 2):S137-S142. [doi: 10.1097/JTO.0000000000000299] [Medline: 25396311]
34. Perrino M, Voulaz E, Balin S, et al. Autoimmunity in thymic epithelial tumors: a not yet clarified pathologic paradigm associated with several unmet clinical needs. Front Immunol 2024;15:1288045. [doi: 10.3389/fimmu.2024.1288045] [Medline: 38629065]
35. Xin Z, Lin M, Hao Z, et al. The immune landscape of human thymic epithelial tumors. Nat Commun 2022 Sep 17;13(1):5463. [doi: 10.1038/s41467-022-33170-7] [Medline: 36115836]
36. Maniar R, Loehrer PJ. Understanding the landscape of immunotherapy in thymic epithelial tumors. Cancer 2023 Apr 15;129(8):1162-1172. [doi: 10.1002/cncr.34678]
37. Sutanto H, Safira A, Fetarayani D. From tumor to tolerance: a comprehensive review of immune checkpoint inhibitors and immune-related adverse events. Asia Pac Allergy 2024 Aug;14(3):124-138. [doi: 10.5415/apallergy.0000000000000146] [Medline: 39220570]
38. Nguyen LS, Cooper LT, Kerneis M, et al. Systematic analysis of drug-associated myocarditis reported in the World Health Organization pharmacovigilance database. Nat Commun 2022 Jan 10;13(1):25. [doi: 10.1038/s41467-021-27631-8] [Medline: 35013204]
39. Power JR, Alexandre J, Choudhary A, et al. Association of early electrical changes with cardiovascular outcomes in immune checkpoint inhibitor myocarditis. Arch Cardiovasc Dis 2022 May;115(5):315-330. [doi: 10.1016/j.acvd.2022.03.003]
40. Wang DY, Salem JE, Cohen JV, et al. Fatal toxic effects associated with immune checkpoint inhibitors: a systematic review and meta-analysis. JAMA Oncol 2018 Dec 1;4(12):1721-1728. [doi: 10.1001/jamaoncol.2018.3923] [Medline: 30242316]
41. Johnson DB, Balko JM, Compton ML, et al. Fulminant myocarditis with combination immune checkpoint blockade. N Engl J Med 2016 Nov 3;375(18):1749-1755. [doi: 10.1056/NEJMoa1609214] [Medline: 27806233]
42. Salem JE, Manouchehri A, Moey M, et al. Cardiovascular toxicities associated with immune checkpoint inhibitors: an observational, retrospective, pharmacovigilance study. Lancet Oncol 2018 Dec;19(12):1579-1589. [doi: 10.1016/S1470-2045(18)30608-9] [Medline: 30442497]
43. Fenioux C, Abbar B, Boussouar S, et al. Thymus alterations and susceptibility to immune checkpoint inhibitor myocarditis. Nat Med 2023 Dec;29(12):3100-3110. [doi: 10.1038/s41591-023-02591-2] [Medline: 37884625]
44. Nguyen LS, Bretagne M, Arrondeau J, et al. Reversal of immune-checkpoint inhibitor fulminant myocarditis using personalized-dose-adjusted abatacept and ruxolitinib: proof of concept. J Immunother Cancer 2022 Apr;10(4):e004699. [doi: 10.1136/jitc-2022-004699] [Medline: 35383117]
45. Axelrod ML, Meijers WC, Screever EM, et al. T cells specific for α-myosin drive immunotherapy-related myocarditis. Nature New Biol 2022 Nov;611(7937):818-826. [doi: 10.1038/s41586-022-05432-3] [Medline: 36385524]
46. Blum SM, Zlotoff DA, Smith NP, et al. Immune responses in checkpoint myocarditis across heart, blood and tumour. Nature New Biol 2024 Dec;636(8041):215-223. [doi: 10.1038/s41586-024-08105-5] [Medline: 39506125]
47. Sakaeda T, Tamon A, Kadoyama K, Okuno Y. Data mining of the public version of the FDA Adverse Event Reporting System. Int J Med Sci 2013;10(7):796-803. [doi: 10.7150/ijms.6048] [Medline: 23794943]

Abbreviations

AE: adverse event
FAERS: US Food and Drug Administration Adverse Event Reporting System
FDA: Food and Drug Administration

XSL•FO
RenderX

ICI: immune checkpoint inhibitor
irAEs: immune-related adverse events
MedDRA: Medical Dictionary for Regulatory Activities
ORR: objective response rate
PD-1: programmed death-1
PD-L1: programmed death-ligand 1
PFS: progression-free survival
PT: preferred term
ROR: reporting odds ratio
TET: thymic epithelial tumor
WSP: Weibull Shape Parameter

Edited by M Balcarras, N Cahill; submitted 04.May.2025; peer-reviewed by S Matsuda, V Giunchi, Y Zhai; revised version received 28.Dec.2025; accepted 02.Jan.2026; published 25.Feb.2026.

Please cite as:
Chen R, Chen H, Sun L, Cao Y, Lin L
Immunotherapy Safety in Thymic Epithelial Tumors: Disproportionality Analysis of the Food and Drug Administration Adverse Event Reporting System
JMIR Cancer 2026;12:e76908
URL: https://cancer.jmir.org/2026/1/e76908
doi:10.2196/76908

© Ruilian Chen, Hanrui Chen, Lingling Sun, Yang Cao, Lizhu Lin. Originally published in JMIR Cancer (https://cancer.jmir.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Hemophagocytic Lymphohistiocytosis and Describing Outcomes Using Computable Phenotypes: Retrospective Cohort Study

Suheyla Ocak[1,2], MD; Martin Yi[3], MDA; Agata Wolochacz[3], BSc; Ida Mehrdadi[3], BSc; Ahmed Naqvi[1], MD; Sumit Gupta[1,3], MD, PhD; Lillian Sung[1,3], MD, PhD; Adam P Yan[1,3], MD, MBI

[1]Division of Haematology/Oncology, Hospital for Sick Children, 555 University Avenue, Toronto, ON, Canada

[2]Department of Pediatric Hematology and Oncology, Cerrahpasa Medical Faculty, Istanbul University-Cerrahpasa, Istanbul, Turkey

[3]Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, ON, Canada

Corresponding Author:
Adam P Yan, MD, MBI
Division of Haematology/Oncology, Hospital for Sick Children, 555 University Avenue, Toronto, ON, Canada

Abstract

Background: Hemophagocytic lymphohistiocytosis (HLH) is a life-threatening hyperinflammatory syndrome that requires rapid diagnosis and intervention. However, identifying these patients is difficult because the HLH-2004 diagnostic criteria are complex and not always captured systematically in electronic health records (EHRs). Furthermore, it is unclear how clinicians use these criteria to diagnose HLH and make treatment decisions. There is a critical need for validated computable phenotypes to accurately identify patients and study treatment-related outcomes in HLH.

Objective: The aim of this study is to compare different approaches to using the EHR to build computable phenotypes of patients with HLH and to evaluate characteristics and outcomes of patients meeting the HLH-2004 diagnostic criteria who received HLH-directed therapies compared to those who did not.

Methods: Three approaches to computable phenotype development in the EHR were taken by identifying patients (1) with an HLH-specific *International Statistical Classification of Diseases, Tenth Revision* (*ICD-10*) code, (2) with an HLH-specific treatment plan, and (3) meeting the HLH-2004 clinical criteria for diagnosis of HLH. Among patients who met the HLH-2004 criteria, we evaluated the characteristics and outcomes of patients who received HLH-directed therapies compared to those who did not. HLH treatment was defined as either any chemotherapy or HLH-specific therapy (dexamethasone, methylprednisolone, anakinra, ruxolitinib, cyclosporine, etoposide, or emapalumab).

Results: We identified 388 patients with possible HLH across the three cohorts. An HLH *ICD-10* diagnosis (n=220) and meeting 5 or more clinical criteria (n=245) were much more common than a HLH treatment plan (n=42). Among the patients meeting HLH-2004 clinical criteria, 193 (79%) received HLH-directed therapy. There was no difference in any specific HLH criteria between those who did and did not receive HLH-directed therapy. In-hospital mortality was very high among both groups and was 15% among those who received HLH-directed therapy and 13.5% among those who did not receive HLH-directed therapy. Among 1325 patients with an elevated ferritin and fever, only 252 (19%) met >5 clinical criteria.

Conclusions: Constructing HLH cohorts from EHR data is challenging, with diagnosis codes, treatment plans, and clinical criteria each capturing distinct but overlapping populations.

(JMIR Cancer 2026;12:e87347) doi:10.2196/87347

KEYWORDS

HLH; electronic health record; pediatrics; computable phenotype; oncology; hemophagocytic lymphohistiocytosis

Introduction

Hemophagocytic lymphohistiocytosis (HLH) is a rare but often life-threatening syndrome represented by excessive inflammation resulting in organ dysfunction and potentially death [1]. Early recognition and treatment of HLH are critical to patient outcomes. Traditionally, HLH has been classified as primary, where there is a documented or presumed genetic etiology [2,3], or secondary, where excessive inflammation can be attributed to a trigger such as infection, malignancy, or a rheumatological condition in the absence of a genetic etiology [4]. Although there are an increasing number of genes identified responsible for primary HLH [5], there are patients with presumed primary HLH without a known mutation. Further, patients with both primary and secondary HLH may require HLH-directed therapy. Consequently, diagnostic criteria have

XSL•FO
RenderX

been developed for HLH (Multimedia Appendix 1), which often guide decision-making about treatment initiation. Criteria were initially proposed in 2004 [2,3] and recently updated in 2024 [3].

In general, diagnosis of HLH can be made based upon molecular or functional cellular findings consistent with HLH in combination with meeting at least 5 clinical criteria (Textbox 1). While these criteria have been widely used for diagnosis and treatment decision-making, there have been questions raised about the specificity of these criteria, particularly as they relate to secondary HLH [4]. HLH frequently presents with nonspecific clinical features that overlap with severe infection, sepsis, malignancy, and other inflammatory syndromes, leading to substantial diagnostic uncertainty in real-world clinical practice. Most studies of HLH have focused on small cohorts of participants who have received a clinical diagnosis of HLH. To evaluate the utility of HLH criteria, it may be informative to determine the association of each individual HLH criterion against an overall clinical HLH diagnosis, whether HLH-specific treatments received or an HLH outcome such as mortality.

XSL•FO
RenderX

Textbox 1. HLH-2004 and HLH-2024 diagnostic criteria.

HLH-2004 [3]

Criterion 1 or 2 is fulfilled.

1. A molecular diagnosis consistent with hemophagocytic lymphohistiocytosis (HLH)
2. Diagnostic criteria for HLH fulfilled (5 of the 8 criteria below):

- Fever
- Splenomegaly
- Cytopenias (affecting ≥2 of 3 lineages in the peripheral blood)
  - Hemoglobin <90 g/L (hemoglobin <100 g/L in infants <4 wk)
  - Platelets $<100\times10^9$/L
  - Neutrophils $<1.0\times10^9$/L
- Hypertriglyceridemia and/or hypofibrinogenemia
  - Fasting triglycerides ≥3.0 mmol/L (ie, ≥265 mg/dL)
  - Fibrinogen ≤1.5 g/L
- Hemophagocytosis in bone marrow or spleen or lymph nodes. No evidence of malignancy.
- Low or no natural killer cell activity (according to local laboratory reference)
- Ferritin ≥500 µg/L
- CD25 (ie, soluble IL-2 receptor) ≥2400 U/mL

HLH-2024 [2]

Criterion 1, 2, or 3 below is fulfilled.

1. Functional cellular findings consistent with familial HLH (FHL) in a patient with signs/symptoms suggestive of HLH
2. A molecular diagnosis consistent with FHL in a patient with signs/symptoms suggestive of HLH
3. Clinical diagnostic criteria for FHL with at least 5 of the 7 criteria below fulfilled:

- Fever ≥38.5°C
- Splenomegaly (≥2 cm below the costal margin)
- Cytopenias (affecting ≥2 of 3 lineages in the peripheral blood)
  - Hemoglobin <90 g/L (hemoglobin <100 g/L in infants <4 wk)
  - Platelets $<100\times10^9$/L
  - Neutrophils $<1.0\times10^9$/L
- Hypertriglyceridemia and/or hypofibrinogenemia
  - Fasting triglycerides ≥3.0 mmol/L
  - Fibrinogen ≤1.5 g/L
- Hemophagocytosis
- Ferritin ≥500 µg/L
- sCD25 (ie, soluble IL-2 receptor) ≥2400 U/mL

Most studies of patients with HLH use small, highly curated datasets. Such studies are therefore limited in their ability to evaluate how well diagnostic criteria perform in broader, unselected clinical populations. Few studies have leveraged electronic health record (EHR) data, which may represent broader populations and contain a wider range of clinical information. Leveraging EHR data to facilitate clinical research of HLH requires the construction of an accurate patient cohort. Given that EHR data are not entered with the goal of facilitating future downstream research, manually entered data such as diagnosis codes can often be incorrect or missing. Computable phenotypes are machine-evaluable definitions for a given

XSL•FO
RenderX

condition developed using EHR data. Multisource computable phenotypes incorporating laboratory results, medications, procedures, and clinical events may improve case ascertainment and reduce bias compared with diagnosis codes alone [6]. To our knowledge, no attempt has been made to use diverse EHR data to develop a machine-evaluable approach to HLH identification [7-9]. It is unknown how alternative EHR-based definitions perform for a clinically heterogeneous and dynamically evolving syndrome such as HLH.

The primary objective of this study was to compare various approaches to using EHR data to build a computable phenotype of patients with HLH—specifically, we compared three approaches: (1) HLH-specific diagnosis code usage, (2) HLH-specific treatment plan application, and (3) HLH clinical criteria, among all patients at a pediatric hospital. Our hypothesis was that the use of diagnosis codes and treatment plans would undercapture patients and the use of clinical criteria would overcapture patients. Among patients who met at least five HLH clinical criteria, the secondary objective was to explore characteristics and outcomes among those who received HLH-directed therapy versus those who did not.

Methods

Ethical Considerations

This observational study was approved by the Research Ethics Board at The Hospital for Sick Children (SickKids). The requirement for informed consent and assent were waived given the retrospective nature of the study. In keeping with our Research Ethics Board approval, no potentially identifying information of any participant is disclosed in this report. All data are presented as aggregates.

Data Source

The data source was the SickKids Enterprise-wide Data in Azure Repository (SEDAR) [10]. SEDAR is a curated and validated version of the Epic Clarity database that lives in Microsoft Azure. To create SEDAR, we transformed Clarity first into a Filtered Schema that retains only valid SickKids patients, then into a curated schema where the data are structured into clinically relevant units such as patients, encounters, laboratory tests, and medication administrations. During SEDAR creation, all data were validated via random sampling. This project focused on the following SEDAR tables: patient, diagnoses, cancer treatment plans, hospital encounter, nonhospital encounter, medication administration, prescriptions, laboratory tests, pathology results, flowsheets, and notes.

Operationalizing HLH Criteria

We used the HLH-2004 criteria (as opposed to the HLH-2024 criteria) as these would have been the criteria in place for most of the patient cohort. The time window to evaluate clinical criteria was centered on an episode, which was usually an inpatient admission spanning admission to discharge. A criterion was considered met if it occurred at any point during the episode. Five of the criteria were based on the laboratory results table (ferritin, cytopenia, hypertriglyceridemia or hypofibrinogenemia, sCD25, and low natural killer cell activity) and were considered met if the patient ever had a value above or below the criteria, as appropriate. Fever was defined as an oral temperature of at least 38.3°C once or 38.0°C to 38.2°C for at least one hour [11]. If a laboratory test was not performed or a criterion was not met, the criteria was categorized as not abnormal.

Two of the criteria required searching of text. Hemophagocytosis was identified by searching for “h(a)emophagocytosis” or “h(a)emophagocytic” in all pathology results. These reports were manually reviewed to identify true hemophagocytosis. Splenomegaly was identified by searching all notes for any of the following terms “splenomegaly,” “big spleen,” “organomegaly,” and “enlarged spleen.” Negative terms were excluded using the following terms within 3 words previous to the splenomegaly term: “no,” “none,” “absence,” “without,” and “negative.” The number of mentions were too high to manually review each note. Thus, a random sample of 20 notes underwent chart review to validate the approach. We found that 19/20 were correct. One of 20 was incorrect. We considered this satisfactory to proceed.

Eligibility Criteria

We established three cohorts of HLH “diagnosis” based upon encounters that occurred between June 2, 2018, and May 31, 2025. First, patients with a coded diagnosis of HLH were those with an *International Statistical Classification of Diseases, Tenth Revision* (*ICD-10*) code of D76.1 (hemophagocytic lymphohistiocytosis) or D76.2 (hemophagocytic syndrome, infection-associated). Second, cancer treatment plans are electronic care plans that are a component of Epic’s Beacon oncology module. All chemotherapy at our institution must be ordered within a treatment plan; however, treatment plan use is restricted to use by oncologists. To that end, an oncologist would therefore order a drug such as emapalumab within a treatment plan, while a rheumatologist using the same drug would not. We manually identified either treatment plan or protocol display names that included “HLH” or “h(a)emophagocytosis.” Use of a treatment plan was defined as having an HLH-specific treatment plan applied in Epic.

The third approach consisted of identifying the number of HLH-2004 clinical criteria within an encounter. The HLH-specific encounter was the encounter with the maximum number of criteria. If there was more than one encounter with this number of criteria, the first encounter was selected. For establishment of HLH based on clinical criteria, we identified encounters with at least 5 criteria within that encounter regardless of encounter length.

Given that there is no gold-standard adjudicated HLH cohort, determining which computable phenotype for cohort creation was most accurate was not possible with this data set.

Procedure

We were interested in two measures of whether patients received HLH treatment; these might be administered within or outside of an HLH treatment plan. First, we considered any chemotherapy. Second, we considered HLH-directed therapy, which we defined as receipt of any dosage of dexamethasone, methylprednisolone, anakinra, ruxolitinib, cyclosporine, etoposide, or emapalumab [12-16].

XSL•FO
RenderX

We compared demographic features, HLH-related variables, and clinical outcomes among those with at least 5 HLH clinical criteria for those who received HLH-directed therapy (yes vs no) during the encounter where the criteria were met. The goal of this analysis was description and hypothesis generation given the possibility of confounding by indication for treatment. Demographic features were sex and age group (0-<1, 1 - 4, 5 - 14, and ≥15 years). HLH-related variables were whether there was an *ICD-10* HLH diagnosis associated with that encounter, timing of *ICD-10* HLH diagnosis relative to the encounter, and specific clinical criteria met. Clinical outcomes were services involved, length of stay, intensive care unit admission, in-hospital mortality, 30-day mortality, and whether the patient underwent stem cell transplantation following the encounter. To calculate 30-day mortality, we identified all patients with a date of death within 30 days of the visit where they met the HLH-2004 criteria. We also evaluated the 5 most common admission diagnoses among the cohort with at least 5 clinical criteria. Primary admission diagnosis was derived from the admitting diagnosis field in our EHR.

As an exploratory objective, we hypothesized that high ferritin and fever are common conditions in pediatric patients. Thus, we identified all encounters that met the HLH-2004 ferritin and fever criteria. Any encounter where the patient met both criteria at any point during the encounter were counted. We then described the distribution of the other HLH-2004 criteria in this cohort and stratified encounters by whether there were at least 5 clinical criteria present during an HLH-specific encounter.

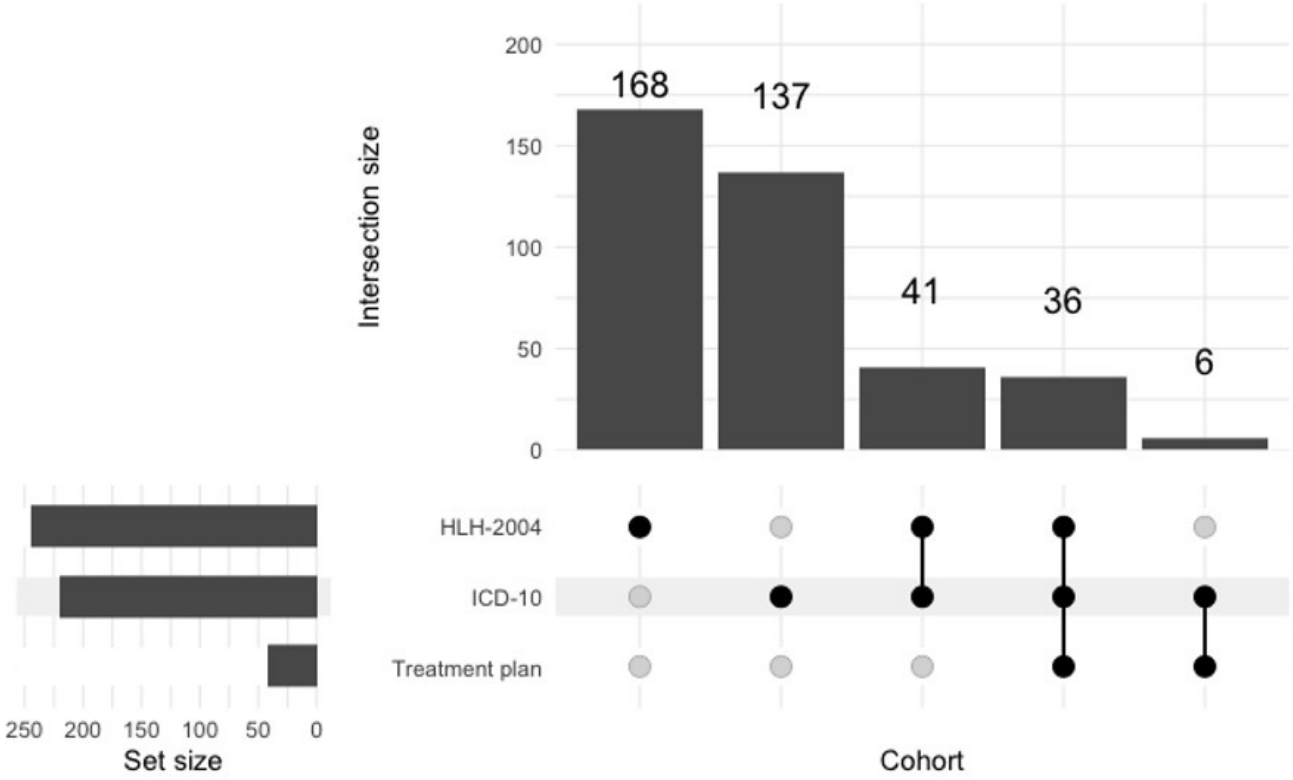
Analysis

To compare patients who received HLH-directed therapy (yes versus no) among patients who met at least 5 clinical criteria, we used the Wilcoxon rank-sum test for continuous variables and χ^2 or Fisher exact test for categorical variables. Absolute risk differences in in-hospital mortality, 30-day mortality, and median length of stay were calculated with 95% confidence intervals. Analyses were performed with R (version 4.2.2; R Foundation for Statistical Computing).

Results

Figure 1 shows the intersection of the 388 patients identified as having HLH using our three approaches to HLH-cohort creation using three different computable phenotypes (diagnosis code, treatment plan, or HLH criteria). All patients with an HLH-specific treatment plan also had either an *ICD-10* HLH diagnosis or met clinical HLH criteria. Only 36 patients were identified by all three methods as having HLH. Table 1 shows the number of patients with an HLH *ICD-10* diagnosis code, an HLH treatment plan, or who met at least 5 HLH clinical criteria. An HLH *ICD-10* diagnosis (n=220 patients) and meeting 5 or more clinical criteria (n=245) were much more common than an HLH treatment plan (n=42 patients). Among the 220 patients with an HLH *ICD-10* code, only 35 patients received chemotherapy during the HLH-specific encounter. However, 85 patients received HLH-directed therapy during the HLH-specific encounter. Conversely, among the 245 patients who met at least 5 clinical criteria, 78 patients received chemotherapy and 193 patients received HLH-directed therapy during the HLH-specific encounter.

Figure 1. Intersection of HLH cohorts derived using three different approaches to cohort creation (N=388 patients). HLH: hemophagocytic lymphohistiocytosis; *ICD-10*: *International Statistical Classification of Diseases, Tenth Revision*.



XSL•FO
RenderX

Table . Number of patients with possible hemophagocytic lymphohistiocytosis (HLH) by different approaches.

	HLH *ICD-10*[a] diagnosis code	HLH treatment plan	Five or more HLH criteria
Total number of patients	220	42	245
Number of patients with chemotherapy during HLH-specific encounter[b]	35	37	78
Number of patients with chemotherapy during any encounter	73	38	99
Number of patients with chemotherapy or HLH-directed therapy[c] during HLH-specific encounter	85	41	198
Number of patients with chemotherapy or HLH-directed therapy during any encounter	143	42	213
Number of patients with HLH-directed therapy during HLH-specific encounter	85	41	193
Number of patients with HLH-directed therapy during any encounter	142	42	212

[a]*ICD-10*: *International Statistical Classification of Diseases, Tenth Revision.*

[b]HLH-specific encounter is associated with first HLH diagnosis, first HLH treatment plan, or first encounter with the maximum number of criteria met.

[c]HLH-directed therapy includes dexamethasone, methylprednisolone, anakinra, ruxolitinib, cyclosporine, etoposide, and emapalumab.

Table 2 shows the demographic features of unique patients who met 5 or more HLH criteria stratified by receipt of HLH-directed therapy versus not. The median age of the cohort was 7.7 years (IQR 1.4 - 13.9). Of the 193 patients who received HLH-directed therapies, 67.9% (n=131) only received steroids and 29% (n=56) received steroids in combination with an additional HLH-directed therapy. Those with sepsis were less likely to receive HLH-directed therapy (13.5% vs 4.7%; *P*=.05). Those patients receiving HLH-directed therapy were more likely to have an *ICD-10* HLH diagnosis code during that encounter (*P*=.001). Notably, there was no difference in any specific HLH criteria between those who did and did not receive HLH-directed therapy although high ferritin was almost universal in both groups. The absolute difference in in-hospital mortality and 30-day mortality rates between the groups was 1.5% (95% CI –1.1 to 1.1) and 1.2% (95% CI –1.2 to 1.1), respectively. The absolute difference in median length of stay between the two groups was 12.6 days (95% CI 4.1 to 22.4).

XSL•FO
RenderX

Table . Number of patients with at least 5 hemophagocytic lymphohistiocytosis (HLH) criteria stratified by receipt of HLH-directed therapy.

	HLH-directed therapy (n=193)	No HLH-directed therapy (n=52)	*P* value
Demographic features			
Male sex, n (%)	96 (49.7)	27 (51.9)	.90
Age group in years, n (%)			.34
0-<1	43 (22.3)	13 (25)	
1 - 4	44 (22.8)	12 (23.1)	
5 - 14	65 (33.7)	14 (26.9)	
≥15	41 (21.2)	12 (23.1)	
Primary admission diagnosis[a], n (%)			
Fever	37 (19.2)	14 (26.9)	.30
Sepsis	9 (4.7)	7 (13.5)	.05
Leukemia	12 (6.2)	2 (3.8)	.75
Acute lymphoblastic leukemia	8 (4.1)	4 (7.7)	.49
Neutropenia	6 (3.1)	3 (5.8)	.62
HLH-related variables			
ICD-10[b] HLH diagnosis, n (%)	77 (31.4)	5 (9.6)	<.001
ICD-10 HLH diagnosis relative to encounter, n (%)			.001
Before	8 (4.1)	0 (0)	
During	55 (28.5)	3 (5.8)	
After	9 (4.7)	2 (3.8)	
Never had HLH diagnosis	121 (62.7)	47 (90.4)	
HLH criteria present, n (%)			
High ferritin	190 (98.4)	52 (100)	.85
Fever	177 (91.7)	47 (90.4)	.98
Cytopenia of at least 2 lineages	176 (91.2)	48 (92.3)	>.99
Hypertriglyceridemia or hypofibrinogenemia	183 (94.8)	47 (90.4)	.39
Splenomegaly	175 (90.7)	46 (88.5)	.83
sCD25[c]	91 (47.2)	26 (50)	.84
Hemophagocytosis	33 (17.1)	3 (5.8)	.07
Low natural killer cell activity[c]	13 (6.7)	1 (1.9)	.32
Clinical outcomes			
Services involved, n (%)			
Oncology	109 (56.5)	22 (42.3)	.01
Hematology	103 (53.4)	33 (63.5)	.25
Rheumatology	89 (46.1)	16 (30.8)	.07
Infectious diseases	159 (82.4)	45 (86.5)	.62
Immunology	46 (23.8)	13 (25)	>.99
Neurology	69 (35.8)	17 (32.7)	.80
Median length of stay (IQR)	33.4 (81.1-14.3)	20.8 (33-8.5)	.001
Intensive care unit, n (%)	111 (57.5)	22 (42.3)	.07
In-hospital mortality, n (%)	29 (15)	7 (13.5)	.95

XSL•FO
RenderX

	HLH-directed therapy (n=193)	No HLH-directed therapy (n=52)	*P* value
Mortality within 30 days, n (%)	32 (16.6)	8 (15.4)	>.99
Stem cell transplant following encounter, n (%)	13 (6.7)	1 (1.9)	.32

[a]Five most common admitting diagnoses for patients with at least 5/8 HLH criteria (n=245).

[b]*ICD-10*: *International Statistical Classification of Diseases, Tenth Revision*.

[c]Number of encounters where measured for HLH-directed therapy (yes versus no): sCD25: 135 versus 36; natural killer cell activity: 53 versus 7.

Table 3 shows that among the 1325 encounters (where individual patients can have more than one encounter) with high ferritin and fever, cytopenia, hypertriglyceridemia or hypofibrinogenemia, and splenomegaly were very common. Of the 1325 encounters, 19% (n=252) had >5 HLH criteria. Even among those with less than 5 clinical criteria, cytopenia was present in 44% (n=475) of encounters while hypertriglyceridemia or hypofibrinogenemia and splenomegaly were present in 24% (n=255) and 27% (n=290) of encounters, respectively.

Table . Distribution of criteria among encounters with high ferritin and fever[a].

	All encounters (N=1325 encounters), n (%)	5 or more criteria (n=252 encounters), n (%)	<5 criteria (n=1073 encounters), n (%)
Cytopenia of at least 2 lineages	706 (53)	231 (92)	475 (44)
Hypertriglyceridemia or hypofibrinogenemia	485 (37)	230 (91)	255 (24)
Splenomegaly	519 (39)	229 (91)	290 (27)
sCD25[b]	159 (12)	114 (45)	45 (4)
Hemophagocytosis	27 (2)	27 (11)	0 (0)
Low natural killer cell activity[b]	12 (1)	11 (4)	1 (0.1)

[a]The table represents all encounters where ferritin and fever met the threshold for HLH diagnosis without considering if the patient was diagnosed with HLH by any approach. There may be multiple episodes per patient.

[b]Number of encounters where measured for ≥5 versus <5 criteria were as follows: sCD25: 172 (68%) versus 165 (15%); natural killer cell activity: 55 (22%) versus 24 (2%).

Multimedia Appendix 1 shows the distribution of clinical criteria among those with an HLH *ICD-10* code and among those with at least 5 clinical criteria. Some patients with their first HLH *ICD-10* diagnosis code had no clinical criteria.

Discussion

In this study, we compared three approaches to constructing an EHR-derived cohort of pediatric patients with HLH: (1) *ICD-10* diagnosis codes, (2) HLH-specific treatment plans, and (3) fulfillment of HLH-2004 clinical criteria. Our findings demonstrate substantial variation in the populations identified by each approach, with diagnosis codes and clinical criteria identifying a far greater number of encounters than HLH-specific treatment plans. We also identified that among patients meeting at least five HLH-2004 clinical criteria, a large proportion received HLH-directed therapy, and that among a cohort of patients with fever and an elevated ferritin, a substantial portion have other findings of HLH such as cytopenia, hypertriglyceridemia or hypofibrinogenemia or splenomegaly.

Comparing our three approaches to EHR-based cohort construction yields numerous interesting insights. First, we observed that HLH *ICD-10* codes were relatively common in our dataset. A large proportion of patients with an HLH *ICD-10* code did not receive HLH-directed therapy, suggesting that coded diagnoses may be used for suspected but unconfirmed HLH. Conversely, while HLH treatment plans were highly specific, they were rarely used. This is likely because many pediatric patients with HLH do not require chemotherapy and receive HLH-directed medications that are ordered outside of a chemotherapy plan. Thus, the use of treatment plans, which are likely the most accurate way to identify HLH-directed medications, likely underestimates the true HLH population. Clinical criteria, while nonspecific, identified the broadest cohort, consistent with the recognition that HLH represents a clinical syndrome overlapping with other hyperinflammatory states. The exclusive use of these criteria can also be problematic in an EHR cohort, where certain laboratory tests may be done at a referring institution such as for a patient with HLH being referred for a hematopoietic stem cell transplant. This is supported by the fact that we had some patients who met zero HLH clinical criteria but had an HLH *ICD-10* code applied. Future research leveraging EHR data to assemble an HLH cohort needs to consider the advantages and disadvantages of different techniques for cohort assembly to select the best approach for a given project. For example, a researcher interested in understanding the diagnostic odyssey of patients with HLH prior to treatment to improve time to diagnosis might use diagnostic criteria as this would more broadly capture patients compared to using a more narrow approach like diagnostic codes. Future work should also explore the usage of a

XSL•FO
RenderX

combination of approaches such as having both HLH-criteria met and an HLH diagnosis code and using additional filters in the treatment approach such as dosages or duration for medications that might be used for indications other than HLH management such as dexamethasone. It will also be important to validate these approaches at other institutions to understand generalizability.

Using our newly developed EHR-based cohort of children with HLH, we were able to generate new insights into the management and outcomes of patients with HLH. Of the 245 patients who met the HLH-2004 criteria for HLH, 27% (n=52) received no HLH-directed therapy. No differences were identified in the occurrence of HLH-specific criteria such as ferritin or fever between the two cohorts. From this work, it remains unclear as to what drove the difference in clinical decision-making to start HLH-directed therapy for patients meeting 5 criteria. We did identify that patients who did not receive HLH-directed therapy were more likely to have sepsis than those who did [17]. This is not surprising given that we know that sepsis and HLH have an overlapping presentation. Patients who received HLH-directed therapy also had a significantly longer length of stay (33.4 vs 20.8 d). It is important to note that regardless of whether or not a patient who met 5 criteria received therapy, their 30-day mortality risk was approximately 1 in 6. This is similar to other contemporary studies of patients treated for HLH, where mortality rates have ranged from 23% to 55% [14,18-20]. Given the high mortality rate in those who did not receive therapy, it raises the question of whether or not we are failing to treat a subset of patients with HLH with appropriate HLH-directed therapies.

We also evaluated the distribution of HLH criteria within an EHR-derived cohort of patients with high ferritin and fever. An elevated ferritin and fever are often used as the entrance criteria into a "rule-out" HLH algorithm at many institutions that results in the ordering of additional HLH-specific labs or consultation of an HLH-specific team such as hematology/oncology and/or rheumatology [16,21]. In this cohort that mirrors a potential cohort of patients where an HLH work-up and diagnosis may be considered, we identified that a significant proportion of patients will meet additional HLH criteria. However, only 20% (n=252) of the patients in this cohort met >5 HLH criteria, suggesting that if fever and ferritin are used as entrance criteria into an HLH algorithm, 5 patients would need to be evaluated to identify 1 patient meeting the HLH clinical criteria.

The strengths of this study are access to all data in a curated and validated EHR data set, which allows for the inclusion of patients without an HLH diagnosis. Unlike prior studies restricted to curated clinical cohorts, our approach provides insight into the challenges of applying HLH diagnostic frameworks in real-world, unselected populations. However, limitations of this study include that it was done at a single pediatric institution, so results may not generalize to other institutions. Specifically, in the Canadian context, we may underutilize ICD codes compared to American peers given different implications on billing and remuneration [22]. Second, we were limited in validating the performance of our cohorts against a gold-standard cohort, as no such curated dataset exists that includes both primary and secondary patients with HLH. Third, EHR-derived data are prone to misclassification. Diagnosis codes may not reflect confirmed diagnoses, laboratory results may be incomplete (as SickKids is a large referral center so some labs may have been done at a referring institution and may not have been included in our dataset), and text-mining strategies (eg, for splenomegaly) have imperfect specificity. Fourth, we centered our analysis on the HLH-2004 criteria given these were the criteria being used by clinicians during our study period. The HLH-2024 criteria have since been published, which differ slightly from the HLH-2004 criteria by removal of the natural killer cell activity criterion. Given that only 5% (n=60) of patients had natural killer cell activity testing performed in our cohort, removal of this criteria is unlikely to impact results. Finally, treatment receipt was defined by medication administration records, which may not fully capture the intent of medication prescribing, and this analysis may be confounded by medication indication.

Constructing HLH cohorts from EHR data is challenging, with diagnosis codes, treatment plans, and clinical criteria each capturing distinct but overlapping populations. While clinical criteria identify the broadest group, they lack specificity, underscoring the limitations of applying rigid diagnostic frameworks to real-world EHR data. Developing accurate, machine-evaluable HLH phenotypes will be critical to advancing both clinical care and research in this rare but important condition.

Acknowledgments

Generative artificial intelligence was not used in drafting any portion of this manuscript.

Funding

LS is supported by the Canada Research Chair in Pediatric Oncology Supportive Care. This project was supported by the SickKids Foundation at the Hospital for Sick Children.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Distribution of HLH clinical criteria.

XSL•FO
RenderX

[DOCX File, 22 KB - cancer_v12i1e87347_app1.docx]

References

1. Janka GE, Lehmberg K. Hemophagocytic lymphohistiocytosis: pathogenesis and treatment. Hematology Am Soc Hematol Educ Program 2013;2013:605-611. [doi: 10.1182/asheducation-2013.1.605] [Medline: 24319239]
2. Henter JI, Sieni E, Eriksson J, et al. Diagnostic guidelines for familial hemophagocytic lymphohistiocytosis revisited. Blood 2024 Nov 28;144(22):2308-2318. [doi: 10.1182/blood.2024025077] [Medline: 39046779]
3. Henter JI, Horne A, Aricó M, et al. HLH-2004: Diagnostic and therapeutic guidelines for hemophagocytic lymphohistiocytosis. Pediatr Blood Cancer 2007 Feb;48(2):124-131. [doi: 10.1002/pbc.21039] [Medline: 16937360]
4. Naymagon L, Roehrs P, Hermiston M, et al. Perspectives on the current diagnostic and treatment paradigms in secondary hemophagocytic lymphohistiocytosis (HLH). Orphanet J Rare Dis 2025 Apr 26;20(1):200. [doi: 10.1186/s13023-025-03698-0] [Medline: 40287693]
5. Sepulveda FE, de Saint Basile G. Hemophagocytic syndrome: primary forms and predisposing conditions. Curr Opin Immunol 2017 Dec;49:20-26. [doi: 10.1016/j.coi.2017.08.004] [Medline: 28866302]
6. Shivade C, Raghavan P, Fosler-Lussier E, et al. A review of approaches to identifying patient phenotype cohorts using electronic health records. J Am Med Inform Assoc 2014;21(2):221-230. [doi: 10.1136/amiajnl-2013-001935] [Medline: 24201027]
7. Richesson RL, Smerek MM, Blake Cameron C. A framework to support the sharing and reuse of computable phenotype definitions across health care delivery and clinical research applications. EGEMS (Wash DC) 2016;4(3):1232. [doi: 10.13063/2327-9214.1232] [Medline: 27563686]
8. Phillips CA, Razzaghi H, Aglio T, et al. Development and evaluation of a computable phenotype to identify pediatric patients with leukemia and lymphoma treated with chemotherapy using electronic health record data. Pediatr Blood Cancer 2019 Sep;66(9):e27876. [doi: 10.1002/pbc.27876] [Medline: 31207054]
9. Hochheiser H, Castine M, Harris D, Savova G, Jacobson RS. An information model for computable cancer phenotypes. BMC Med Inform Decis Mak 2016 Sep 15;16(1):121. [doi: 10.1186/s12911-016-0358-4] [Medline: 27629872]
10. Guo LL, Calligan M, Vettese E, et al. Development and validation of the SickKids Enterprise-wide Data in Azure Repository (SEDAR). Heliyon 2023 Nov;9(11):e21586. [doi: 10.1016/j.heliyon.2023.e21586] [Medline: 38027579]
11. Freifeld AG, Bow EJ, Sepkowitz KA, et al. Clinical practice guideline for the use of antimicrobial agents in neutropenic patients with cancer: 2010 update by the Infectious Diseases Society of America. Clin Infect Dis 2011 Feb 15;52(4):e56-e93. [doi: 10.1093/cid/cir073]
12. Zhang Q, Zhao YZ, Ma HH, et al. A study of ruxolitinib response-based stratified treatment for pediatric hemophagocytic lymphohistiocytosis. Blood 2022 Jun 16;139(24):3493-3504. [doi: 10.1182/blood.2021014860] [Medline: 35344583]
13. Locatelli F, Jordan MB, Allen C, et al. Emapalumab in children with primary hemophagocytic lymphohistiocytosis. N Engl J Med 2020 May 7;382(19):1811-1822. [doi: 10.1056/NEJMoa1911326] [Medline: 32374962]
14. Eloseily EM, Weiser P, Crayne CB, et al. Benefit of anakinra in treating pediatric secondary hemophagocytic lymphohistiocytosis. Arthritis Rheumatol 2020 Feb;72(2):326-334. [doi: 10.1002/art.41103] [Medline: 31513353]
15. Shakoory B, Carcillo JA, Chatham WW, et al. Interleukin-1 receptor blockade is associated with reduced mortality in sepsis patients with features of macrophage activation syndrome: reanalysis of a prior phase III trial. Crit Care Med 2016 Feb;44(2):275-281. [doi: 10.1097/CCM.0000000000001402] [Medline: 26584195]
16. Halyabar O, Chang MH, Schoettler ML, et al. Calm in the midst of cytokine storm: a collaborative approach to the diagnosis and treatment of hemophagocytic lymphohistiocytosis and macrophage activation syndrome. Pediatr Rheumatol Online J 2019 Feb 14;17(1):7. [doi: 10.1186/s12969-019-0309-6] [Medline: 30764840]
17. Machowicz R, Janka G, Wiktor-Jedrzejczak W. Similar but not the same: differential diagnosis of HLH and sepsis. Crit Rev Oncol Hematol 2017 Jun;114:1-12. [doi: 10.1016/j.critrevonc.2017.03.023] [Medline: 28477737]
18. Gregory J, Greenberg J, Basu S. Outcomes analysis of children diagnosed with hemophagocytic lymphohistiocytosis in the PICU. Pediatr Crit Care Med 2019 Apr;20(4):e185-e190. [doi: 10.1097/PCC.0000000000001827] [Medline: 30520798]
19. Dao D, Xoay TD, Galeano BK, Phuc PH, Ouellette Y. Etiologies and clinical outcomes of patients with secondary hemophagocytic lymphohistiocytosis at a tertiary PICU. Pediatr Crit Care Med 2019 Jul;20(7):e311-e318. [doi: 10.1097/PCC.0000000000001980] [Medline: 31149968]
20. Henter JI, Samuelsson-Horne A, Aricò M, et al. Treatment of hemophagocytic lymphohistiocytosis with HLH-94 immunochemotherapy and bone marrow transplantation. Blood 2002 Oct 1;100(7):2367-2373. [doi: 10.1182/blood-2002-01-0172] [Medline: 12239144]
21. Taylor ML, Hoyt KJ, Han J, et al. An evidence-based guideline improves outcomes for patients with hemophagocytic lymphohistiocytosis and macrophage activation syndrome. J Rheumatol 2022 Sep;49(9):1042-1051. [doi: 10.3899/jrheum.211219] [Medline: 35840156]
22. Guo LL, Morse KE, Aftandilian C, et al. Characterizing the limitations of using diagnosis codes in the context of machine learning for healthcare. BMC Med Inform Decis Mak 2024 Feb 14;24(1):51. [doi: 10.1186/s12911-024-02449-8] [Medline: 38355486]

XSL•FO
RenderX

Abbreviations

EHR: electronic health record
HLH: hemophagocytic lymphohistiocytosis
***ICD-10*:** *International Statistical Classification of Diseases, Tenth Revision*
SEDAR: SickKids Enterprise-wide Data in Azure Repository
SickKids: The Hospital for Sick Children

Edited by M Balcarras; submitted 07.Nov.2025; peer-reviewed by A Ramzi, RG Sussman, X Ding; revised version received 11.Feb.2026; accepted 03.Mar.2026; published 26.Mar.2026.

Please cite as:
Ocak S, Yi M, Wolochacz A, Mehrdadi I, Naqvi A, Gupta S, Sung L, Yan AP
Identifying Hemophagocytic Lymphohistiocytosis and Describing Outcomes Using Computable Phenotypes: Retrospective Cohort Study
JMIR Cancer 2026;12:e87347
URL: https://cancer.jmir.org/2026/1/e87347
doi:10.2196/87347

© Suheyla Ocak, Martin Yi, Agata Wolochacz, Ida Mehrdadi, Ahmed Naqvi, Sumit Gupta, Lillian Sung, Adam P Yan. Originally published in JMIR Cancer (https://cancer.jmir.org), 26.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predicting Overall Survival in Patients With Multiple Primary Lung Cancer: Nomogram Development and Validation Study

Wenzhi Luo[1*], MMed; Kengliang Rao[1*], MMed; Hongjia Chen[1*], MBBS; Cailian Hu[1], MBBS; Shengming Liu[1], MD; Jing Wang[2*], MMed; Dongdong Zhang[3*], MMed; Li Chen[1*], MD

[1]Department of Pulmonary and Critical Care Medicine, First Affiliated Hospital of Jinan University, 613 Huangpu Avenue West, Tianhe District, Guangzhou, Guangdong, China

[2]Department of Pulmonary and Critical Care Medicine, Meizhou People's Hospital, Meizhou, Guangdong, China

[3]Department of Thoracic Surgery, The First Affiliated Hospital of Jinan University, Guangzhou, Guangdong, China

[*]these authors contributed equally

Corresponding Author:
Li Chen, MD
Department of Pulmonary and Critical Care Medicine, First Affiliated Hospital of Jinan University, 613 Huangpu Avenue West, Tianhe District, Guangzhou, Guangdong, China

Abstract

Background: With the rapid development of medical technology and the emphasis on early lung cancer screening, the detection rate of multiple primary lung cancer (MPLC) has increased in recent years. However, the prognostic determinants and clinical characteristics of patients with MPLC remain poorly characterized.

Objective: This study aimed to develop and validate a nomogram for predicting overall survival (OS) in patients with MPLC using data from the Surveillance, Epidemiology, and End Results database.

Methods: This study was reported in accordance with the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines. A cohort of 4177 patients with MPLC (2007 - 2015) was obtained from the Surveillance, Epidemiology, and End Results database. The patients were randomly divided into training (n=2923) and validation (n=1254) cohorts at a 7:3 ratio. Backward stepwise Cox regression identified 11 independent risk factors, which were integrated into a nomogram predicting 3-, 5-, and 8-year OS rates.

Results: The nomogram demonstrated superior discriminative ability compared to the American Joint Committee on Cancer staging system, with higher area under the receiver operating characteristic curve values for 3-, 5-, and 8-year OS predictions in both cohorts (training cohort: 0.743, 0.751, and 0.759, respectively; validation cohort: 0.737, 0.734, and 0.695, respectively). Calibration curves and decision curve analysis confirmed its clinical utility.

Conclusions: This study establishes a validated nomogram incorporating clinical and socioeconomic variables to optimize prognostic assessment and personalized treatment planning for patients with MPLC.

(JMIR Cancer 2026;12:e87275) doi:10.2196/87275

KEYWORDS

multiple primary lung cancer; Surveillance, Epidemiology, and End Results database; SEER database; nomogram; survival prediction; overall survival

Introduction

Lung cancer remains a leading cause of cancer-related mortality worldwide [1,2]. Multiple primary lung cancer (MPLC), defined as the presence of 2 or more distinct primary lung tumors either synchronously or metachronously, accounts for 3.5% to 15% of newly diagnosed cases [3-6]. With advancements in imaging and an emphasis on early lung cancer screening, the incidence rate of MPLC is increasing worldwide.

MPLC can be classified into metachronous MPLC and synchronous MPLC according to the time of lesion occurrence [7-9]. While advancements in imaging modalities have improved MPLC detection, differentiating it from intrapulmonary metastasis remains clinically challenging, particularly in histologically similar lesions [5-7,10]. Notably, patients with MPLC exhibit significantly better survival outcomes than those with intrapulmonary metastasis, with reported 5-year postresection survival rates ranging from 77.3% to 87% [6,11-17]. Moreover, the prognosis of surgically treated MPLC has been reported to be comparable to that of stage-matched solitary lung cancer [18,19]. Therefore, accurate diagnosis and prognostic assessment of MPLC are of substantial clinical importance.

XSL•FO
RenderX

The TNM staging system plays an important role in treatment selection and prognostic evaluation for patients with lung cancer. However, substantial survival heterogeneity still exists among patients with similar histological subtypes and TNM stages, particularly in MPLC [20]. In MPLC, each lesion is evaluated separately for the tumor, node, and metastasis categories according to current staging recommendations [21]. Furthermore, staging strategies differ among various forms of multifocal lung cancer and among MPLC cases with different radiological manifestations [21]. Thus, patients with MPLC may be inadvertently upstaged to T3 and/or T4 or even M1a, leading to erroneous staging, inaccurate prognosis, and improper treatment [20,22].

Several prognostic nomograms for MPLC have been developed using Surveillance, Epidemiology, and End Results (SEER)–based cohorts [15,23]. However, most existing models primarily focus on clinicopathological variables and provide limited evaluation of socioeconomic factors or the distinct prognostic contributions of the first primary lung cancer (FPLC) and second primary lung cancer (SPLC). In addition, few studies have established long-term survival prediction models for patients with MPLC, and the prognostic applicability of the conventional TNM and American Joint Committee on Cancer (AJCC) staging systems in this population remains limited.

Therefore, this study aimed to identify independent prognostic factors associated with overall survival (OS) in patients with MPLC using the SEER database and develop and validate a prognostic nomogram for predicting 3-, 5-, and 8-year OS. Furthermore, the predictive performance of the nomogram was compared with that of the AJCC staging system to evaluate its potential clinical utility in individualized prognostic assessment and risk stratification for patients with MPLC.

Methods

Ethical Considerations

This study analyzed publicly available, deidentified data from the SEER database. The use of this dataset is exempt from institutional review board approval, and the requirement for informed consent was waived for this analysis. The study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. This study was reported in accordance with the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines.

Patient Selection and Data Processing

From 465,552 lung cancer cases in the SEER database (2007 - 2015), a total of 8354 (1.8%) tumor records corresponding to 4177 unique patients were identified as meeting the 2013 American College of Chest Physicians criteria for MPLC. These 4177 patients were included in the final analysis. Exclusion criteria encompassed age below 18 years, diagnosis via autopsy or death certificate, and incomplete variables. Patients with MPLC were identified using *International Classification of Diseases for Oncology, Third Edition*, site codes for lung and bronchus tumors (C33.9 and C34.0-C34.9). The SEER sequence number variable was used to distinguish FPLC from SPLC, with sequence numbers 01 and 02 representing the first and second primary lung tumors, respectively. Patients with sequence numbers of 03 or higher were excluded. Synchronous MPLC and metachronous MPLC were defined according to a diagnostic interval of 6 months or less and more than 6 months, respectively. To minimize potential misclassification of intrapulmonary metastasis, cases with identical histological subtypes were conservatively excluded if tumors were located within the same lobe, accompanied by N1 or N2 disease, or associated with distant metastasis (M1). Eligible patients were randomized into training (70%) and validation (30%) cohorts. The following variables were evaluated: age, sex, AJCC stage, FPLC and SPLC histology, SPLC tumor size, FPLC and SPLC surgery status, average household income, marital status, race, survival status, and survival time. Age (<65 vs ≥65 years), SPLC tumor size (≤3, 3 - 5, and >5 cm), and household income (<US $50,000, US $50,000 - 80,000, and >US $80,000) were categorized according to the SEER database records for model construction. All TNM stages were uniformly converted to the Eighth Edition of the *AJCC Cancer Staging Manual* using the standard mapping algorithm provided by the SEER database.

Construction and Validation of the Nomogram

To improve the prediction power of the nomogram, 4177 patients were randomly assigned to the training cohort (n=2923) and the validation cohort (n=1254) at a ratio of 7:3. The prognostic nomogram in this study was developed using the Cox proportional hazards regression model, a conventional statistical method. No machine learning, artificial intelligence, or deep learning algorithms were used for model construction, data analysis, or outcome prediction. Multivariate Cox proportional hazards regression analysis was performed to identify variables that affect OS in the training cohort. Using these predictors of OS, we constructed a nomogram for predicting 3-, 5-, and 8-year survival rates in patients with MPLC.

We subsequently evaluated the discrimination ability of the nomogram by calculating the index of concordance (C-index) and the area under the receiver operating characteristic curve (AUC). The calibration curves show the associations between the predicted and actual probabilities. Both discrimination and calibration were evaluated via bootstrapping with 1000 resamples. The net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were used to compare the accuracy of the nomogram with that of the traditional AJCC staging model. Decision curve analysis (DCA) was used to assess the actual clinical benefits. Finally, AUC values and Kaplan-Meier survival curves were used to compare the nomogram with AJCC staging alone.

Statistical Analysis

The primary end point was OS, defined as the time from diagnosis of the FPLC to death from any cause or last follow-up. All the statistical analyses were performed via SPSS (version 27.0; IBM Corp) and R (version 4.0.3; R Foundation for Statistical Computing). The statistical tests were 2 sided, and a *P* value of less than .05 was considered to indicate statistical significance.

XSL•FO
RenderX

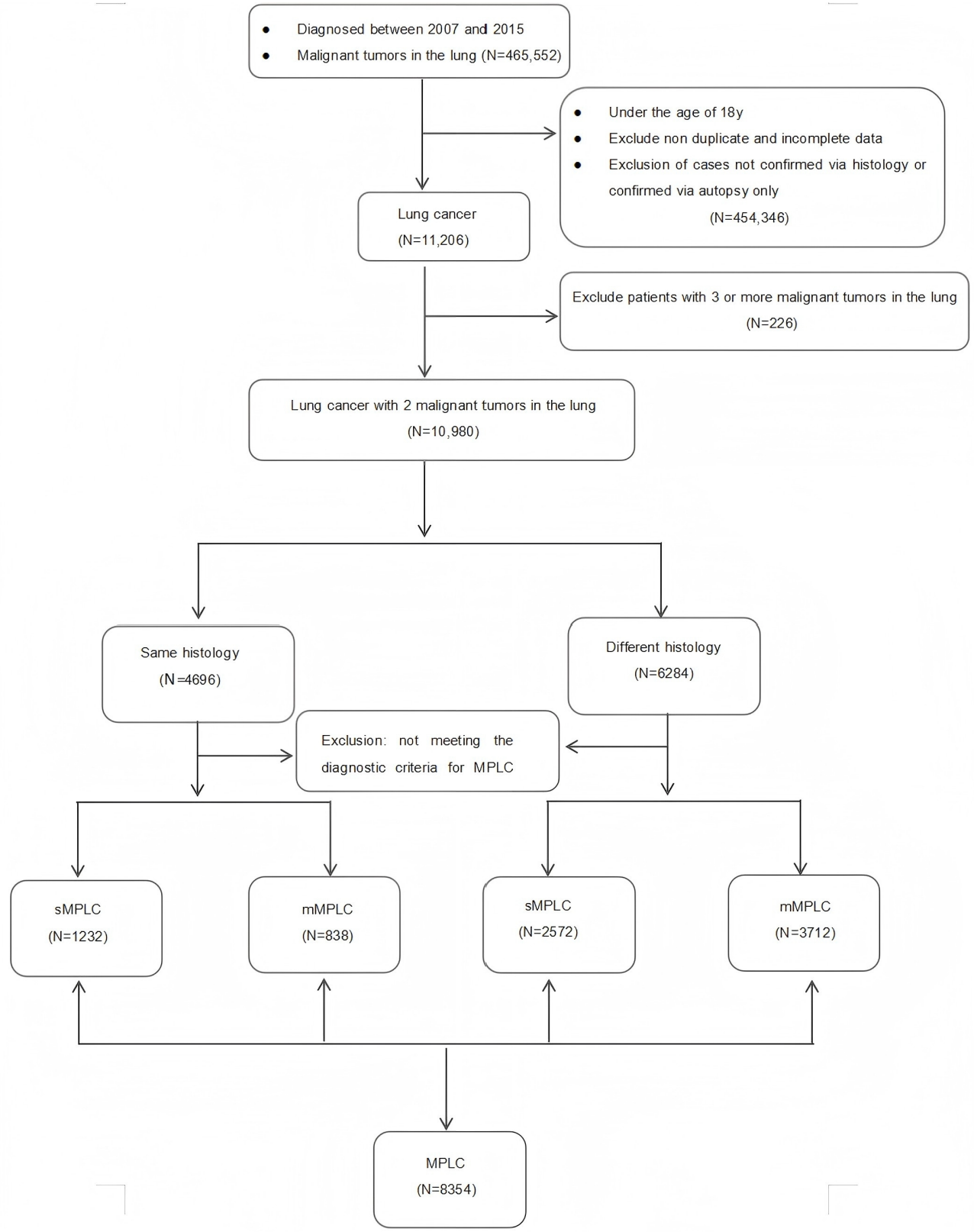
Results

Baseline Characteristics of the Study Participants

We evaluated a total of 4177 patients with MPLC who were randomly divided into a training cohort (n=2923, 70%) and a validation cohort (n=1254, 30%). The selection process of patients with MPLC is shown in Figure 1. The baseline characteristics of the patients in the training and validation cohorts are presented in Multimedia Appendix 1. In the training cohort, most patients (2277/2923, 77.9%) were aged 65 years or older, and 54% (1578/2923) were female. AJCC stage 1 was the most common stage (1459/2923, 49.9%), followed by stage 3 (775/2923, 26.5%). Adenocarcinoma was the predominant histological subtype in both FPLC (1696/2923, 58%) and SPLC (1681/2923, 57.5%). In addition, 61.4% (1796/2923) of patients presented with bilateral lesions. The surgical resection rate was lower for SPLC (1566/2923, 53.6%) than for FPLC (2147/2923, 73.5%). Detailed demographic and clinicopathological characteristics are summarized in Multimedia Appendix 1. During the follow-up period, a total of 2354 deaths were observed, with a maximum follow-up duration of 119 months. For the entire cohort of 4177 patients, the median OS was 68 (95% CI 66 - 71) months. The estimated 3-, 5-, and 8-year OS rates were 70.8% (95% CI 69.4% - 72.2%), 55.3% (95% CI 53.7% - 56.9%), and 32.6% (95% CI 30.8% - 34.4%), respectively.

XSL•FO
RenderX

Figure 1. Flow diagram describing the evaluation and selection of patients with multiple primary lung cancer (MPLC). The final cohort consisted of 8354 tumor records representing 4177 unique patients. mMPLC: metachronous MPLC; sMPLC: synchronous MPLC.



Univariate and Multivariate Analyses of OS in the Training Cohort

Through univariate and multivariate Cox proportional hazards regression, we identified 11 independent risk factors in the training cohort as follows: age, sex, AJCC stage, histology of FPLC and SPLC, tumor size of SPLC, surgery of both FPLC and SPLC, average household income, marital status, and race, as shown in Table 1.

XSL•FO
RenderX

Table . Univariable and multivariable analyses of overall survival in the training cohort.

Variable	Univariate analysis		Multivariate analysis	
	HR[a] (95% CI)	*P* value	HR (95% CI)	*P* value
Age (y)				
<65	Reference	Reference	Reference	Reference
≥65	1.467 (1.295-1.662)	<.001	1.416 (1.245-1.610)	<.001
Sex				
Male	Reference	Reference	Reference	Reference
Female	0.674 (0.612-0.743)	<.001	0.735 (0.662-0.815)	<.001
AJCC[b] stage				
1	Reference	Reference	Reference	Reference
2	1.393 (1.176-1.650)	<.001	1.218 (1.025 - 1.447)	.03
3	1.642 (1.459 - 1.847)	<.001	1.353 (1.196 - 1.530)	<.001
4	2.712 (2.370 - 3.104)	<.001	2.132 (1.850 - 2.456)	<.001
Histology of FPLC[c]				
Squamous cell carcinoma	Reference	Reference	Reference	Reference
Adenocarcinoma	0.653 (0.587 - 0.726)	<.001	0.873 (0.780 - 0.976)	.02
SCLC[d]	1.123 (0.928 - 1.358)	.23	1.080 (0.888 - 1.313)	.44
Others	0.775 (0.613 - 0.980)	.03	0.951 (0.751 - 1.206)	.68
Histology of SPLC[e]				
Squamous cell carcinoma	Reference	Reference	Reference	Reference
Adenocarcinoma	0.623 (0.557 - 0.697)	<.001	0.830 (0.737 - 0.936)	.002
SCLC	1.745 (1.503 - 2.028)	<.001	1.618 (1.386 - 1.888)	<.001
Others	0.968 (0.755 - 1.242)	.80	1.150 (0.893 - 1.482)	.28
Nodule size of FPLC (cm)				
≤3	Reference	Reference	—[f]	—
3-5	1.246 (1.108 - 1.402)	<.001	—	—
>5	1.367 (1.168 - 1.601)	<.001	—	—
Nodule size of SPLC (cm)				
≤3	Reference	Reference	Reference	Reference
3-5	1.989 (1.753 - 2.256)	<.001	1.630 (1.432 - 1.854)	<.001
>5	2.641 (2.245 - 3.107)	<.001	1.983 (1.675 - 2.347)	<.001
Surgery of FPLC				
No	Reference	Reference	Reference	Reference
Yes	0.568 (0.511 - 0.632)	<.001	0.869 (0.770 - 0.980)	.02
Surgery of SPLC				
No	Reference	Reference	Reference	Reference
Yes	0.453 (0.410 - 0.500)	<.001	0.633 (0.564 - 0.711)	<.001
Income (US $)				
<50,000	Reference	Reference	Reference	Reference
50,000-80,000	0.730 (0.657 - 0.810)	<.001	0.810 (0.727 - 0.902)	<.001
>80,000	0.699 (0.585 - 0.836)	<.001	0.835 (0.695 - 1.002)	.05
Marital status				

XSL•FO
RenderX

Variable	Univariate analysis		Multivariate analysis	
	HR[a] (95% CI)	*P* value	HR (95% CI)	*P* value
Married	Reference	Reference	Reference	Reference
Divorced or separated	1.108 (0.963 - 1.275)	.15	1.147 (0.994 - 1.324)	.06
Other	1.157 (1.037 - 1.291)	.009	1.142 (1.017 - 1.281)	.02
Race				
White	Reference	Reference	Reference	Reference
Black	1.056 (0.889 - 1.254)	.54	0.944 (0.792 - 1.124)	.52
Others	0.707 (0.554 - 0.901)	.005	0.756 (0.591 - 0.968)	.03

[a]HR: hazard ratio.

[b]AJCC: American Joint Committee on Cancer.

[c]FPLC: first primary lung cancer.

[d]SCLC: small-cell lung cancer.

[e]SPLC: second primary lung cancer.

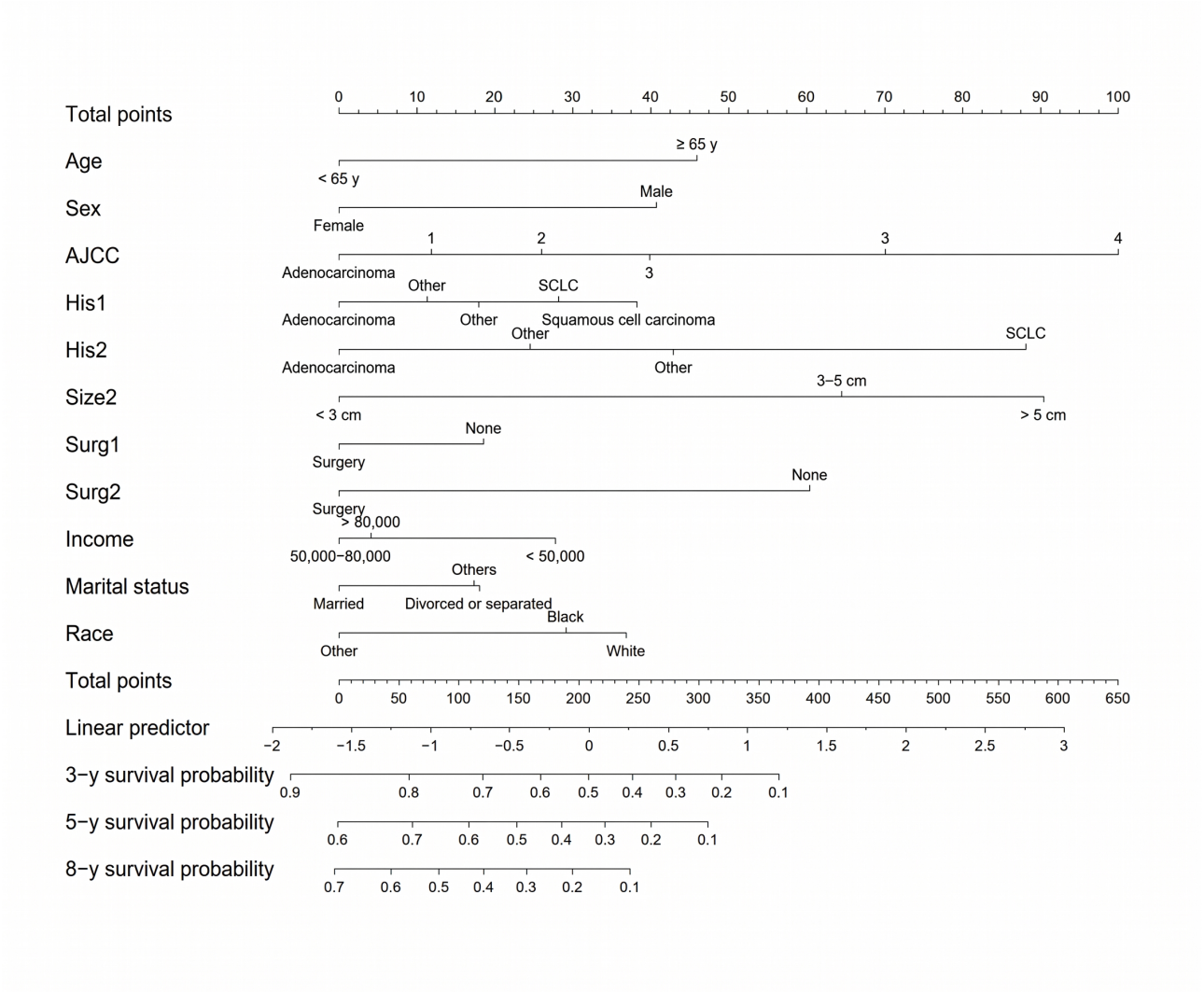
[f]Variables excluded from the multivariate analysis.

Construction and Validation of the Nomogram

A nomogram was constructed for prediction based on the 11 independent risk factors, which were statistically significant prognostic factors in the Cox proportional hazards model. The nomogram revealed that the AJCC stage had the greatest contribution to survival outcomes, followed by histology, tumor size, and SPLC surgery. We obtained a total score by adding the total points for the score assigned to each variable. The total score corresponded to the prediction of 3-, 5-, and 8-year OS for each patient with MPLC. The higher the score, the worse the survival prognosis. The nomogram is shown in Figure 2.

XSL•FO
RenderX

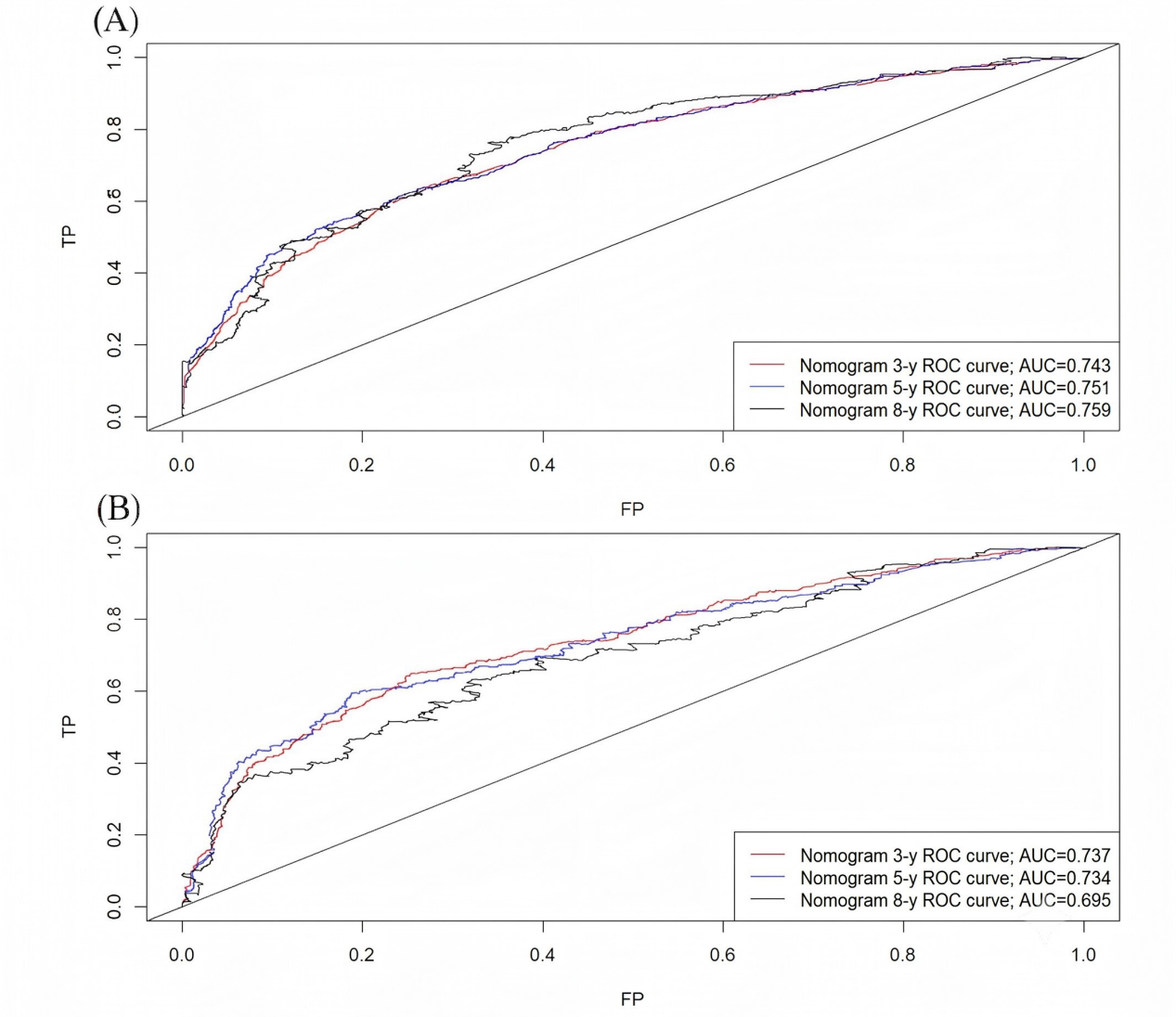
Figure 2. Nomogram predicting 3-, 5-, and 8-year overall survival of patients with multiple primary lung cancer. The factors of age, sex, American Joint Committee on Cancer (AJCC) stage, histology of first primary lung cancer (FPLC; His1) and second primary lung cancer (SPLC; His2), tumor size of SPLC (Size2), surgery for both FPLC (Surg1) and SPLC (Surg2), average household income, marital status, and race were included in the model. SCLC: small-cell lung cancer.



The C-index of the nomogram in the training cohort was 0.702, whereas it was 0.699 in the validation cohort. The AUCs used to predict 3-, 5-, and 8-year OS in the training cohort were 0.743, 0.751, and 0.759, respectively (Figure 3A). In the validation cohort, these values were 0.737, 0.734, and 0.695, respectively (Figure 3B). Furthermore, the calibration curves of the nomogram demonstrated a favorable fit for the 3-, 5-, and 8-year OS probabilities in both the training (Figure 4A-C) and validation (Figure 4D-F) cohorts. There was good consistency between the predicted values and the actual observed values. These findings demonstrated the high accuracy and discriminatory capacity of the nomogram.

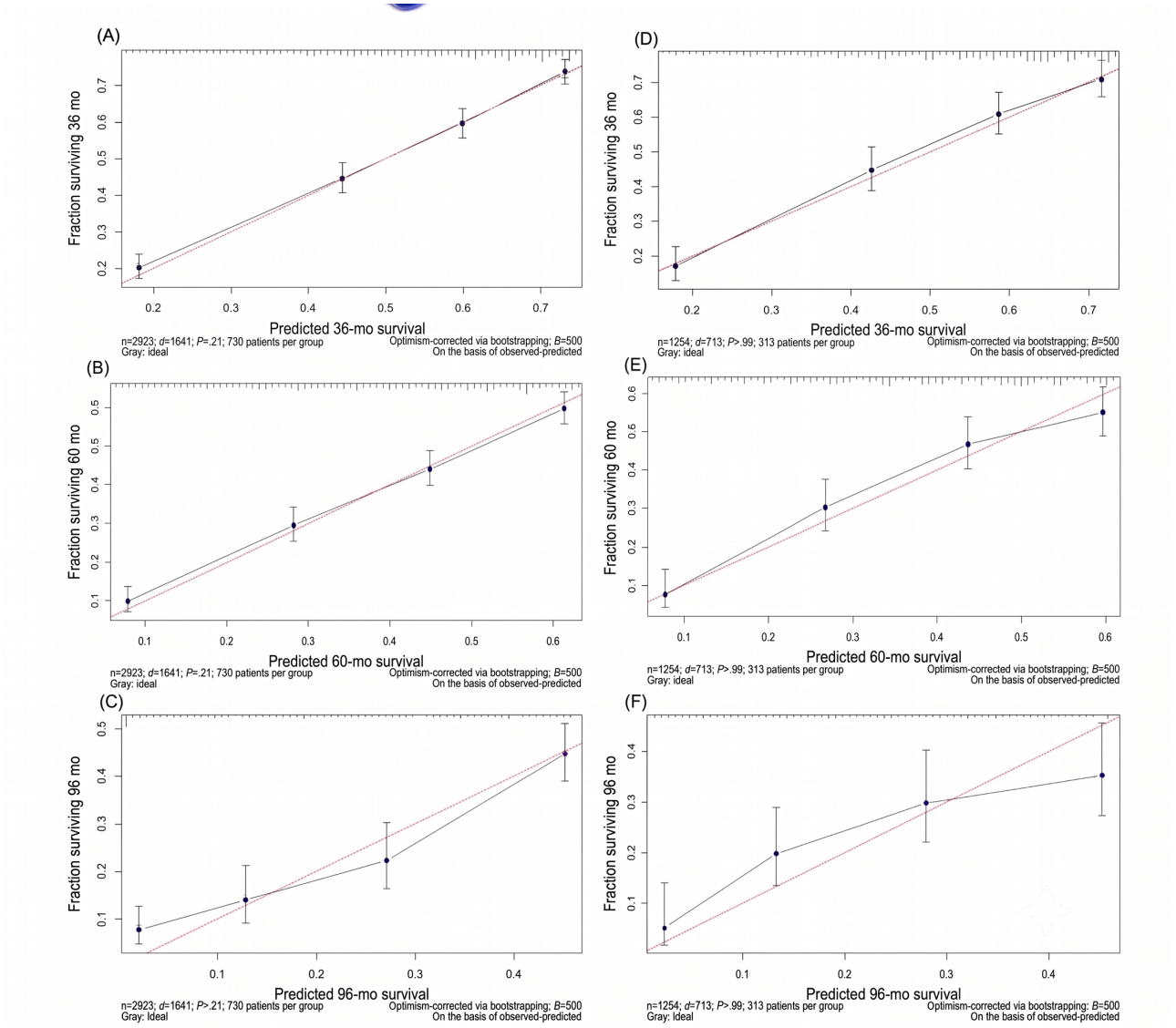
XSL•FO
RenderX

Figure 3. Receiver operating characteristic (ROC) curves of the nomogram for 3-, 5-, and 8-year overall survival prediction: (A) training cohort and (B) validation cohort. AUC: area under the ROC curve; FP: false positive; TP: true positive.



XSL•FO
RenderX

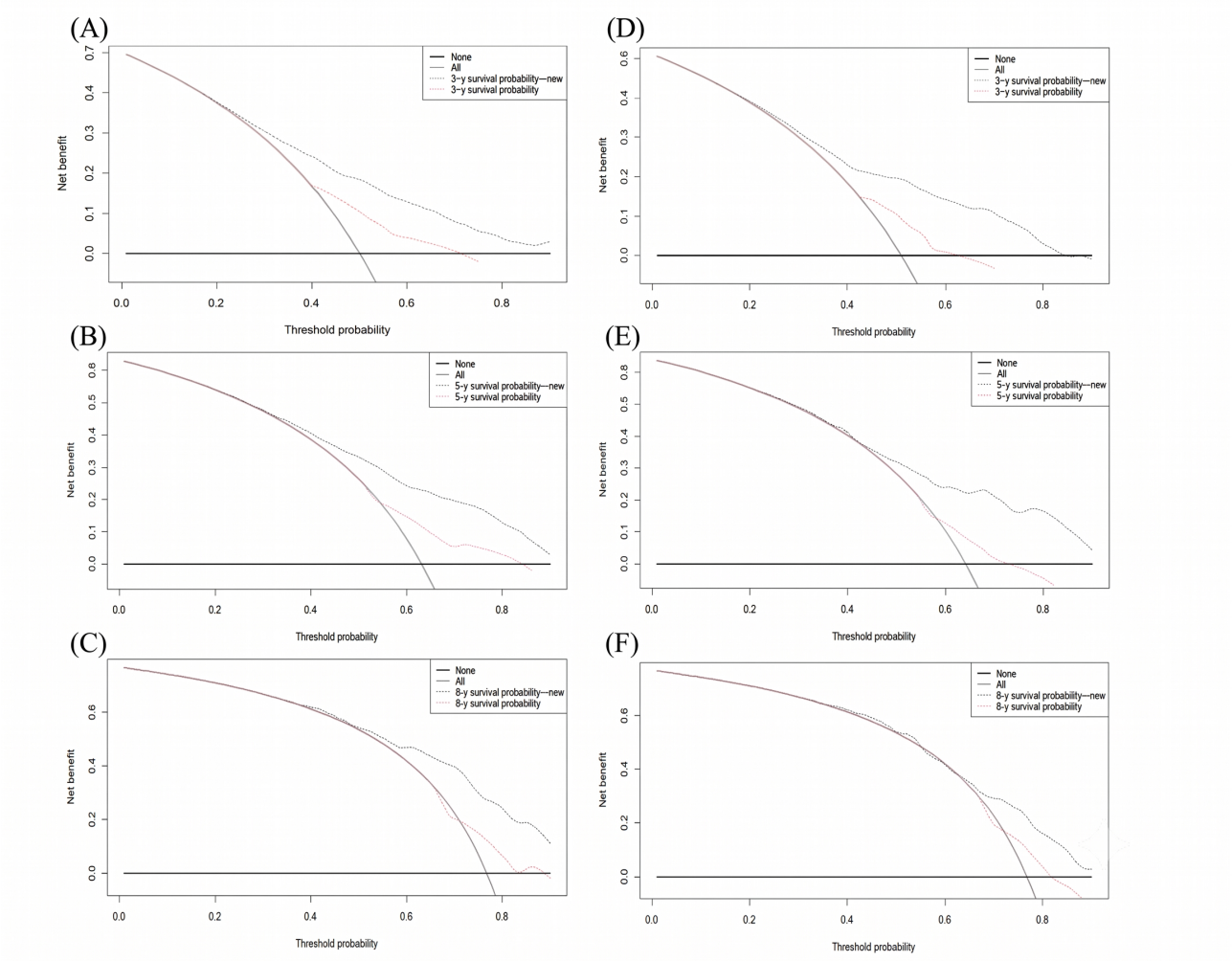
Figure 4. The calibration curves for predicting patient overall survival at 3, 5, and 8 years in the training cohort (A-C) and validation cohort (D-F).



In addition, the NRI values at 3, 5, and 8 years in the training cohort were 0.590 (95% CI 0.505 - 0.664), 0.576 (95% CI 0.522 - 0.688), and 0.577 (95% CI 0.467 - 0.762), respectively. The IDI values for 3-, 5-, and 8-year survival were 0.097, 0.091, and 0.070, respectively, in the training cohort and 0.102, 0.090, and 0.066, respectively, in the validation cohort. All the IDI values were greater than 0 (P<.001), indicating a significantly improved prediction performance of the nomogram compared with the AJCC stage model. In both the training cohort and the validation cohort, DCA strongly improved the prediction of 3-, 5-, and 8-year OS (Figure 5).

XSL•FO
RenderX

Figure 5. Decision curve analysis for the nomogram and American Joint Committee on Cancer stage of multiple primary lung cancer at 3, 5, and 8 years in the training cohort (A-C) and validation cohort (D-F).

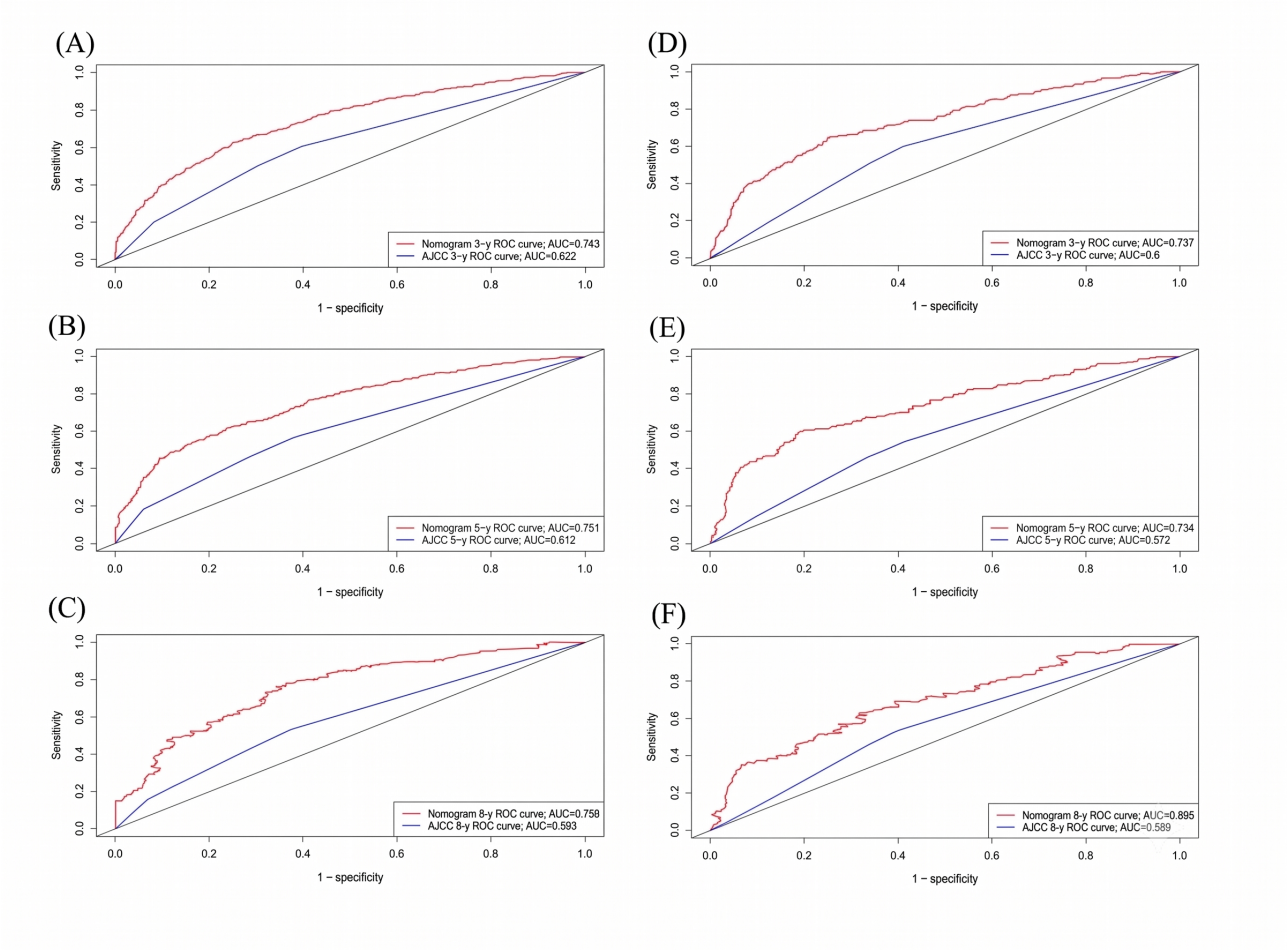


Comparison of the Predictive Accuracy Between the Novel Risk Stratification Model and the Conventional Prognostic Model

Using AUC values, we compared the accuracy between the nomogram and traditional AJCC staging system. The AUC values of the nomogram and the traditional prediction model were 0.743 vs 0.622, 0.751 vs 0.612, and 0.759 vs 0.583, respectively, in the training cohort (Figure 6A-C) and 0.737 vs 0.60, 0.734 vs 0.572, and 0.695 vs 0.569, respectively, in the validation cohort (Figure 6D-F). We observed that the nomogram had higher AUC values than the AJCC staging model regardless of the prediction of 3-, 5-, and 8-year OS in both the training and validation cohorts.

XSL•FO
RenderX

Figure 6. Receiver operating characteristic (ROC) curves of the nomogram and American Joint Committee on Cancer (AJCC) stage prognostic system for 3-, 5-, and 8-year overall survival prediction in the training (A-C) and validation (D-F) cohorts. AUC: area under the ROC curve.



Discussion

Principal Findings

In recent years, the survival prognosis of patients with MPLC has garnered significant research interest. Current evidence indicates that patients with MPLC exhibit substantially improved OS compared to those with intrapulmonary metastasis [5,24,25]. Notably, female patients demonstrate a markedly elevated MPLC incidence (54.2%-73.2%) relative to SPLC (27.9%) or intrapulmonary metastasis (31.6%) [5,20,24], a finding corroborated by our study. However, studies on the prognosis of MPLC are mostly based on small sample sizes, and some important factors affecting the prognosis of patients with tumors are ignored [4-6,12].

In this retrospective study, we developed a Cox proportional hazards regression model and prognostic nomogram for patients with MPLC. Our analysis identified 4 key independent risk factors: AJCC stage, histological subtype of SPLC, secondary tumor size, and surgical management of SPLC. Additionally, female sex, higher socioeconomic status, and race emerged as favorable prognostic indicators. On the basis of these factors, a nomogram predicting the probability of survival calibrated well and was a good measure of discrimination.

Compared with previously published SEER-based prognostic nomograms for MPLC, our model incorporated both clinicopathological and socioeconomic variables and separately evaluated the histological characteristics of FPLC and SPLC. In addition, this nomogram provided simultaneous prediction of 3-, 5-, and 8-year OS and demonstrated improved predictive performance compared with the conventional AJCC staging system. These findings suggest that the model may offer additional value for individualized prognostic assessment in patients with MPLC.

While the AJCC staging system partially predicts lung cancer outcomes, its prognostic utility in MPLC requires refinement. A prior study by Tanvetyanon et al [15] highlighted sex, younger age, smaller tumor size, absence of nodal involvement, and bilateral presentation as significant prognosticators in MPLC. Our findings align with these observations and further establish AJCC stage as the most influential prognostic determinant [26]. Notably, Yu et al [18] identified tumor size as the sole prognostic factor in node-negative patients with MPLC, whereas subsequent studies [15,27-29] confirmed that survival outcomes in early-stage MPLC (T1-T2) parallel those of solitary lung nodules, significantly surpassing stage 3 and 4 non–small-cell lung cancer outcomes based on the Eighth Edition of the *AJCC Cancer Staging Manual*.

Histological characteristics, particularly those of SPLC, significantly impacted survival outcomes. Adenocarcinoma histology was independently correlated with improved prognosis, consistent with prior reports [15,30]. However,

XSL•FO
RenderX

neither the aforementioned study by Tanvetyanon et al [15] nor a study by Song et al [23] determined that the histology from FPLC or SPLC had a greater impact on the prognosis of MPLC. We conducted an independent analysis of the histology of 2 primary lung cancers in MPLC. The nomogram revealed that the contribution of histological variables to SPLC was significantly greater than that to FPLC. The survival prognosis of patients with MPLC was related mainly to the histological type of SPLC.

In clinical evaluation, distinguishing between synchronous and metachronous MPLC is of great significance. Although the SEER database does not provide variables that can be used to clearly define the sequence of onset, the proportion of bilateral lesions in the training cohort of this study was as high as 61.4% (1796/2923), suggesting that this patient population was predominantly characterized by synchronous or near-synchronous onset. Regarding pathological consistency, the model in this study included the histological types of FPLC and SPLC as 2 independent variables in the analysis rather than simply combining them into a binary "pathologically consistent/inconsistent" indicator. This design allows the line plot to more fully reflect the prognostic value of the corresponding histological types of the 2 tumors, particularly highlighting the importance of the histological subtype of SPLC for prognosis without being confounded by whether the histologies of the 2 primary cancers were consistent. In the future, it will be necessary to further validate the results of this study in different subgroups based on more detailed information regarding the sequence of onset.

Surgical intervention for SPLC conferred superior prognostic benefits compared to FPLC management. Cox regression analysis revealed substantially worse outcomes in nonsurgical cohorts, potentially attributable to misdiagnosis as metastatic or recurrent disease leading to therapeutic delays. Current consensus emphasizes surgical resection as the main approach to MPLC management [3], supported by evidence of enhanced long-term survival vs stage 4 non–small-cell lung cancer or metastatic disease [30,31]. The research by Stella et al [31] further advocates for aggressive surgical approaches for eligible patients with adequate cardiopulmonary reserve in MPLC or multiple lesions. Moreover, the differences in imaging assessment, inconsistencies in pathological diagnostic criteria, interobserver variability, variability in tissue sampling methods, lack of standardized definitions for distinguishing MPLC from intrapulmonary metastasis, and differences in diagnostic workup across institutions or time periods could bias the diagnosis of MPLC considering that current guidelines focus mainly on the recurrence of intrapulmonary metastasis rather than the development of MPLC. Establishing a risk profile for the survival prognosis of patients with MPLC would be beneficial.

In our study, being older; male; from a low-income background; divorced, separated, or single; or White were poor prognostic factors for MPLC. MPLC is common in patients aged 65 years and older, and the prognosis of patients with MPLC becomes worse with age, similar to most malignant tumors. Although there were more female patients than male patients in this study, the risk of death in female patients is lower than in male patients, which is due to the emotional and psychological stress and hormone levels of female patients. In terms of economic value, low-income individuals do not pay attention to the disease and cannot afford medical expenses, which is an important reason for the high risk of death of patients with MPLC among low-income individuals. Married patients may have better survival outcomes than other patients. However, with respect to the impact of race on the prognosis of MPLC, because the information was collected from the SEER database, White patients comprised most of the data, and more real-world data are needed to support the conclusion that White individuals have a high incidence of MPLC.

According to the results of this study, the nomogram demonstrated good distinguishing ability and high accuracy, with training and validation cohort C-indexes of 0.702 and 0.699, respectively. Time-dependent AUC values exceeded 0.70 for 3-, 5-, and 8-year survival predictions. The calibration curve was close to the reference line and fell along the diagonal line of 45 degrees, which shows that the nomogram was in good agreement with the actual results and that the degree of fit was high. The NRI and IDI values were greater than 0, indicating that the nomogram had improved prediction performance over traditional AJCC staging. The DCA curve demonstrated that the nomogram yielded a high net benefit compared to alternative strategies in both the training and validation cohorts, indicating that it meets the practical needs of clinical decision-making and can be widely used in clinical analysis. This nomogram provides continuous estimates of survival probability rather than assigning patients into predefined risk categories. Preserving continuous probability estimation avoids the loss of prognostic information associated with arbitrary cutoff selection and enables more individualized risk assessment. The continuous scores generated by the model may assist clinicians in personalized follow-up surveillance, prognostic evaluation, and clinical risk stratification for patients with MPLC. For example, patients with progressively increasing model-derived risk scores may benefit from more intensive follow-up monitoring, thereby avoiding the limitations of the traditional "one-size-fits-all" approach inherent in coarse risk stratification systems. Therefore, we frame this nomogram as a clinical decision support tool for individualized prognostic assessment and risk stratification rather than as a stand-alone prognostic classifier.

Limitations

This study has certain limitations, and caution should be exercised when interpreting and applying its findings. First, as this was a retrospective analysis based on the SEER database, some clinical information was incomplete, which may introduce selection bias; key prognostic factors such as smoking, alcohol consumption, family history of cancer, genotyping, detailed imaging characteristics, and laboratory test results were missing, making it impossible to adjust for these as confounding factors in the analysis. Second, due to limitations in the database itself, treatment details were limited. Data on specific regimens for radiotherapy, chemotherapy, targeted therapy, and immunotherapy were lacking, making it difficult to accurately assess the impact of different treatments on patient prognosis. Third, this study included only cases of dual primary lung cancer and did not include cases with 3 or more primary tumors; therefore, the conclusions cannot be extrapolated to a broader

XSL•FO
RenderX

population of patients with MPLC. Fourth, the constructed nomogram has not yet undergone external validation in independent cohorts from other medical institutions; its generalizability across different populations and clinical settings requires further verification. Fifth, this study used OS as the primary end point rather than tumor-specific survival; non–cancer-related deaths may introduce confounding factors into the survival analysis results. Additionally, the model's C-index was approximately 0.70, indicating moderate but clinically meaningful discriminatory ability rather than perfect prediction accuracy. Therefore, this nomogram should be considered a clinical decision support tool for individualized prognostic assessment and risk stratification rather than an independent prognostic classifier. In summary, the model may assist clinicians in developing individualized follow-up and treatment strategies, but it should not be used alone for final clinical decision-making.

Conclusions

In summary, we constructed a nomogram for patients with MPLC using the SEER database that has good ability to predict the individual survival of patients with MPLC more efficiently, accurately, and easily. This information can help clinicians evaluate the prognosis of patients better than before and provide individualized treatment.

Acknowledgments

During the preparation of this manuscript, the authors used ChatGPT for minor linguistic polishing to improve sentence expression, language fluency and overall readability. After this language refinement work, the authors thoroughly reviewed, revised and edited the entire manuscript content as needed, and take full responsibility for all academic content, analysis, arguments and conclusions presented in this publication.

Funding

This research was supported by Science and Technology Projects in Guangzhou (No. 2025A03J4224).

Data Availability

The datasets generated or analyzed during this study are publicly available in the Surveillance, Epidemiology, and End Results repository [32].

Authors' Contributions

Conceptualization: WL, KR, HC
Data curation: HC, LC, CH
Methodology: WL, KR, HC, SL
Writing—original draft: WL, KR, HC, CH, JW
Writing—review and editing: JW, DZ, LC, SL
WL, KR, and HC contributed equally to the study design. JW, DZ, and LC contributed equally to the manuscript review and editing and are co-corresponding authors for this paper. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Baseline characteristics of the patients in the training and validation cohorts.
[PDF File, 80 KB - cancer_v12i1e87275_app1.pdf]

References

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin 2025;75(1):10-45. [doi: 10.3322/caac.21871] [Medline: 39817679]
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
3. Huang Y, Zhang HM, Cai HR, Hu JF. Advances in the diagnosis and treatment of multiple primary lung cancers. Asian J Surg 2024 Jun;47(6):2635-2636. [doi: 10.1016/j.asjsur.2024.03.022] [Medline: 38548539]
4. Murphy SJ, Harris FR, Kosari F, et al. Using genomics to differentiate multiple primaries from metastatic lung cancer. J Thorac Oncol 2019 Sep;14(9):1567-1582. [doi: 10.1016/j.jtho.2019.05.008] [Medline: 31103780]
5. Liu Z, Wang L, Gao S, et al. Plasma metabolomics study in screening and differential diagnosis of multiple primary lung cancer. Int J Surg 2023 Mar 1;109(3):297-312. [doi: 10.1097/JS9.0000000000000006] [Medline: 36928390]

XSL•FO
RenderX

6. Adebonojo SA, Moritz DM, Danby CA. The results of modern surgical therapy for multiple primary lung cancers. Chest 1997 Sep;112(3):693-701. [doi: 10.1378/chest.112.3.693] [Medline: 9315801]
7. Martini N, Melamed MR. Multiple primary lung cancers. J Thorac Cardiovasc Surg 1975 Oct;70(4):606-612. [doi: 10.1016/S0022-5223(19)40289-4] [Medline: 170482]
8. Allgäuer M, Kluck K, Christopoulos P, et al. Advancing lung cancer staging: integrating IASLC recommendations and bioinformatics to delineate tumor origins. J Thorac Oncol 2026 Feb;21(2):267-282. [doi: 10.1016/j.jtho.2025.10.010] [Medline: 41135642]
9. Xu R, Zhang G, Che Y, et al. Ablation combined with video-assisted thoracic surgery hybrid technique for multiple primary lung cancer. iScience 2025 May 20;28(6):112703. [doi: 10.1016/j.isci.2025.112703] [Medline: 40546939]
10. Chou TY, Dacic S, Wistuba I, et al. Differentiating separate primary lung adenocarcinomas from intrapulmonary metastases with emphasis on pathological and molecular considerations: recommendations from the International Association for the Study of Lung Cancer Pathology Committee. J Thorac Oncol 2025 Mar;20(3):311-330. [doi: 10.1016/j.jtho.2024.11.016] [Medline: 39579981]
11. Jiang L, He J, Shi X, et al. Prognosis of synchronous and metachronous multiple primary lung cancers: systematic review and meta-analysis. Lung Cancer 2015 Mar;87(3):303-310. [doi: 10.1016/j.lungcan.2014.12.013] [Medline: 25617985]
12. Takamochi K, Oh S, Matsuoka J, Suzuki K. Clonality status of multifocal lung adenocarcinomas based on the mutation patterns of EGFR and K-ras. Lung Cancer 2012 Mar;75(3):313-320. [doi: 10.1016/j.lungcan.2011.08.007] [Medline: 22209037]
13. Ishikawa Y, Nakayama H, Ito H, et al. Surgical treatment for synchronous primary lung adenocarcinomas. Ann Thorac Surg 2014 Dec;98(6):1983-1988. [doi: 10.1016/j.athoracsur.2014.07.006] [Medline: 25443005]
14. Tanvetyanon T, Finley DJ, Fabian T, et al. Prognostic factors for survival after complete resections of synchronous lung cancers in multiple lobes: pooled analysis based on individual patient data. Ann Oncol 2013 Apr;24(4):889-894. [doi: 10.1093/annonc/mds495] [Medline: 23136230]
15. Tanvetyanon T, Finley DJ, Fabian T, et al. Prognostic nomogram to predict survival after surgery for synchronous multiple lung cancers in multiple lobes. J Thorac Oncol 2015 Feb;10(2):338-345. [doi: 10.1097/JTO.0000000000000400] [Medline: 25330145]
16. Wang Z, Yuan X, Sun K, et al. Optimizing the NGS-based discrimination of multiple lung cancers from the perspective of evolution. NPJ Precis Oncol 2025 Jan 14;9(1):14. [doi: 10.1038/s41698-024-00786-5] [Medline: 39809905]
17. Liu A, Wang Z, Yang Y, et al. Preoperative diagnosis of malignant pulmonary nodules in lung cancer screening with a radiomics nomogram. Cancer Commun (Lond) 2020 Jan;40(1):16-24. [doi: 10.1002/cac2.12002] [Medline: 32125097]
18. Yu YC, Hsu PK, Yeh YC, et al. Surgical results of synchronous multiple primary lung cancers: similar to the stage-matched solitary primary lung cancers? Ann Thorac Surg 2013 Dec;96(6):1966-1974. [doi: 10.1016/j.athoracsur.2013.04.142] [Medline: 24021769]
19. Dai L, Yang HL, Yan WP, et al. The equivalent efficacy of multiple operations for multiple primary lung cancer and a single operation for single primary lung cancer. J Thorac Dis 2016 May;8(5):855-861. [doi: 10.21037/jtd.2016.03.42] [Medline: 27162659]
20. Shintani Y, Okami J, Ito H, et al. Clinical features and outcomes of patients with stage I multiple primary lung cancers. Cancer Sci 2021 May;112(5):1924-1935. [doi: 10.1111/cas.14748] [Medline: 33236385]
21. Detterbeck FC, Nicholson AG, Franklin WA, et al. The IASLC Lung Cancer Staging Project: summary of proposals for revisions of the classification of lung cancers with multiple pulmonary sites of involvement in the forthcoming eighth edition of the TNM classification. J Thorac Oncol 2016 May;11(5):639-650. [doi: 10.1016/j.jtho.2016.01.024] [Medline: 26940528]
22. Eberhardt WE, Mitchell A, Crowley J, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the M descriptors in the forthcoming eighth edition of the TNM classification of lung cancer. J Thorac Oncol 2015 Nov;10(11):1515-1522. [doi: 10.1097/JTO.0000000000000673] [Medline: 26536193]
23. Song CK, Guo ZX, Shen XY, et al. Prognostic factors analysis and nomogram construction of dual primary lung cancer: a population study. Biomed Res Int 2020;2020:7206591. [doi: 10.1155/2020/7206591] [Medline: 32149127]
24. Hu C, Zhao L, Liu W, et al. Genomic profiles and their associations with TMB, PD-L1 expression, and immune cell infiltration landscapes in synchronous multiple primary lung cancers. J Immunother Cancer 2021 Dec;9(12):e003773. [doi: 10.1136/jitc-2021-003773] [Medline: 34887263]
25. Rami-Porta R, Asamura H, Travis WD, Rusch VW. Lung cancer - major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin 2017 Mar;67(2):138-155. [doi: 10.3322/caac.21390] [Medline: 28140453]
26. Tian H, Bai G, Yang Z, et al. Multiple primary lung cancer: updates of clinical management and genomic features. Front Oncol 2023;13:1034752. [doi: 10.3389/fonc.2023.1034752] [Medline: 36910635]
27. Trousse D, Barlesi F, Loundou A, et al. Synchronous multiple primary lung cancer: an increasing clinical occurrence requiring multidisciplinary management. J Thorac Cardiovasc Surg 2007 May;133(5):1193-1200. [doi: 10.1016/j.jtcvs.2007.01.012] [Medline: 17467428]

XSL•FO
RenderX

28. Lv J, Zhu D, Wang X, Shen Q, Rao Q, Zhou X. The value of prognostic factors for survival in synchronous multifocal lung cancer: a retrospective analysis of 164 patients. Ann Thorac Surg 2018 Mar;105(3):930-936. [doi: 10.1016/j.athoracsur.2017.09.035] [Medline: 29394993]
29. Voltolini L, Rapicetta C, Luzzi L, et al. Surgical treatment of synchronous multiple lung cancer located in a different lobe or lung: high survival in node-negative subgroup. Eur J Cardiothorac Surg 2010 May;37(5):1198-1204. [doi: 10.1016/j.ejcts.2009.11.025] [Medline: 20022516]
30. Wang SX, Lei L, Guo HH, Shrager J, Kunder CA, Neal JW. Synchronous primary lung adenocarcinomas harboring distinct MET Exon 14 splice site mutations. Lung Cancer 2018 Aug;122:187-191. [doi: 10.1016/j.lungcan.2018.06.019] [Medline: 30032829]
31. Stella F, Luciano G, Dell'Amore A, et al. Pulmonary metastases from NSCLC and MPLC (multiple primary lung cancers): management and outcome in a single centre experience. Heart Lung Circ 2016 Feb;25(2):191-195. [doi: 10.1016/j.hlc.2015.07.016] [Medline: 26525847]
32. Surveillance, Epidemiology, and End Results Program. National Cancer Institute. URL: https://seer.cancer.gov/ [accessed 2026-06-18]

Abbreviations

AJCC: American Joint Committee on Cancer
AUC: area under the receiver operating characteristic curve
DCA: decision curve analysis
FPLC: first primary lung cancer
IDI: integrated discrimination improvement
MPLC: multiple primary lung cancer
NRI: net reclassification improvement
OS: overall survival
SEER: Surveillance, Epidemiology, and End Results
SPLC: second primary lung cancer
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

Edited by M Balcarras; submitted 10.Nov.2025; peer-reviewed by S Zhang, X Ding; revised version received 17.May.2026; accepted 19.May.2026; published 26.Jun.2026.

Please cite as:
Luo W, Rao K, Chen H, Hu C, Liu S, Wang J, Zhang D, Chen L
Predicting Overall Survival in Patients With Multiple Primary Lung Cancer: Nomogram Development and Validation Study
JMIR Cancer 2026;12:e87275
URL: https://cancer.jmir.org/2026/1/e87275
doi:10.2196/87275

© Wenzhi Luo, Kengliang Rao, Hongjia Chen, Cailian Hu, Shengming Liu, Jing Wang, Dongdong Zhang, Li Chen. Originally published in JMIR Cancer (https://cancer.jmir.org), 26.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Latent Dirichlet Allocation Topic Modeling to Uncover Latent Research Topics and Trends in Renal Cell Carcinoma: Bibliometric Review

Javier De La Hoz-M[1], PhD; Karime Montes-Escobar[2,3], PhD; Carlos Alfredo Salas-Macias[4], PhD; Martha Fors[5], MD, PhD; Santiago J Ballaz[6], PhD

[1]Universidad del Magdalena, Santa Marta, Colombia

[2]Departamento de Ciencias Agrícolas, Facultad de Ingeniería Agrícola, Universidad Técnica de Manabí, Portoviejo, Ecuador

[3]Departamento de Formación y Desarrollo Científico en Ingeniería, Facultad de Ingeniería, Ciencia y Tecnología, Universidad Bernardo O'Higgins, Santiago de Chile, Chile

[4]Laboratory of Agroecosystems Functioning and Climate Change – FAGROCLIM, Departamento de Ciencias Agronómicas, Facultad de Ingenierías Agroambientales, Universidad Técnica de Manabí, Santa Ana, Ecuador

[5]Facultad de Ciencias de la Salud, Escuela de Medicina, Universidad de Las Américas, Granados vía a Nayón sin número, Quito, Ecuador

[6]Facultad de Ciencias de la Salud "Dr. Enrique Ortega Moreira", Universidad Espíritu Santo, Samborondón, Ecuador

Corresponding Author:
Martha Fors, MD, PhD
Facultad de Ciencias de la Salud, Escuela de Medicina, Universidad de Las Américas, Granados vía a Nayón sin número, Quito, Ecuador

Abstract

Background: Renal cell carcinoma (RCC) is a common, often lethal kidney cancer that originates in the renal cortex. Its incidence is rising, and major factors include smoking, obesity, and hypertension, though its etiology is uncertain. While surgery is effective for localized RCC, treatments for metastatic RCC have advanced significantly due to better diagnostic, prognostic, and predictive tools. Despite this progress, challenges remain, including long-term drug resistance and the complexity of RCC as a diverse group of diseases rather than a single entity.

Objective: The aim of this bibliometric review was a comprehensive analysis of the topics and trends in RCC research, offering a foundation for future investigations.

Methods: We used R "Bibliometrix" to conduct a bibliographic search in Scopus and PubMed covering publications from 1975 to 2023 to statistically assess the distribution of publications associated with RCC by year, journal, and country. Topic modeling of RCC research was conducted using latent Dirichlet allocation, a Bayesian network-based probabilistic algorithm that identifies unobserved thematic clusters in a collection of text documents. Trends in the retrieved themes were then characterized by using regression slopes over time, across countries, and in different journals. These trends were visualized as a heatmap, which was then used for hierarchical clustering to group similar topics based on their correlation strengths.

Results: A total of 35,228 documents from 3070 sources were found, with a steady yearly growth of 9.86% and 118 participating countries. Thirty topics with the best coherence score were found in 8 crucial domains: treatment and therapies, biomolecular and genetic characteristics, disease characteristics and progression, diagnosis and evaluation, metastasis and dissemination, epidemiology and risk factors, related conditions, and pathological features. The pertinent clustergrams that resulted from the heatmaps mirrored the latent Dirichlet allocation's algorithm identification of major RCC research subjects.

Conclusions: Over 50 years, RCC research's focus has shifted from diagnosis and assessment to a more thorough understanding of disease characteristics and progression. Because many patients are diagnosed with abdominal imaging studies, an emerging topic in RCC is diagnostic imaging and radiological evolution. The advances in omics technologies and the function of microRNA signature in the progression, diagnosis, therapy targeting, and prognosis of RCC have garnered a lot of attention. The discovery of the genetic background has enhanced our understanding of the growth of RCC. Drug resistance, local RCC ablation, and postoperative surveillance of RCC recurrence following nephrectomy are key future research avenues. The next generation of drug-targeted therapy and immunotherapy will make it possible to successfully treat metastatic RCC following nephrectomy. Neglected topics include the association between ferroptosis and RCC, the long-term assessment of novel treatments, and the application of artificial intelligence on RCC. Our bibliographic review delivered pertinent data for clinical decision-making and the planning of future RCC research.

(JMIR Cancer 2026;12:e78797) doi:10.2196/78797

XSL•FO
RenderX

KEYWORDS

bibliometrics; carcinoma renal cell; knowledge discovery; machine learning; research

Introduction

Renal cell carcinoma (RCC) is the term coined to describe the malignant transformation of proximal renal tubular epithelium within the renal cortex [1]. RCC accounts for approximately 90% of all renal malignancies [2] and for approximately 2% of all cancer diagnoses and cancer deaths worldwide [3]. With a 2:1 ratio of new diagnoses, men are more likely than women to be affected by RCC, whose incidence rises significantly with age. RCC is a serious health problem since it is often a lethal kidney cancer with an increasing incidence worldwide. Its relevance is also tied to challenges in early diagnosis and the development of aggressive subtypes like those involving tumor thrombus. RCC's etiology is uncertain. The 3 main risk factors for RCC include being overweight, having hypertension, and smoking cigarettes [3]. Medical disorders such as chronic kidney disease, hemodialysis, kidney transplantation, polycystic kidney disease, and renal stones are additional risk factors for RCC. Numerous dietary, occupational, environmental, and lifestyle factors have also been linked to the development of RCC [4]. Even though the majority of RCCs are sporadic, 3% - 5% of all RCC diagnoses occur in patients younger than 46 years, which suggests an underlying RCC form that is inherited [5].

In terms of treatment, surgery is a successful strategy for managing localized RCC, but conventional chemotherapy is ineffective for treating metastatic RCC. Thankfully, during the past 10 years, amazing progress has been made in treating metastatic RCC, resulting in a significant drop in the cancer's death rates despite a continuous rise in the number of individuals receiving a diagnosis. The main factor for the improvement in RCC survival over the past few decades has been the wide diagnostic, prognostic, and predictive methodologies that are currently available [6,7]. Despite these advancements, long-term RCC drug resistance is still a problem [8]. While accelerating RCC cures is critically needed, research on RCC faces various obstacles. Prior to the discovery of the *VHL* gene, kidney cancer was regarded as a single disease [9]. In fact, kidney cancer is a multitude of diverse diseases, each with its own genetic makeup [10]. This has delayed the search for a cure by impeding reproducibility across studies and appropriate interpretation of the research [11].

RCC offers an intricate and challenging research landscape that hinders scientific progress. The goal of this study was to conduct a comprehensive and up-to-date summary of the topic structure, novel research avenues, study trends, and knowledge gaps in RCC research. Conventional bibliometric methods, scientific mappings, and network visualization studies fall short in offering this kind of text analysis [12], since they frequently necessitate manual categorization or extensive, subjective human intervention [13-16]. Instead, we sought a corpus of text-based data for research trends and topics by using a topic algorithm model called latent Dirichlet allocation (LDA) [17]. We intended for RCC research documents "to tell the story themselves" and for topics to emerge on their own, without human intervention, and only based on their statistical characteristics.

Methods

Search Strategy and Data Collection

This study was based on data obtained from PubMed [18] and Scopus [19] as of April 6, 2023. We chose Scopus over other databases because of its superior coverage in the health sciences field, accurate indexing, and "federated search interface" (ie, functionality), which enables us to query the content found across its sources using a common or standardized search form. The authoritative and comprehensive PubMed is the top choice for searching medical and health sciences literature. Medical Subject Headings (MeSH) terms were used for the PubMed search to improve comprehensiveness. Raw data were stored in *TXT* and *CSV* files, respectively. The R "Bibliometrix" tool of the R Statistical software [20] was used to clean data and integrate the 2 databases' unique publications into a combined dataset (with and without assigned DOI) [18,19,21-23]. At this point, duplicate documents with an assigned DOI (1809 in total) were eliminated. The inclusion criteria for the outcome were all the research documents, written in English in peer-reviewed journals, that were published between 1974 (the earliest article we found) and 2023, and that dealt with RCC. Books, book chapters, gray literature, and reports were not included to avoid noise. The search strings, which used a Boolean computation, are indicated in Table 1. Only those articles with the term "renal cell carcinoma" in their title or abstract were selected. The leader and the other authors reviewed the complete list of all possible acceptable publications. Their reliability and value to the field were based on criteria like the journal's impact factor, author affiliations, and citation count.

XSL•FO
RenderX

Table . Enhanced information retrieval for research on renal cell carcinoma.

Database	Search data	Search string	Results, n
PubMed	April 6, 2024	“renal cell carcinoma”[Title/Abstract] AND “english”[Language] AND “journal article”[Publication Type] AND 1974/01/01:2023/12/31[Date - Publication]	38,577
Scopus	April 6, 2024	TITLE-ABS (“renal cell carcinoma”) AND PUBYEAR>1973 AND PUBYEAR<2024 AND (LIMIT-TO (DOCTYPE , “ar”) OR LIMIT-TO (DOCTYPE , “re”)) AND (LIMIT-TO (SRCTYPE , “j”)) AND (LIMIT-TO (LANGUAGE , “English”))	40,479

Bibliometrix Analysis

Our bibliometric review procedure adhered to best practice guidance published elsewhere [24] (Checklist 1). A preliminary descriptive analysis of the retrieved information was carried out using the R package bibliometrix [22]. This open-source application analyzes publication and citation metrics using mathematical and statistical methods to obtain a broad picture of the scientific output that was within the purview of the study. The questions answered were as follows: (Q1) What are the primary research topics in RCC? (Q2) How have RCC research questions changed over time? (Q3) How are these research topics distributed across countries and scientific journals? Three levels of analysis—countries, sources, and authors—were included at this stage to answer the above questions.

The annual growth rate of publications was calculated using the Bibliometrix package in R, which computes the compound annual growth rate (CAGR) using the following equation:

$$(1) CAGR=(Vf/Vi)1n-1,$$

where V_f is the number of publications in the final year of the study period, V_i is the number of publications in the initial year of the study period, and n is the number of years between the initial and final year.

Latent Dirichlet Allocation

The unsupervised machine learning algorithm LDA [17] was applied to identify topics. Considered an extension of the probabilistic latent semantic analysis, it has its roots in Bayesian models [17,25]. Topics in LDA are thought of as multinomial distributions of vocabulary terms, in which each word has a given probability of occurring in a topic. This leads to the prominence of words that are more frequently used in a topic, creating clusters that reflect specific underlying themes. LDA does not require prior knowledge of the topics or the way they are presented in the texts. Rather, topics merely flow from the statistical properties of the data and the model’s underlying assumptions. For thematic analysis, it was decided to use abstracts rather than entire texts because topics are more coherent and ranked higher in large document collections. Inaccurate or noisy terms have less impact on topic word distribution [26].

The LDA model was validated using the Cv metric, which is grounded in the distributional hypothesis stating that words with similar meanings tend to coexist in similar contexts [27]. In other words, Cv rates the semantic similarity of words within a topic (ie, topic interpretability). The Cv score was calculated by looking at word co-occurrence statistics in a reference corpus and their conformity to human-like semantic interpretation. The package textmineR was used for such analysis, which made it easier to determine the ideal number of topics (k) for the study. Because higher scores indicate better interpretability, a model with the highest coherence score among those in the study rank (from k=4 to 50) was selected to reach a balance between granularity and thematic clarity.

Identifying Research Topics

The procedure for identifying topics through LDA was divided into 3 stages: (1) preprocessing, (2) construction of the LDA model, and (3) assigning labels to topics. LDAShiny [28], an open-source R package that uses Bayesian inference for LDA and machine learning algorithms to improve the analytical process, was selected for the first 2 stages.

Preprocessing

Converting all documents into a standardized format for ease of handling was the objective of the stage known as “Text refining” [29]. Initially, textual data consisted solely of character sets. To enhance topic coherence, each abstract underwent tokenization using bigrams, which are consecutive unigram combinations. This process involves converting text to lowercase and removing punctuation marks, dashes, brackets, numbers, spaces, and “stop words.” The list of stop words was extracted from standard libraries such as Natural Language ToolKit and Snowball and was modified to include unrelated terms unique to the medical and technical domains.

The preprocessed data result in the creation of a document-term matrix in which each document is represented as a vector containing an unordered collection of words. If the corpus contains a total of V words, each document becomes a V-dimensional vector, with the value of each element representing the frequency of the corresponding word in the document.

XSL•FO
RenderX

Construction of the LDA Model

LDA assumes that topics are shared by all documents in the collection, while subject proportions vary stochastically between documents, as they are randomly extracted from a Dirichlet distribution [30]. Establishing the expected number of topics was done a priori, making it a nontrivial task to choose the right number of topics (*k*) for a given collection of items. Since the optimal number of topics was unknown beforehand, we generated different models ranging from 4 to 40 topics. We ran 1000 iterations for Gibbs sampling [31] and utilized the default values of the LDAShiny package for Dirichlet parameters α and β. We used Cv as the topic coherence measure of the topics generated by LDA models [32].

Assigning Labels to Topics

The LDA model generates topics without semantic labels. Given that algorithmic analyses are not always able to fully capture the implicit meanings of human language, manual labeling is widely recognized as a normal practice in topic modeling [32]. The manual topic labeling involved a diverse team of 7 experts, including the authors and independent scholars with backgrounds in oncology research and bibliometric analysis. Team members were provided with 2 sources of information: the lists of most frequently occurring words (presumably) provided by the model and a sample of 3 document titles with their corresponding summaries classified by the algorithm. They were asked to independently validate and summarize the identified topics and investigate existing literature to identify research trends, gaps, and influential works. Discrepancies between team members were resolved through remote communications to minimize bias, improve rigor, and ensure the theme structures and semantic interpretation were aligned with the research objective. Reference [32] provides a guide to the procedures used to ensure the trustworthiness of the labels issued, with the difference that we used up to 16 annotators rating candidate labels.

By providing the most relevant and thematically aligned examples within each topic, these articles guaranteed readability and clarity. The 2 articles chosen allowed for a concise and efficient summary of each topic's key ideas without overwhelming the analysis and the reader. This approach strikes a balance between realistic representation and practical interpretability. These articles were then condensed into succinct summaries that encapsulated the essence of each topic. This manual approach not only provided a gold-standard reference [32] but also ensured interpretability and utility in the context of the study.

Quantitative Indices

For each topic, additional characteristics were revealed, especially at the journal and country levels, through statistical description based on the probability distributions of document-topic and topic-word acquired through LDA. To make results and findings more evident, we used certain quantitative indices suggested by Xiong and colleagues [33], which were obtained by adding document-topic and topic-word distributions. The indexes were described as follows.

The distribution of topics over time was obtained by the following equation:

$$(2)\theta ky=\sum m\in y\theta mk/ny,$$

where *m j* represents articles published each year, θ_{mk} is the proportion of the *k*th topic in each item, and n^y is the total number of articles published in the year.

Topic distribution across journals was defined as the ratio of the *k*th topic in the journal *j*: θkj as indicated in the following equation:

$$(3)\theta kj=\sum m\in j\theta mk/nj,$$

where *m j* represents the articles in a particular journal, θ_{mk} is the proportion of the *k*th topic on each item, and n^j is the total number of articles published in the journal *j*.

Topic distribution across countries was defined as the ratio of the *k*th topic in the country *c*, as in the following equation:

$$(4)\theta kc=\sum m\in c\theta mk/nc,$$

where *m c* represents the articles in a particular country, θ_{mk} is the proportion of the *k*th topic on each item, and n^c is the total number of articles published in the country *c*.

Statistics

With the purpose of facilitating the characterization of the topics in terms of their tendency, topic datasets were also submitted to simple regression slopes where the year, country, and journal were the dependent variables, while the proportion of the topics in the corresponding year, country, and journal was the response variable [34]. From the regression slopes, we determined the directionality of these trends and set a significance threshold of $P<.01$ (Eq2). Topics that showed statistically significant positive slopes were identified as having upward trends, while those with statistically significant negative slopes were in decline. Tendencies were finally visualized using the *ggcorrplot* library of R to represent correlation strengths as a heatmap matrix. The color-mapped matrix was subjected to advanced hierarchical clustering analysis in order to investigate and compile correlation datasets, given its visual form like a tree-shaped dendrogram. The *Agnes* function with *Ward's* method showed the agglomerative hierarchical clustering of variables. Each leaf of the dendrogram corresponded to one observation (variable), and the fusion height showed the dissimilarity between 2 observations on the vertical axis. A cut height for cluster identification was calculated using the Average Silhouette method [35].

Results

Overview of the Dataset

The consolidated dataset was obtained by combining the results and removing duplicates, totaling 39,856 articles (Figure 1). After the relevant Excel file was created, 4628 articles lacking titles, abstracts, or affiliations were removed. Within this extensive dataset, 35,228 documents were assembled, demonstrating an annual growth rate of 9.86%. The extensive summary of the key descriptive characteristics pertaining to RCC from 1974 to 2023 can be found in Table 2. A substantial

XSL•FO
RenderX

number of information sources, 3070 in total, were revealed by the data. Given that the document's average age was 11.6 years (the time since the publication of the examined articles), it is likely that much of the research was carried out some time ago. On average, however, each document received 32.35 citations, demonstrating their influence and recognition in the field (Table 2).

Figure 1. The workflow for article selection and bibliometric analysis in renal cell carcinoma using the PubMed and Scopus databases.

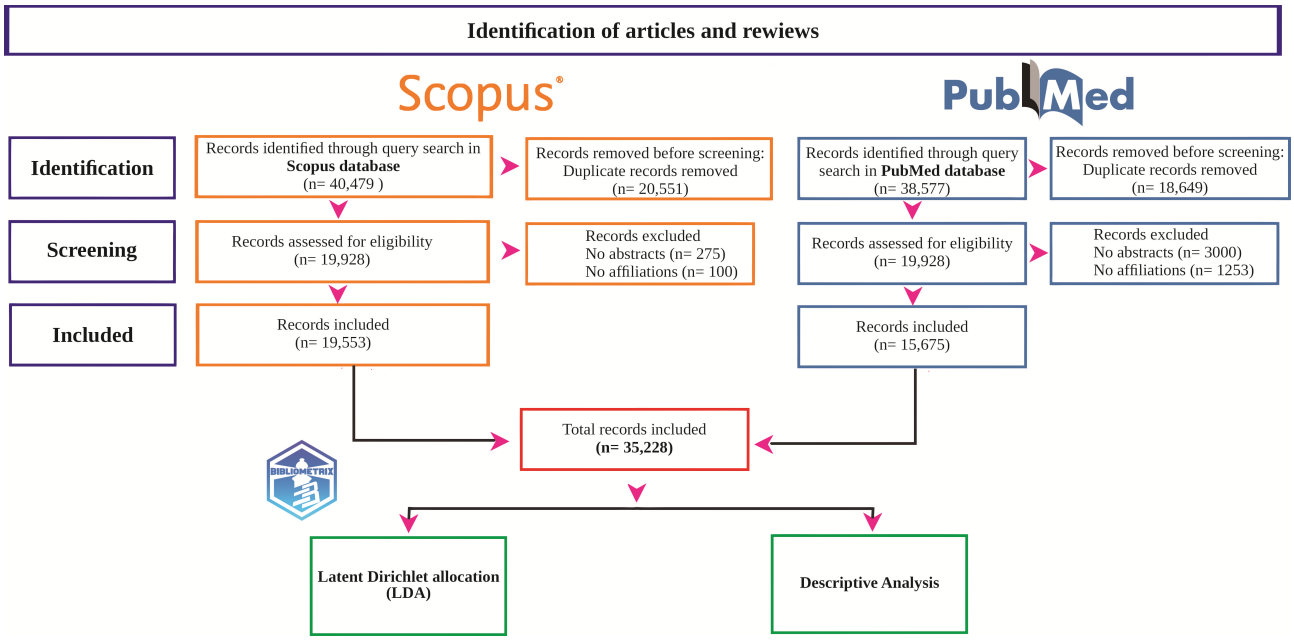


Table . Comprehensive overview of key descriptive characteristics and publication metrics on renal cell carcinoma from 1974 to 2023. Retrieved from the PubMed and Scopus databases.

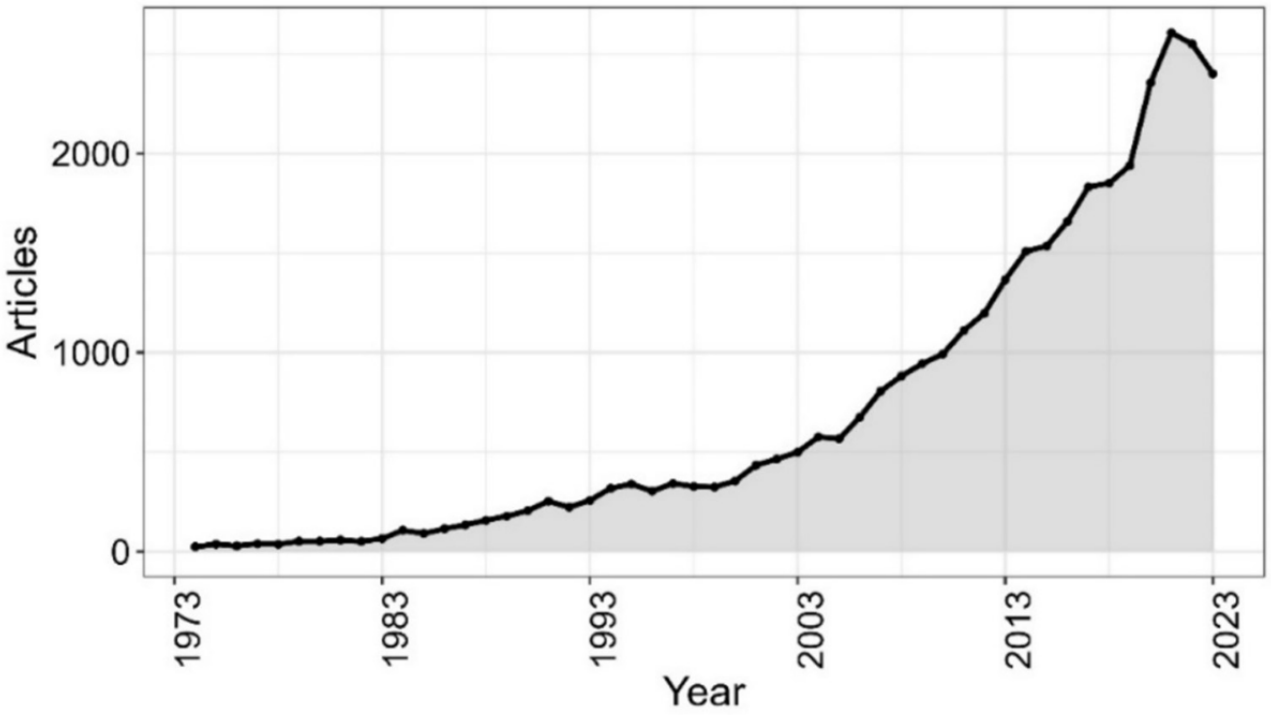
Description	Results
Main information about data	
Timespan, range	1974 - 2023
Sources (journals, books, etc), n	3070
Total documents, n	35,228
Annual growth rate (%), mean	9.86
Document age (y), mean	11.6
Citations per document, mean	32.35
Document contents, n	
Keywords plus	52,792
Author's keywords	31,248
Authors, n	
Total authors	95,238
Authors of single-authored documents	608
Author collaboration	
Single-authored docs, n	769
Coauthors per doc, mean	7.55
Document types	
Original research	30,913
Review	4315

With respect to the yearly output of documents on RCC, Figure 2 provides an overview of increased production from 1974 (24 articles) to 2023 (2401 articles). Throughout the 1980s and 1990s, the document production grew modestly. However, the 2000s saw a significant increase in production, reaching a peak in 2017 (1832 articles) and indicating a solid trend in recent

XSL•FO
RenderX

years. Table 3 lists the top 30 scientific journals, while a global map (Figure 3) shows the 118 countries that were involved in RCC research. With 10,308 publications, the United States clearly was the top contributor.

Figure 2. Annual production of documents on renal cell carcinoma from 1974 to 2023.



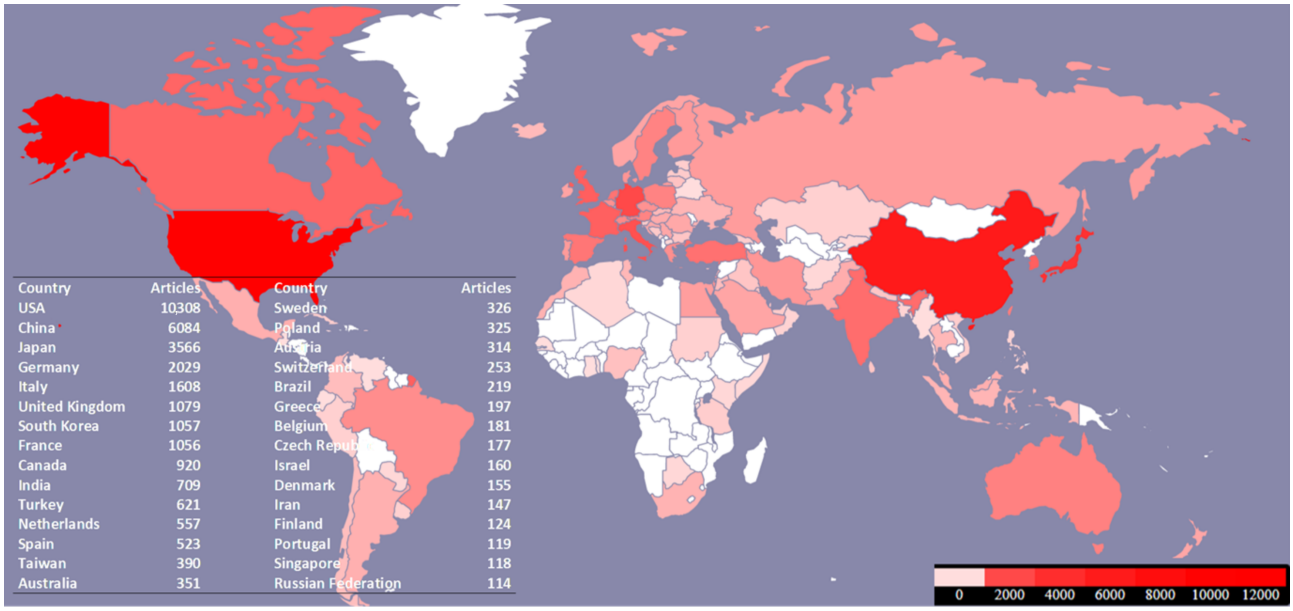
XSL•FO
RenderX

Table . Top 30 scientific journals for research on renal cell carcinoma based on articles published between 1974 and 2023.

Source	Abbreviation	Articles, n
Journal of Urology	*J Urol*	1161
Urology	*Urology*	839
Urologic Oncology: Seminars and Original Investigations	*Urol Oncol*	560
European Urology	*Eur Urol*	542
Frontiers in Oncology	*Front Oncol*	481
Cancer	*Cancer*	478
Clinical Genitourinary Cancer	*Clin Genitourin Cancer*	474
BJU International	*BJU Int*	451
International Journal of Urology	*Int J Urol*	422
Oncotarget	*Oncotarget*	363
PLOS One	*PLoS One*	362
Cancers	*Cancers*	350
Urologia Internationalis	*Urol Int*	316
International Journal of Cancer	*Int J Cancer*	311
Clinical Cancer Research	*Clin Cancer Res*	308
Oncology Letters	*Oncol Lett*	307
World Journal of Urology	*World J Urol*	297
British Journal of Cancer	*Brit J Cancer*	291
BMC Cancer	*BMC Cancer*	273
Scientific Reports	*Sci Rep*	248
American Journal of Surgical Pathology	*Am J Surg Pathol*	236
American Journal of Roentgenology	*Am J Roentgenol*	210
Journal of Clinical Oncology	*J Clin Oncol*	206
Cancer Research	*Cancer Res*	200
Human Pathology	*Hum Pathol*	196
Oncology Reports	*Oncol Rep*	189
International Urology and Nephrology	*Int Urol Nephrol*	188
International Journal of Molecular Sciences	*Int J Mol Sci*	186
Medicine (United States)	*Med (United States)*	172
Journal of Endourology	*J Endourol*	165

XSL•FO
RenderX

Figure 3. Distribution of geographical origins in the analysis of 35,228 published articles on renal cell carcinoma from 1974 to 2023. The table displays the top 30 countries with the highest research production.



Country	Articles	Country	Articles
USA	10308	Sweden	326
China	6084	Poland	325
Japan	3566	Austria	314
Germany	2029	Switzerland	253
Italy	1608	Brazil	219
United Kingdom	1079	Greece	197
South Korea	1057	Belgium	181
France	1056	Czech Republic	177
Canada	920	Israel	160
India	709	Denmark	155
Turkey	621	Iran	147
Netherlands	557	Finland	124
Spain	523	Portugal	119
Taiwan	390	Singapore	118
Australia	351	Russian Federation	114

Latent Dirichlet Allocation

The methodological rigor, reproducibility, and accuracy of LDA were ensured using advanced bibliometric tools such as Bibliometrix and Textminer. There were 30 topics with the best coherence score in the LDA model (Figure 4). The terms with the highest probabilities and semantically relevant labels for each latent topic are shown in Table 4.

XSL•FO
RenderX

Table . Topics discovered from 35,228 articles on renal cell carcinoma published between 1974 and 2023.

Terms	Article numbers, n	Top terms	Label	Themes	Prevalence (%)
t_1	1411	inhibitor, sunitinib, treatment, target, drug, kinas, vegf, growth, factor, tki, sorafenib, resist, agent, tyrosin, growth_factor	Inhibitors in treatment	Treatment and therapies	3.455
t_2	563	rcc, level, patient, serum, control, rcc_patient, increas, group, concentr, elev, blood, distribut, plasma, healthi, compar	Serum levels and patient control	Diagnosis and evaluation	2.414
t_3	1240	express, tissu, normal, protein, tumor, level, correl, mrna, sampl, kidnei, posit, compar, marker, normal_tissu, express_level	Protein expression in tissues and tumors	Biomolecular and genetic characteristics	3.866
t_4	1370	gene, mutat, tumor, chromosom, genet, dna, loss, alter, famili, sequenc, region, identifi, methyl, variant, r	Genetic mutations and chromosomal alterations	Biomolecular and genetic characteristics	3.067
t_5	2113	rcc, clear, tumor, papillari, type, case, subtyp, carcinoma, featur, clear_rcc, prcc, posit, chromophob, histolog, pattern	Histological features and subtypes	Disease characteristics and progression	4.537
t_6	1293	tumor, surgic, thrombu, complic, oper, nephrectomi, surgeri, resect, laparoscop, blood, vena, ivc, postop, perform, approach	Surgical approaches and complications	Pathological features	3.053
t_7	939	vhl, Î±, hif, factor, Î², hif_Î±, protein, hypoxia, activ, induc, von, lindau, hippel, hippel_lindau, von_hippel	Hypoxia factors and related proteins	Pathological features	2.654
t_8	575	effect, treatment, ablat, control, local, search, evid, review, includ, meta, radiat, perform, systemat, outcom, percutan	Treatment effects and local ablation	Treatment and therapies	2.281
t_9	770	immun, pd, respons, immunotherapi, combin, nivolumab, checkpoint, ici, inhibitor, immun_checkpoint, therapi, checkpoint_inhibitor, treatment, death, anti	Immunotherapy and immune responses	Treatment and therapies	2.416
t_10	544	model, predict, score, risk, base, valid, group, curv, clinic, perform, cohort, featur, characterist, set, auc	Risk prediction models and clinical assessment	Epidemiology and risk factors	2.782

XSL•FO
RenderX

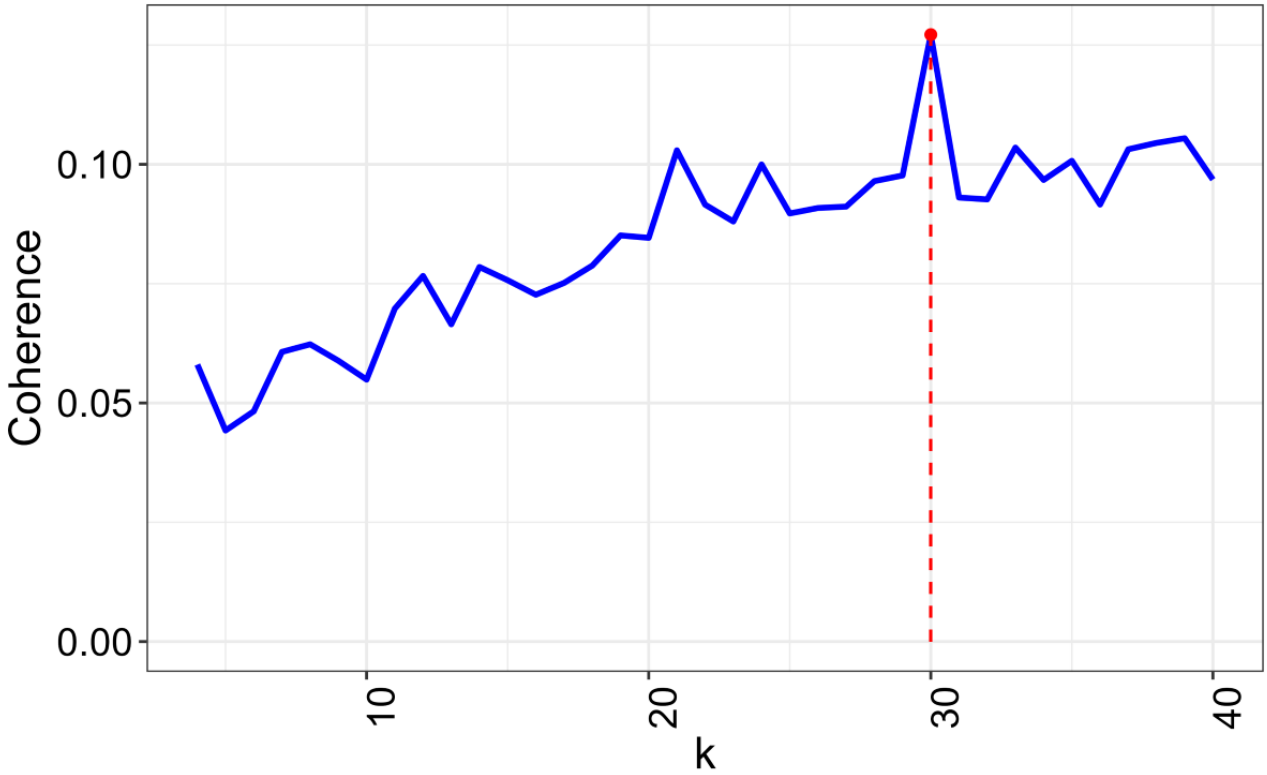
Terms	Article numbers, n	Top terms	Label	Themes	Prevalence (%)
t_11	1593	imag, ct, enhanc, lesion, mass, contrast, evalu, tomographi, mri, comput, detect, phase, comput_tomographi, scan, find	Diagnostic imaging and radiological evaluation	Diagnosis and evaluation	3.633
t_12	1794	patient, month, surviv, median, mrcc, progress, o, group, pf, treat, metastat, free, line, progress_free, receiv	Survival and disease progression	Disease characteristics and progression	4.504
t_13	125	tumor, node, lymph, lymph_node, invas, metastasi, distant, involv, stage, crcc, posit, distant_metastasi, presenc, patient, node_metastasi	Metastasis and lymph node involvement	Metastasis and dissemination	1.776
t_14	724	metastat, metastas, metastasi, primari, patient, bone, site, rcc, lung, lesion, brain, primari_tumor, resect, pet, diseas	Metastasis and lesions in other organs	Metastasis and dissemination	2.724
t_15	324	patient, syndrom, develop, symptom, diseas, clinic, common, earli, relat, occur, sever, adult, hypertens, infect, manifest	Syndromes and clinical manifestations	Pathological features	2.108
t_16	1969	tumor, activ, human, effect, induc, increas, mice, antibodi, antigen, anti, deriv, specif, growth, line, cytotox	Tumor activity and immune responses	Disease characteristics and progression	4.218
t_17	2190	patient, nephrectomi, year, recurr, group, surgeri, follow, rate, rang, underw, local, radic, month, partial, diseas	Nephrectomy and recurrence	Metastasis and dissemination	5.242
t_18	1147	rcc, risk, ci, increas, associ, ag, patient, incid, compar, ratio, popul, year, mortal, interv, data	Risk factors and epidemiology	Epidemiology and risk factors	3.858
t_19	2049	patient, respons, treatment, dose, toxic, dai, week, event, advers, receiv, diseas, progress, evalu, phase, efficaci	Toxicity and adverse events in treatments	Pathological features	4.779
t_20	102	bladder, prostat, urolog, urinari, health, urotheli, tsc, prostat_cancer, malign, kluwer, lippincott, wolter, wolter_kluwer, william, wilkin	Urological cancers and related conditions	Related conditions	1.628

XSL•FO
RenderX

Terms	Article numbers, n	Top terms	Label	Themes	Prevalence (%)
t_21	351	kidnei, diseas, long, term, transplant, function, long_term, develop, chronic, kidnei_diseas, egfr, diabet, dialysi, donor, recipi	Chronic kidney disease and transplant	Related conditions	1.967
t_22	1917	ccrcc, gene, clear, identifi, express, relat, prognosi, clear_ccrcc, ccrcc_patient, cancer, pathwai, biomark, potenti, data, genom	Gene expression and prognosis	Biomolecular and genetic characteristics	4.362
t_23	487	develop, molecular, potenti, recent, therapeut, provid, approach, clinic, understand, import, review, research, base, applic, strategi	Molecular advances and therapeutics	Pathological features	3.026
t_24	1388	surviv, patient, prognost, factor, independ, multivari, specif, cox, o, prognosi, free, prognost_factor, outcom, predictor, specif_surviv	Prognostic factors and survival	Disease characteristics and progression	4.137
t_25	202	cancer, type, lung, breast, kidnei_cancer, melanoma, cancer_patient, includ, lung_cancer, small, breast_cancer, research, acid, metabol, colorect	Nonrenal cancers and comparative analysis	Pathological features	2.354
t_26	555	malign, tumor, diagnosi, biopsi, benign, pancreat, neoplasm, thyroid, mass, case, diagnost, lesion, diagnos, carcinoma, small	Diagnosis and characterization of tumors	Diagnosis and evaluation	2.559
t_27	3022	rcc, mir, prolifer, inhibit, express, role, regul, assai, line, invas, effect, target, apoptosi, promot, migrat	Gene regulation and microRNA expression	Biomolecular and genetic characteristics	5.728
t_28	604	tumor, stage, grade, size, patholog, pt, tumor_size, low, nuclear, clinic, fuhrman, correl, necrosi, fuhrman_grade, tnm	Staging and pathological features	Pathological features	2.986
t_29	1168	therapi, treatment, clinic, target, improv, trial, advanc, system, review, manag, metastat, rcc, target_therapi, benefit, diseas	Advanced therapies and management of metastatic disease	Treatment and therapies	3.795
t_30	2699	case, report, year, present, rare, reveal, kidnei, left, mass, diagnosi, adren, report_case, examin, diagnos, literatur	Clinical presentation and diagnosis of rare cases	Diagnosis and evaluation	4.089

XSL•FO
RenderX

Figure 4. Evaluation of coherence scores for topic models in oncology across different numbers of topics (*k*). Coherence: measures the semantic quality of topics. Higher values indicate stronger word relationships. Topic models: identify thematic patterns in texts; in oncology, they reveal key research areas; *k* represents the number of topics in the model. Helps determine the optimal value for thematic analysis.



The 30 identified topics were in 8 crucial domains of RCC research:

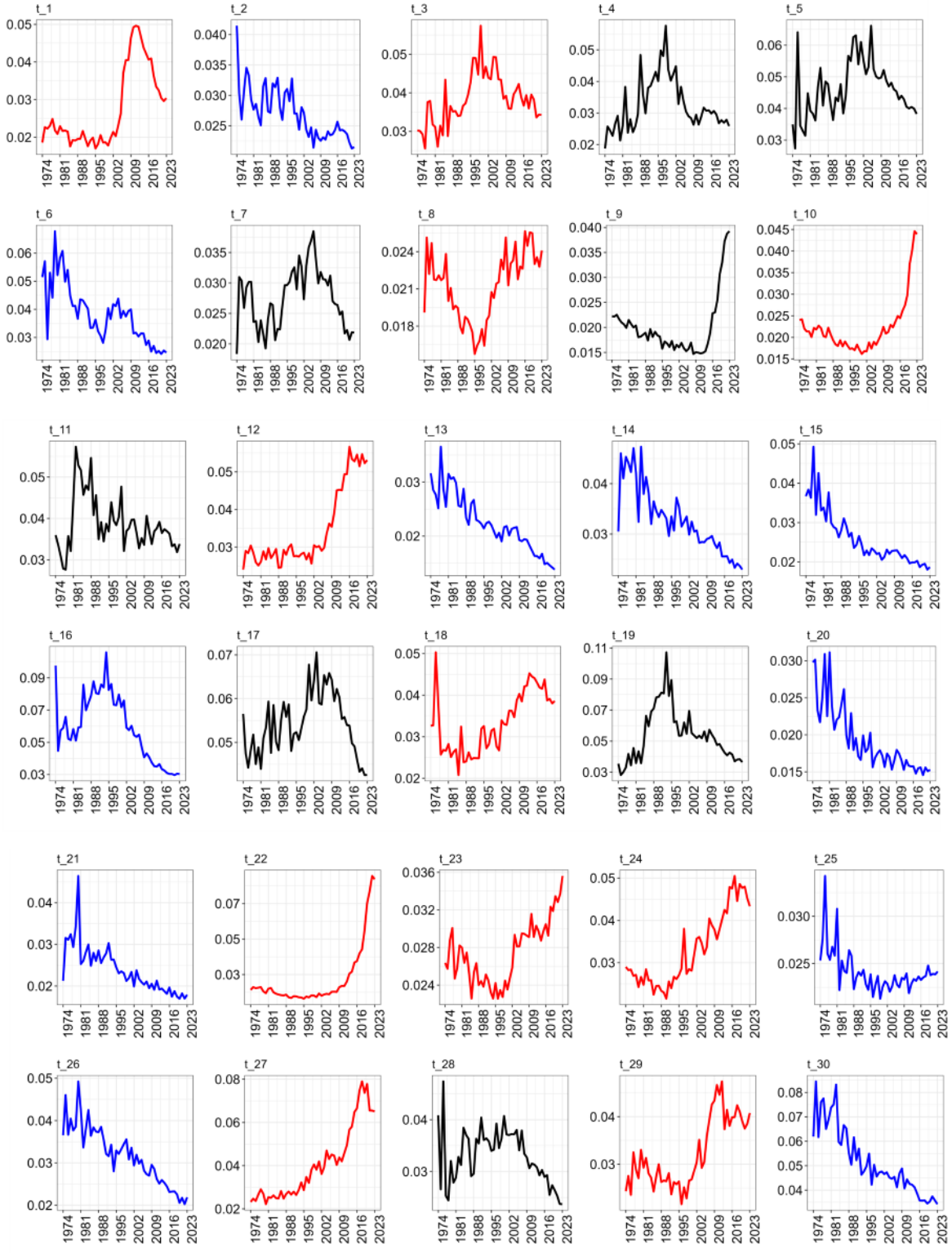
- Treatment and Therapies: This category focused on different approaches to treating RCC, including targeted inhibitors (t_1), local ablation effects (t_8), immunotherapy responses (t_9), and advanced management of metastatic disease (t_29).
- Biomolecular and Genetic Characteristics: Here, the emphasis was on understanding the molecular and genetic makeup of RCC, covering topics like protein expression in tissues and tumors (t_3, t_22, and t_27) and genetic mutations and chromosomal alterations (t_4).
- Disease Characteristics and Progression: This category delved into the histological features and subtypes of RCC (t_5), as well as the dynamics of tumor activity and immune responses (t_16), and survival rates and disease progression (t_12 and t_24) and pathological features (t_28).
- Diagnosis and Evaluation: Topics in this category included diagnostic imaging and radiological evaluation (t_2 and t_11) for RCC detection and the characterization of tumors (t_26 and t_30) for accurate diagnosis.
- Metastasis and Dissemination: Here, the focus was on understanding how RCC spreads, including its involvement with lymph nodes (t_13), lesions in other organs (t_14), and the recurrence of the disease post nephrectomy (t_17).
- Epidemiology and Risk Factors: This category examined the risk prediction models and clinical assessment tools (t_10) used to evaluate RCC risk, as well as the epidemiological factors associated with the disease (t_18).
- Related Conditions: Topics here explored conditions related to RCC, such as urological cancers (t_20), chronic kidney disease, and transplant issues (t_21).
- Pathological Features: Finally, this category encompassed various pathological features of RCC, including surgical approaches and complications (t_6), hypoxia factors and related proteins (t_7), syndromes and clinical manifestations (t_15), toxicity and adverse events in treatments (t_19), molecular advances and therapeutics (t_23), and comparisons with nonrenal cancers (t_25).

Topic Trends

The topic distribution by document θm was added to compute the average probability θky of all the articles published in a particular year to identify the trends (Figure 5). We found that the probabilities of some topics steadily increased over time (red). Black indicates topics with no discernible trend, whereas blue denotes topics with a decreasing behavior.

XSL•FO
RenderX

Figure 5. Trends of research topics in renal cell carcinoma between 1973 and 2023: increasing (red), decreasing (blue), and stable (black) topic dynamics over time.



Heatmaps

Although the discovery of 30 unique themes with high coherence scores, a granular and nuanced analysis of the research landscape, was made possible by the LDA method, there was a need for validating the theme output by a form of visual data analytics. Heatmaps helped examine and understand how thematic patterns were related to variables like publication year, country, and journal. Red highlighting indicates the strongest associations among variables, reflecting higher correlation levels.

Figure 6A illustrates correlations between specific topics and years. Red highlighting indicates the strongest associations among variables, reflecting higher correlation levels. For example, within Cluster 4, Topic 16 (t_16) “Tumor Activity and Immune Responses” is primarily associated with the years

XSL•FO
RenderX

1993, 1985, 1988, 1990, 1992, 1991, 1989, 1995, 1994, 1996, and 1999. Topic 17 (t_17) "Nephrectomy and Recurrence" exhibits stronger correlations with the years 1993, 1990, 1992, 1991, 1989, 1995, and 1994. In Cluster 3, Topic 22 (t_22) "Gene Expression and Prognosis" significantly correlates with the years 2020, 2021, 2022, and 2023, while Topic 27 (t_27), also addressing "Gene Regulation and microRNA Expression," shows significant associations with the years 2017, 2018, 2019, and 2020. Finally, in group 3, Topic Thirty30 (t_30) "Clinical Presentation and Diagnosis of Rare Cases" is linked with the years 1982, 1983, 1981, 1978, 1977, and 1975.

Figure 6. Heatmaps to correlate topics with year (A), country (B), and source (C).

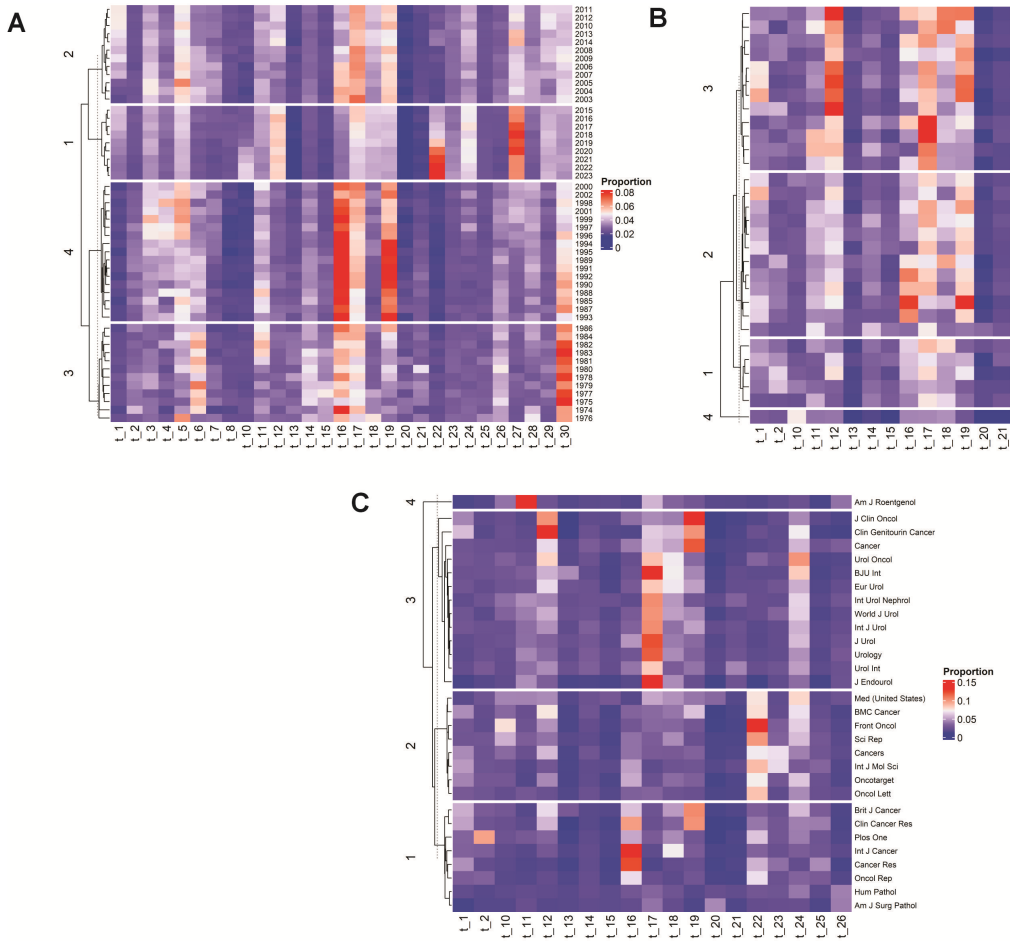


Figure 6B presents the interactions between countries and topics across different groups. In group 1, Topic 27 (t_27), "Gene Regulation and microRNA Expression," is predominantly correlated with Taiwan. Group 2's topics, "Tumor Activity and Immune Responses" and "Toxicity and Adverse Events in Treatments," are closely associated with the Netherlands. For group 3, "Survival and Disease Progression" shows significant correlations with the Czech Republic, Denmark, Belgium, France, and Italy, and "Nephrectomy and Recurrence" is notably linked with Israel and South Korea. Lastly, in group 4, "Gene Expression and Prognosis" and "Gene Regulation and microRNA Expression" are primarily connected with China.

Figure 6C outlines the dominant publication patterns for specific topics within various journals. In group 1, "Tumor Activity and Immune Responses" is predominantly linked with the *International Journal of Cancer* and *Cancer Research*. Group 2's topic, "Gene Expression and Prognosis," has a significant association with *Frontiers in Oncology*. In group 3, "Survival and Disease Progression" relates to *Clinical Genitourinary Cancer*, and "Nephrectomy and Recurrence" is associated with *BJU International*, *Journal of Endourology*, and *Urology*. Lastly, in group 4, "Diagnostic Imaging and Radiological Evaluation" is prominently linked to the *American Journal of Roentgenology*. LDA algorithm's identification of the key RCC research themes was mirrored in the relevant clustergrams that emerged from the hierarchical cluster analysis of the heatmaps.

Discussion

Principal Findings

By displaying each topic as a group of related words, LDA was able to identify latent (hidden) topics within a corpus of documents and demonstrate how each document may be represented as a combination of these topics. According to this method, RCC research has evolved over the past 50 years from concentrating on surgery to comprehending its genetic underpinnings and the impact of new treatments like immune checkpoint inhibitors and targeted therapies, which have improved the prognosis of metastatic RCC.

The disease characterization, progression, and pathological features have dominated the RCC research landscape for the past 50 years (t_16 and t_19) [10]. Kidney cancer was regarded

XSL•FO
RenderX

as a single disease until the *VHL* gene was discovered [9]. Since then, scientists have realized that kidney cancer is a multitude of diverse diseases, each with its own genetic makeup. Although the characterization of genetic mutations and chromosomal alterations linked to the growth of RCC [36,37] has recently improved our knowledge of this kidney cancer (t_22), RCC research does emphasize the necessity of assessment of treatment effects' safety and efficacy, as well as surveillance of RCC recurrence following nephrectomy. Therefore, a better understanding of resistance mechanisms, molecular prognosis (t_22), and immunological responses (t_16) is essential.

The advances in omics technologies over the last 10 years constitute a promising area of personalized RCC cure [1,38,39]. In addition, the microRNA signature in RCC and its function in progression, diagnosis, therapy targeting, and prognosis of RCC (t_27) have also received particular attention [40,41]. Recent RCC research also moves toward earlier cancer detection through broad imaging and radiological evolution (t-11). RCC has a difficult pathological classification, since the histological analysis reveals three most recurrent sporadic types: clear-cell RCC (70% - 75%), papillary RCC (10% - 15%), and chromophobe RCC (5%) [42]. It is predicated on morphologies, architecture, underlying genetic abnormalities, and tumoral protein expression [43,44]. Pathologists can detect these malignancies more accurately and provide better treatment plans and patient outcomes if they are aware of the significance of these markers (t_12 and t_19).

A major cause for worry is the postoperative surveillance of RCC recurrence following nephrectomy (t_17) [45]. Research on local RCC ablation and assessment of treatment effects on safety and efficacy is still ongoing (t_12, t_16, and t_19). The discovery of tailored medication like tyrosine kinase inhibitors (also called TKIs), such as Sunitinib and Sorafenib, has been beneficial in treating metastatic RCC [46,47]. Another well-established component of RCC treatment is checkpoint inhibitor immunotherapy [48], which shows its superior therapeutic efficacy when combined with TKIs [49]. The next generation of TKIs and immunotherapy (t_16) [50] is being developed to overcome some unfavorable outcomes with kinase inhibitors [51,52] and immunotherapy (t_19) [53], as well as the emergence of drug resistance [54]. Only a small number of themes still earn little attention. The topics "Serum Levels and Patient Control" (t_2), "Hypoxia Factors and Related Proteins" (t_7), and "Risk Prediction Models and Clinical Assessment" (t_18) are overlooked and should require additional attention.

In line with the LDA analysis and our own scientific expectations and goals, it was also feasible to identify untapped topics that have not yet been covered by scholarly literature. Intriguing lines of inquiry are the prognostic value of vascular endothelial growth factor [55], endostatin [56], C-reactive protein [57], the hypoxia-induced pathway [58], and ferroptosis [59]. The involvement of chronic inflammation [60] and gut and urinary microbiota in immune modulation of metastatic RCC [61] remains poorly investigated. Clinical judgments and patient stratification in RCC may be enhanced by the creation of new predictive models [62] based on, for instance, genetic biomarkers [63]. The application of artificial intelligence is another potential topic that could help physicians in identifying RCC subtypes by analyzing computed tomography scans, as well as in deconstructing complex epidemiological and environmental factors that influence RCC occurrence, like hypoxia [64].

In contrast to traditional bibliometric analysis, the LDA approach effectively extracted potential themes and inferred implicit information from a large collection of documents. The LDA approach and other topic modeling methods like co-citation and keyword co-occurrence cannot be compared with the same conceptual granularity or depth. LDA goes beyond who cites whom (ie, intellectual connection and research lineage) and uncovers the underlying conceptual themes that bind the literature, which may not be immediately apparent from citation patterns alone. Unlike a list of keywords, which reveals basic relationships, LDA organizes these co-occurring words into meaningful higher-order themes, offering a more detailed knowledge of topic relationships and structure. For this reason, the LDA analysis was not affected by the potential simplicity of the keywords chosen in the study.

The subjectivity involved in manually labeling LDA topics, the possibility of missing publications by using only 2 databases, the linguistic bias introduced by excluding articles written in languages other than English, the constraints of the "bag-of-words" model which disregards grammar and context, and the effects of excluding literature like book chapters are some limitations of the study. Despite the constraints, LDA can disclose "unknown unknowns" by revealing unarticulated or unacknowledged themes. It can also give an overview of the research landscape to identify new topics and interdisciplinary connections, as well as demonstrate how old themes are resurfacing in new ones. As a result, LDA remains the most often used natural language model [13,14,65].

Conclusions

This review offered a thorough summary of how research on RCC has changed over the previous 50 years. LDA helped identify important emerging trends in treatment development to address drug resistance and undesirable side effects, surgical techniques, and immunotherapy advancements, among other topics pertinent to clinical practice and medical research. In summary, this study presents a methodological synthesis of the development of RCC research and delivers pertinent data for clinical decision-making, early identification, and the planning of new biomedical research.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JDLH-M, KM-E
Data curation: KM-E, JDLH-M
Formal analysis: JDLH-M, KM-E, CAS-M
Investigation: JDLH-M, KM-E, CAS-M, MF, SJB
Methodology: JDLH-M, KM-E, CAS-M
Supervision: MF, SJB
Validation: JDLH-M, KM-E, CAS-M, MF, SJB
Writing – original draft: KM-E, SJB
Writing – review & editing: MF, SJB
All the authors equally contributed to the writing of the final version of the manuscript and were responsible for its content.

Conflicts of Interest

None declared.

Checklist 1
Bibliometric analysis checklist.
[PDF File, 97 KB - cancer_v12i1e78797_app1.pdf]

References

1. Hsieh JJ, Purdue MP, Signoretti S, et al. Renal cell carcinoma. Nat Rev Dis Primers 2017 Mar 9;3:17009. [doi: 10.1038/nrdp.2017.9] [Medline: 28276433]
2. Ljungberg B, Campbell SC, Choi HY, et al. The epidemiology of renal cell carcinoma. Eur Urol 2011 Oct;60(4):615-621. [doi: 10.1016/j.eururo.2011.06.049] [Medline: 21741761]
3. Lipworth L, Tarone RE, Lund L, McLaughlin JK. Epidemiologic characteristics and risk factors for renal cell cancer. Clin Epidemiol 2009 Aug 9;1:33-43. [doi: 10.2147/clep.s4759] [Medline: 20865085]
4. Chow WH, Dong LM, Devesa SS. Epidemiology and risk factors for kidney cancer. Nat Rev Urol 2010 May;7(5):245-257. [doi: 10.1038/nrurol.2010.46] [Medline: 20448658]
5. Haas NB, Nathanson KL. Hereditary kidney cancer syndromes. Adv Chronic Kidney Dis 2014 Jan;21(1):81-90. [doi: 10.1053/j.ackd.2013.10.001] [Medline: 24359990]
6. Hemminki K, Försti A, Hemminki A, Ljungberg B, Hemminki O. Progress in survival in renal cell carcinoma through 50 years evaluated in Finland and Sweden. PLoS One 2021;16(6):e0253236. [doi: 10.1371/journal.pone.0253236] [Medline: 34157049]
7. Bukavina L, Bensalah K, Bray F, et al. Epidemiology of renal cell carcinoma: 2022 update. Eur Urol 2022 Nov;82(5):529-542. [doi: 10.1016/j.eururo.2022.08.019] [Medline: 36100483]
8. Aweys H, Lewis D, Sheriff M, et al. Renal cell cancer - insights in drug resistance mechanisms. Anticancer Res 2023 Nov;43(11):4781-4792. [doi: 10.21873/anticanres.16675] [Medline: 37909991]
9. Cowey CL, Rathmell WK. VHL gene mutations in renal cell carcinoma: role as a biomarker of disease outcome and drug efficacy. Curr Oncol Rep 2009 Mar;11(2):94-101. [doi: 10.1007/s11912-009-0015-5] [Medline: 19216840]
10. Cairns P. Renal cell carcinoma. Cancer Biomark 2011;9(1-6):461-473. [doi: 10.3233/CBM-2011-0176]
11. Choueiri TK, Pal SK, Lewis B, Poteat S, Pels K, Hammers H. The 5th Kidney Cancer Research Summit: research accelerating cures for renal cell carcinoma in 2023. Oncologist 2024 Feb 2;29(2):91-98. [doi: 10.1093/oncolo/oyad322] [Medline: 38048064]
12. Matorevhu A. Bibliometrics: application opportunities and limitations. In: Bibliometrics - An Essential Methodological Tool for Research Projects: IntechOpen; 2024. [doi: 10.5772/intechopen.1005292]
13. Escobar KM, Vicente-Villardon JL, de la Hoz-M J, Useche-Castro LM, Alarcón Cano DF, Siteneski A. Frequency of neuroendocrine tumor studies: using latent Dirichlet allocation and HJ-biplot statistical methods. Mathematics 2021;9(18):2281. [doi: 10.3390/math9182281]
14. De La Hoz-M J, Mendes S, Fernández-Gómez MJ, González Silva Y. Capturing the complexity of COVID-19 research: trend analysis in the first two years of the pandemic using a Bayesian probabilistic model and machine learning tools. Computation 2022;10(9):156. [doi: 10.3390/computation10090156]
15. Pilacuan-Bonete L, Galindo-Villardón P, Delgado-Álvarez F. HJ-Biplot as a tool to give an extra analytical boost for the latent Dirichlet assignment (LDA) model: with an application to digital news analysis about COVID-19. Mathematics 2022;10(14):2529. [doi: 10.3390/math10142529]

XSL•FO
RenderX

16. Yu D, Fang A, Xu Z. Topic research in fuzzy domain: Based on LDA topic modelling. Inf Sci 2023 Nov;648:119600. [doi: 10.1016/j.ins.2023.119600]
17. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res 2003;3:993-1022 [FREE Full text]
18. Sood A, Ghosh AK. Literature search using PubMed: an essential tool for practicing evidence- based medicine. J Assoc Physicians India 2006 Apr;54:303-308. [Medline: 16944614]
19. Burnham JF. Scopus database: a review. Biomed Digit Libr 2006 Mar 8;3(1):1. [doi: 10.1186/1742-5581-3-1] [Medline: 16522216]
20. The R Project for Statistical Computing. URL: https://www.r-project.org/ [accessed 2025-12-24]
21. Grün B, Hornik K. Topicmodels: an R package for fitting topic models. J Stat Softw 2011;40:1-30. [doi: 10.18637/jss.v040.i13]
22. Aria M, Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis. J Informetr 2017 Nov;11(4):959-975. [doi: 10.1016/j.joi.2017.08.007]
23. Caputo A, Kargina M. A user-friendly method to merge Scopus and Web of Science data during bibliometric analysis. J Market Anal 2022 Mar;10(1):82-88. [doi: 10.1057/s41270-021-00142-7]
24. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res 2021 Sep;133:285-296. [doi: 10.1016/j.jbusres.2021.04.070]
25. Glendowne P, Glendowne D. Interpretability of API call topic models: an exploratory study. Presented at: 53rd Hawaii International Conference on System Sciences; Jan 7-10, 2020. [doi: 10.24251/HICSS.2020.793]
26. Syed S, Spruit M. Full-text or abstract? Examining topic coherence scores using latent Dirichlet allocation. Presented at: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA); Oct 19-21, 2017; Tokyo, Japan p. 165-174. [doi: 10.1109/DSAA.2017.61]
27. Tang J, Chang Y, Liu H. Mining social media with social theories. SIGKDD Explor Newsl 2014 Jun 16;15(2):20-29. [doi: 10.1145/2641190.2641195]
28. De la Hoz-M J, Fernández-Gómez MJ, Mendes S. LDAShiny: an R package for exploratory review of scientific literature based on a Bayesian probabilistic model and machine learning tools. Mathematics 2021;9(14):1671. [doi: 10.3390/math9141671]
29. Blei DM, Lafferty JD. Dynamic topic models. Presented at: ICML '06: Proceedings of the 23rd International Conference on Machine Learning; Jun 25-29, 2006; Pittsburgh, PA p. 113-120. [doi: 10.1145/1143844.1143859]
30. Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans Pattern Anal Mach Intell 1984 Jun;6(6):721-741. [doi: 10.1109/tpami.1984.4767596] [Medline: 22499653]
31. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. Presented at: WSDM '15: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining; Feb 2-6, 2015; Shanghai, China p. 399-408. [doi: 10.1145/2684822.2685324]
32. Lau JH, Grieser K, Newman D, Baldwin T. Automatic labelling of topic models. Presented at: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies; Jun 19-24, 2011; Portland, OR p. 1536-1545 URL: https://aclanthology.org/P11-1154/ [accessed 2025-12-19]
33. Xiong H, Cheng Y, Zhao W, Liu J. Analyzing scientific research topics in manufacturing field using a topic model. Comput Ind Eng 2019 Sep;135:333-347. [doi: 10.1016/j.cie.2019.06.010]
34. Griffiths TL, Steyvers M. Finding scientific topics. Proc Natl Acad Sci USA 2004 Apr 6;101(Suppl 1):5228-5235. [doi: 10.1073/pnas.0307752101] [Medline: 14872004]
35. Kaufman L, Rousseeuw PJ. Finding Groups in Data: An Introduction to Cluster Analysis: John Wiley & Sons; 1990. [doi: 10.1002/9780470316801]
36. Quddus MB, Pratt N, Nabi G. Chromosomal aberrations in renal cell carcinoma: an overview with implications for clinical practice. Urol Ann 2019;11(1):6-14. [doi: 10.4103/UA.UA_32_18] [Medline: 30787564]
37. Testa U, Pelosi E, Castelli G. Genetic alterations in renal cancers: identification of the mechanisms underlying cancer initiation and progression and of therapeutic targets. Medicines (Basel) 2020 Jul 29;7(8):44. [doi: 10.3390/medicines7080044] [Medline: 32751108]
38. Li QK, Pavlovich CP, Zhang H, Kinsinger CR, Chan DW. Challenges and opportunities in the proteomic characterization of clear cell renal cell carcinoma (ccRCC): a critical step towards the personalized care of renal cancers. Semin Cancer Biol 2019 Apr;55:8-15. [doi: 10.1016/j.semcancer.2018.06.004] [Medline: 30055950]
39. Sharma R, Kannourakis G, Prithviraj P, Ahmed N. Precision medicine: an optimal approach to patient care in renal cell carcinoma. Front Med (Lausanne) 2022;9:766869. [doi: 10.3389/fmed.2022.766869] [Medline: 35775004]
40. Ghafouri-Fard S, Shirvani-Farsani Z, Branicki W, Taheri M. MicroRNA signature in renal cell carcinoma. Front Oncol 2020;10:596359. [doi: 10.3389/fonc.2020.596359] [Medline: 33330087]
41. Yang L, Zou X, Zou J, Zhang GA. A review of recent research on the role of microRNAs in renal cancer. Med Sci Monit 2021 May 8;27:e930639. [doi: 10.12659/MSM.930639] [Medline: 33963171]
42. Muglia VF, Prando A. Renal cell carcinoma: histological classification and correlation with imaging findings. Radiol Bras 2015;48(3):166-174. [doi: 10.1590/0100-3984.2013.1927] [Medline: 26185343]

XSL•FO
RenderX

43. Athanazio DA, Amorim LS, da Cunha IW, et al. Classification of renal cell tumors – current concepts and use of ancillary tests: recommendations of the Brazilian Society of Pathology. Surg Exp Pathol 2021 Dec;4(1):4. [doi: 10.1186/s42047-020-00084-x]
44. Sanguedolce F, Mazzucchelli R, Falagario UG, et al. Diagnostic biomarkers in renal cell tumors according to the latest WHO classification: a focus on selected new entities. Cancers (Basel) 2024 May 13;16(10):1856. [doi: 10.3390/cancers16101856] [Medline: 38791935]
45. Lam JS, Shvarts O, Leppert JT, Pantuck AJ, Figlin RA, Belldegrun AS. Postoperative surveillance protocol for patients with localized and locally advanced renal cell carcinoma based on a validated prognostic nomogram and risk group stratification system. J Urol 2005 Aug;174(2):466-472. [doi: 10.1097/01.ju.0000165572.38887.da] [Medline: 16006866]
46. Potti A, George DJ. Tyrosine kinase inhibitors in renal cell carcinoma. Clin Cancer Res 2004 Sep 15;10(18 Pt 2):6371S-6376S. [doi: 10.1158/1078-0432.CCR-050014] [Medline: 15448033]
47. Schöffski P, Dumez H, Clement P, et al. Emerging role of tyrosine kinase inhibitors in the treatment of advanced renal cell cancer: a review. Ann Oncol 2006 Aug;17(8):1185-1196. [doi: 10.1093/annonc/mdj133] [Medline: 16418310]
48. Xu W, Atkins MB, McDermott DF. Checkpoint inhibitor immunotherapy in kidney cancer. Nat Rev Urol 2020 Mar;17(3):137-150. [doi: 10.1038/s41585-020-0282-3] [Medline: 32020040]
49. Rassy E, Flippot R, Albiges L. Tyrosine kinase inhibitors and immunotherapy combinations in renal cell carcinoma. Ther Adv Med Oncol 2020;12:1758835920907504. [doi: 10.1177/1758835920907504] [Medline: 32215057]
50. Braun DA, Bakouny Z, Hirsch L, et al. Beyond conventional immune-checkpoint inhibition - novel immunotherapies for renal cell carcinoma. Nat Rev Clin Oncol 2021 Apr;18(4):199-214. [doi: 10.1038/s41571-020-00455-z] [Medline: 33437048]
51. Ravaud A. Treatment-associated adverse event management in the advanced renal cell carcinoma patient treated with targeted therapies. Oncologist 2011;16 Suppl 2(Suppl 2):32-44. [doi: 10.1634/theoncologist.2011-S2-32] [Medline: 21346038]
52. Kamli H, Li L, Gobe GC. Limitations to the therapeutic potential of tyrosine kinase inhibitors and alternative therapies for kidney cancer. Ochsner J 2019;19(2):138-151. [doi: 10.31486/toj.18.0015] [Medline: 31258426]
53. Espi M, Teuma C, Novel-Catin E, et al. Renal adverse effects of immune checkpoints inhibitors in clinical practice: ImmuNoTox study. Eur J Cancer 2021 Apr;147:29-39. [doi: 10.1016/j.ejca.2021.01.005] [Medline: 33607383]
54. Barrueto L, Caminero F, Cash L, Makris C, Lamichhane P, Deshmukh RR. Resistance to checkpoint inhibition in cancer immunotherapy. Transl Oncol 2020 Mar;13(3):100738. [doi: 10.1016/j.tranon.2019.12.010] [Medline: 32114384]
55. Sato K, Tsuchiya N, Sasaki R, et al. Increased serum levels of vascular endothelial growth factor in patients with renal cell carcinoma. Jpn J Cancer Res 1999 Aug;90(8):874-879. [doi: 10.1111/j.1349-7006.1999.tb00829.x] [Medline: 10543260]
56. Schips L, Dalpiaz O, Lipsky K, et al. Serum levels of vascular endothelial growth factor (VEGF) and endostatin in renal cell carcinoma patients compared to a control group. Eur Urol 2007 Jan;51(1):168-173. [doi: 10.1016/j.eururo.2006.06.026] [Medline: 16844285]
57. Beuselinck B, Vano YA, Oudard S, et al. Prognostic impact of baseline serum C-reactive protein in patients with metastatic renal cell carcinoma (RCC) treated with sunitinib. BJU Int 2014 Jul;114(1):81-89. [doi: 10.1111/bju.12494] [Medline: 24215209]
58. Pantuck AJ, Zeng G, Belldegrun AS, Figlin RA. Pathobiology, prognosis, and targeted therapy for renal cell carcinoma: exploiting the hypoxia-induced pathway. Clin Cancer Res 2003 Oct 15;9(13):4641-4652. [Medline: 14581333]
59. Yu L, Qiu Y, Tong X. Ferroptosis in renal cancer therapy: a narrative review of drug candidates. Cancers (Basel) 2024 Sep 11;16(18):3131. [doi: 10.3390/cancers16183131] [Medline: 39335103]
60. Kruk L, Mamtimin M, Braun A, et al. Inflammatory networks in renal cell carcinoma. Cancers (Basel) 2023 Apr 9;15(8):2212. [doi: 10.3390/cancers15082212] [Medline: 37190141]
61. Yang JW, Wan S, Li KP, Chen SY, Yang L. Gut and urinary microbiota: the causes and potential treatment measures of renal cell carcinoma. Front Immunol 2023;14:1188520. [doi: 10.3389/fimmu.2023.1188520] [Medline: 37441065]
62. Sun M, Shariat SF, Cheng C, et al. Prognostic factors and predictive models in renal cell carcinoma: a contemporary review. Eur Urol 2011 Oct;60(4):644-661. [doi: 10.1016/j.eururo.2011.06.041] [Medline: 21741163]
63. Hsieh JJ, Le V, Cao D, Cheng EH, Creighton CJ. Genomic classifications of renal cell carcinoma: a critical step towards the future application of personalized kidney cancer care with pan-omics precision. J Pathol 2018 Apr;244(5):525-537. [doi: 10.1002/path.5022] [Medline: 29266437]
64. Knudsen JE, Rich JM, Ma R. Artificial intelligence in pathomics and genomics of renal cell carcinoma. Urol Clin North Am 2024 Feb;51(1):47-62. [doi: 10.1016/j.ucl.2023.06.002] [Medline: 37945102]
65. Papadia G, Pacella M, Perrone M, Giliberti V. A comparison of different topic modeling methods through a real case study of Italian customer care. Algorithms 2023;16(2):94. [doi: 10.3390/a16020094]

Abbreviations

LDA: latent Dirichlet allocation
MeSH: Medical Subject Headings
RCC: renal cell carcinoma

XSL•FO
RenderX

TKI: tyrosine kinase inhibitor

Edited by N Cahill; submitted 10.Jun.2025; peer-reviewed by G Fan, L Raymond Guo, N Bievre; revised version received 31.Oct.2025; accepted 31.Oct.2025; published 16.Jan.2026.

Please cite as:
De La Hoz-M J, Montes-Escobar K, Salas-Macias CA, Fors M, Ballaz SJ
Using Latent Dirichlet Allocation Topic Modeling to Uncover Latent Research Topics and Trends in Renal Cell Carcinoma: Bibliometric Review
JMIR Cancer 2026;12:e78797
URL: https://cancer.jmir.org/2026/1/e78797
doi:10.2196/78797

© Javier De La Hoz-M, Karime Montes-Escobar, Carlos Alfredo Salas-Macias, Martha Fors, Santiago J Ballaz. Originally published in JMIR Cancer (https://cancer.jmir.org), 16.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Co-Ablation System for Pain Management in Bone Metastases: Retrospective Exploratory Study

Hang Yuan[1], MD; Wei-Li Xia[1], MD, PhD; Ho-Young Song[1], MD, PhD; Lin Zheng[1], MD, PhD; Wei-Jun Fan[2], MD, PhD; Hong-Tao Hu[1], MD, PhD

[1]The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, Henan, China

[2]Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, China

Corresponding Author:
Hong-Tao Hu, MD, PhD
The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital
No. 127, Dongming Road
Zhengzhou, Henan, 450008
China
Phone: 86 13592616373
Email: huhongtaogy@163.com

Abstract

Background: Bone metastasis is a prevalent complication of malignant tumors, often resulting in restricted mobility, severe pain, and diminished quality of life.

Objective: This study aimed to assess the analgesic effect and safety of a co-ablation system that combines elements of hyperthermic ablation and cryoablation in patients with bone metastases.

Methods: This retrospective study included patients with histologically confirmed painful bone metastases treated with the co-ablation system between January and October 2024. Pain intensity was assessed using the numerical rating scale (NRS), and functional status was evaluated using the Karnofsky Performance Status score at baseline and during a 12-week follow-up. Technical success, adverse events, analgesic use, and pain response (defined as a ≥2-point reduction in NRS) were analyzed.

Results: Nine patients were included. Technical success was achieved in all procedures, with no procedure-related adverse events observed. Median NRS scores decreased progressively from 5 (IQR 4-6) at baseline to 2 (IQR 1-3) at 12 weeks. By the fourth week, 8 (88.9%) patients achieved a clinically meaningful pain reduction. Karnofsky Performance Status scores showed gradual improvement during follow-up. No patients required escalation of analgesic therapy, and some experienced dose reductions.

Conclusions: The co-ablation system appeared to be feasible and was associated with short-term pain reduction in patients with bone metastases, with no ablation-related adverse events observed.

(JMIR Cancer 2026;12:e86301) doi:10.2196/86301

KEYWORDS

bone metastases; pain management; co-ablation system; cryoablation; thermal ablation

Introduction

Bone metastasis is a prevalent complication of malignant tumors, often resulting in restricted mobility, severe pain, and diminished quality of life [1]. Pain represents the most frequent symptom of bone metastasis [2]. Inadequate pain management is particularly prevalent among patients with bone metastases; approximately 80% of cancer-related pain originates from bone metastases, and as many as 25% of patients report insufficient symptom control [3]. In addition, chronic pain can cause depression and anxiety [4,5]. Currently, numerous treatment options exist for managing pain associated with bone metastases, including systemic therapies (chemotherapy, hormonal therapy, radiopharmaceuticals, and bisphosphonates), various analgesics, local interventions (radiotherapy, laser-induced interstitial thermotherapy, and ablation therapy), and surgical approaches [6-8].

Currently, the most prevalent form of local ablation used in analgesic treatment includes techniques such as radiofrequency ablation, cryoablation, and microwave ablation. These methods have been extensively validated for their safety and analgesic effect [9,10]. However, each modality has inherent limitations. Hyperthermic ablation techniques (radiofrequency ablation and microwave ablation) may be constrained by heat-sink effects

XSL•FO
RenderX

and carry a risk of thermal injury to adjacent structures [11,12], whereas cryoablation, although allowing improved visualization of the ice ball, does not permit needle tract ablation.

Chinese scientists have developed a co-ablation system that combines elements of hyperthermic ablation and cryoablation, using liquid nitrogen and alcohol as thermal media. The system is capable of achieving temperatures as low as −196 °C in freezing mode and up to approximately 80 °C in heating mode. The system operates by sequential application of freezing and heating, delivering extreme temperature changes to the target tissue. This design is intended to provide controlled thermal injury through alternating cold and heat exposure. Previous studies have demonstrated its effectiveness in cancer treatment [13,14].

However, there is currently limited literature documenting the use of a co-ablation system for the analgesic treatment of bone metastases. To address this knowledge gap, we conducted a retrospective exploratory study to assess the effectiveness and safety of the co-ablation system in managing pain associated with bone metastases.

Methods

Ethical Considerations

This was a retrospective study approved by the ethics committee of Henan Cancer Hospital (2025-KY-0026). Informed consent was waived by the ethics committee. All patients signed the procedural informed consent form and the general informed consent for anonymous use of their information in research before receiving treatment. During statistical analysis, all patient information was anonymized. Therefore, this study does not involve informed consent, privacy, or compensation.

Participants

From January to October 2024, we retrospectively analyzed patients with histologically confirmed bone metastases treated with the co-ablation system.

The exclusion criteria included the following: (1) survival time less than 3 months, (2) lack of a definitive pathological diagnosis, (3) the ablation site was not a bone metastasis, (4) absence of pain before ablation, and (5) loss to follow-up.

Procedure

According to the tumor size, appropriate co-ablation needles (RBL20, RCS17, and RAL26; Hygea Medical Technology Co) were selected preoperatively, and their functionality was verified using normal saline. Patients were positioned in a supine, prone, or lateral position on the computed tomography (CT) examination bed (GE Large Aperture CT; GE Healthcare Bio-Sciences Corp). A conventional CT scan was initially performed to identify the optimal puncture pathway, which was then marked and subsequently sterilized. Under local anesthesia, the bone metastatic lesion was gradually punctured, and the correct positioning was confirmed before connecting the co-ablation system (Hygea Medical Technology Co) for ablation treatment. Relevant parameters, including freezing time and rewarming time, were adjusted according to the extent of the tumor. Immediate postablation CT imaging was performed to evaluate technical adequacy. The assessment included confirmation of complete coverage of the target lesion by the ablation zone; appropriate extension beyond the lesion margins when feasible; and the absence of procedure-related adverse events (AEs), such as hemorrhage, pathological fracture, or injury to adjacent critical structures.

Data Collection and Follow-Up

Patient pain was assessed using a numerical rating scale (NRS) [15], and the Karnofsky Performance Status (KPS) score was used to evaluate patients' physical condition. NRS scores were recorded before the procedure and at 1 day, 1 week, 4 weeks, 8 weeks, and 12 weeks after the procedure. Pain assessment primarily targeted pain related to the treated bone metastatic lesion. Patients were instructed to focus on pain originating from the target lesion identified for the co-ablation system. KPS scores were recorded before the procedure and 1, 4, 8, and 12 weeks after the procedure. For patients outside the hospital, follow-up assessments were conducted via telephone interviews.

Age, gender, primary tumor type, location and number of bone metastases, administration of bone-modifying agents, use of analgesic drugs, dose adjustments, procedure-related AEs, and other concomitant treatments were reviewed through the electronic medical record system. All procedure-related AEs were recorded and classified according to the Society of Interventional Radiology AE classification system [16]. AEs were graded as mild, moderate, severe, life-threatening, or fatal based on clinical severity, imaging findings, and the requirement for medical intervention or prolonged hospitalization. Technical success and pain response were assessed for each procedure.

Technical success was defined as successful placement of the needle into the target lesion and completion of the planned protocol without technical failure or premature termination. Clinical success was predefined as a clinically meaningful pain reduction, defined as a decrease of ≥2 points in the NRS score compared with baseline [17]. Patients meeting this criterion at a given follow-up time point were classified as responders, and the corresponding proportion was reported as the responder rate.

Statistical Analysis

Statistical analyses were performed using GraphPad Prism software (version 9.5; GraphPad Software). Continuous and ordinal variables were presented as median and IQR. Changes in NRS and KPS scores across multiple time points were analyzed using the Friedman test. When the Friedman test indicated a statistically significant difference, post hoc comparisons between baseline and follow-up time points were performed using Dunn multiple comparisons test with adjustment for multiple testing. All tests were 2-sided, and an adjusted $P<.05$ was considered statistically significant.

Results

Participant Characteristics

From January 2024 to October 2024, a total of 16 patients with bone metastases were treated with the co-ablation system, among whom 2 died within 3 months after treatment, 4 had primary

XSL•FO
RenderX

lesions, and 1 was excluded due to loss of follow-up data. Finally, 9 patients were included in the analysis (Figure 1), including 6 male patients and 3 female patients, with a median age of 60 (IQR 55-69) years. Primary malignancies included hepatocellular carcinoma (n=4, 44%), lung adenocarcinoma (n=4, 44%), and colon adenocarcinoma (n=1, 11%). All patients had a single bone metastasis lesion. Metastatic involvement was primarily located in the femur (n=3, 33%), ilium (n=4, 44%), and ribs (n=2, 22%). Treatment regimens varied among patients: 1 (11%) patient received denosumab injections, 5 (56%) patients were treated with zoledronic acid, and 3 (33%) patients received ibandronate. All patients continued to undergo systemic antitumor therapy (Table 1).

Figure 1. Flow diagram of patient enrollment and study inclusion. A total of 16 patients with bone metastases underwent treatment using the co-ablation system. Of these, 2 patients died within 3 months after ablation, 4 patients received ablation for primary lesions, and 1 patient had incomplete clinical data. The remaining 9 patients met the inclusion criteria and were included in the final analysis.

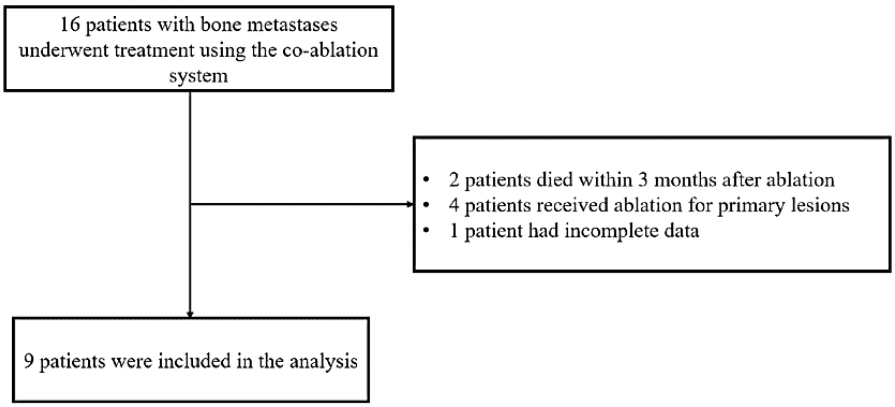


Table 1. Patient characteristics (N=9).

Characteristic	Values
Age (years), median (IQR)	60 (55-69)
Sex, n (%)	
Male	6 (67)
Female	3 (33)
Primary tumor type, n (%)	
Hepatocellular carcinoma	4 (44)
Lung adenocarcinoma	4 (44)
Colon adenocarcinoma	1 (11)
Metastatic site, n (%)	
Femur	3 (33)
Ilium	4 (44)
Rib	2 (22)
Weight-bearing lesion	7 (78)
Baseline NRS[a] score, median (IQR)	5 (4-6)
Baseline KPS[b] score, median (IQR)	60 (50-70)
Baseline analgesic use, n (%)	9 (100)
Use of bone-modifying agents, n (%)	
Denosumab	1 (11)
Zoledronic acid	5 (56)
Ibandronate	3 (33)

[a]NRS: numerical rating scale.

[b]KPS: Karnofsky Performance Status.

XSL•FO
RenderX

Pain Intensity Outcomes

Pain intensity assessed using the NRS showed a significant change over time (P<.001). An overall comparison using the Friedman test demonstrated a statistically significant difference in NRS scores across all assessed time points (baseline, 1 day, 1 week, 4 weeks, 8 weeks, and 12 weeks after treatment; P<.001).

Median NRS scores decreased from 5 (IQR 4-6) at baseline to 4 (IQR 3-5) at 1 day and 3 (IQR 2-4) at 1 week after treatment. Further reductions were observed at 4 weeks (median 2, IQR 2-3) and 8 weeks (median 2, IQR 1-2), with improvement maintained at 12 weeks (median 2, IQR 1-3; Table 2).

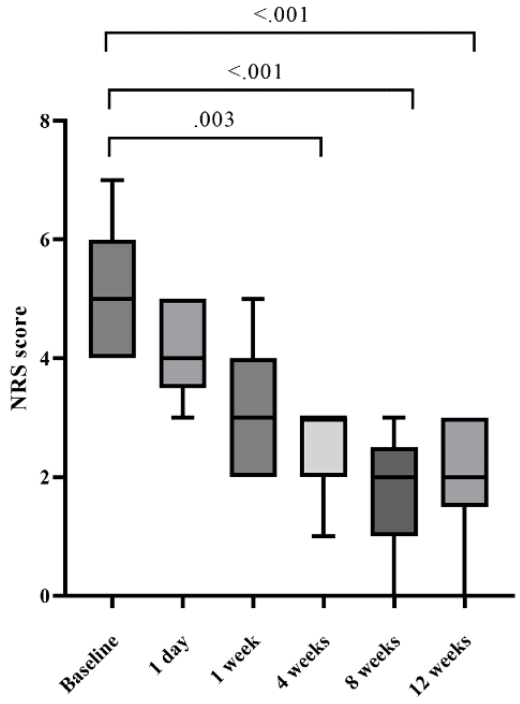
Table 2. Changes in the numerical rating scale and Karnofsky Performance Status scores over time.

Time point	Numerical rating scale, median (IQR)	Karnofsky Performance Status, median (IQR)
Baseline	5 (4-6)	60 (50-70)
1 day after treatment	4 (4-5)	—[a]
1 week after treatment	3 (2-4)	60 (50-80)
4 weeks after treatment	3 (2-3)	70 (60-80)
8 weeks after treatment	2 (1-2)	70 (70-90)
12 weeks after treatment	2 (2-3)	70 (60-90)

[a]Not available.

Compared with baseline, NRS scores were significantly lower at 1 week (adjusted P=.003), 8 weeks (adjusted P<.001), and 12 weeks (adjusted P<.001; Figure 2). By week 4, among the 9 patients, a clinically meaningful pain reduction (≥2-point decrease in NRS) was observed in 8.

Figure 2. Changes in pain intensity (numerical rating scale [NRS] scores) over time after the procedure. Box-and-whisker plots show NRS scores at baseline and at 1 day, 1 week, 4 weeks, 8 weeks, and 12 weeks after treatment. Data are presented as median, IQR (box), and range (whiskers). Overall differences across time points were assessed using the Friedman test. Post hoc comparisons between baseline and follow-up time points were performed using the Dunn multiple comparisons test with adjustment for multiple testing. Adjusted P values for statistically significant comparisons are shown.



Functional Status Outcomes

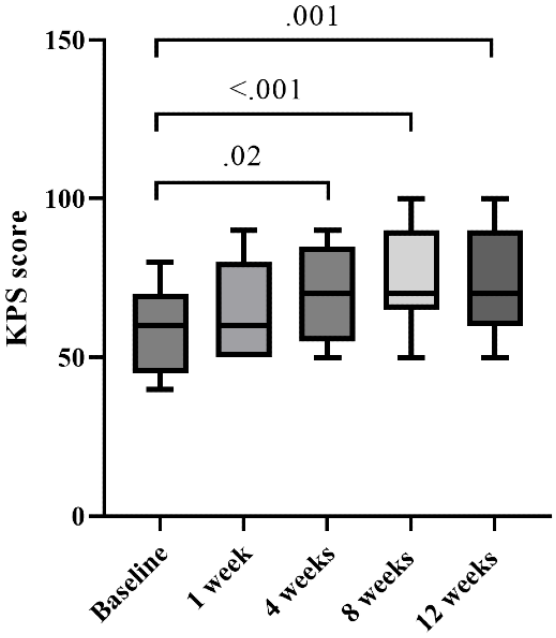
Functional performance, as measured by the KPS score, improved over the course of follow-up. A significant difference in KPS scores was observed across baseline and subsequent assessment time points (P<.001).

At baseline, patients demonstrated limited functional status, with a median KPS score of 60 (IQR 50-70). An increase in KPS was observed at 1 week after treatment (median 60, IQR 50-80); however, this early change was not statistically significant (adjusted P=.54). In contrast, KPS scores showed significant improvement at later follow-up visits, reaching a median of 70 (IQR 60-80) at 4 weeks and 70 (IQR 70-90) at 8 weeks, with these improvements maintained through 12 weeks (median 70, IQR 60-90; Table 2).

XSL•FO
RenderX

Compared with baseline, KPS scores were significantly higher at 4 weeks (adjusted *P*=.02), 8 weeks (adjusted *P*<.001), and 12 weeks (adjusted *P*=.001; Figure 3).

Figure 3. Changes in functional status (Karnofsky Performance Status [KPS] scores) during follow-up. Box-and-whisker plots illustrate KPS scores at baseline and at 1, 4, 8, and 12 weeks after co-ablation. Values are expressed as median, IQR (box), and range (whiskers). Longitudinal changes were analyzed using the Friedman test, followed by the Dunn post hoc test with correction for multiple comparisons. Adjusted *P* values for significant differences compared with baseline are indicated.



AEs and Analgesic Dosage

All patients underwent successful treatment with the co-ablation system, achieving a technical success rate of 100% without any procedure-related AEs. At the 12-week follow-up, no patients switched analgesic agents, 3 patients experienced a 25% reduction in their analgesic dosage, 1 patient experienced a 50% reduction, and no patients required an increase in medication dosage.

Discussion

In this retrospective single-arm exploratory study, the use of the co-ablation system was feasible and well tolerated in patients with bone metastases, with no procedure-related AEs observed. A reduction in pain intensity was noted over the follow-up period, with NRS scores showing a statistically significant overall decrease and most patients achieving a clinically meaningful pain reduction by week 4. Improvements in functional status, as reflected by KPS scores, were observed at later follow-up time points; however, these findings should be interpreted cautiously given the small sample size and study design. No patients required escalation of analgesic therapy during follow-up, and some were able to reduce their baseline analgesic doses. Collectively, these preliminary findings support the technical feasibility and short-term safety of the co-ablation system for palliative pain management in bone metastases.

Several studies have reported favorable pain control outcomes following ablation for bone metastases. Ma et al [18] evaluated microwave ablation, radiofrequency ablation, and cryoablation in patients with bone metastases from non-small cell lung cancer and observed a substantial reduction in pain scores at 4 weeks after treatment. Consistent with these findings, our study demonstrated a marked decrease in NRS scores over time, with median pain scores declining from baseline to 4 weeks after the procedure and most patients achieving a clinically meaningful pain reduction. Although an early reduction in pain was observed as soon as 1 day after treatment in our cohort, this change did not reach statistical significance, which may be related to the small sample size. In contrast, Yang et al [19] reported significant pain relief within 1 day following cryoablation; however, pain was assessed using a visual analog scale rather than NRS, which may limit direct comparability. Additionally, Tomasian et al [20] reported sustained pain relief at 3 months after ablation for painful vertebral metastases, a finding that aligns with the maintained reduction in NRS scores observed through 12 weeks in our study. Taken together, these observations suggest that the co-ablation system may offer pain control outcomes comparable to established ablation modalities, although direct comparisons should be interpreted cautiously.

High-intensity focused ultrasound is a noninvasive image-guided ablation technique that has demonstrated a favorable safety profile and promising analgesic effects in patients with painful bone metastases [21]. Previous prospective studies have shown that magnetic resonance–guided high-intensity focused ultrasound (MR-HIFU) can achieve rapid and clinically meaningful pain reduction without treatment-related AEs [22,23]. Bongiovanni et al [22] reported complete or partial pain responses in all treated patients at 30 days, whereas Napoli et al [23] observed significant improvements in pain severity and interference, with additional evidence of local tumor response in some cases. Although MR-HIFU differs from the co-ablation system evaluated in this study in terms of

XSL•FO
RenderX

invasiveness and ablation mechanism, both approaches support the role of targeted local ablation in pain palliation. In our cohort, short-term pain relief without ablation-related AEs was observed, which is consistent with the analgesic and safety trends reported for MR-HIFU.

Previous studies have shown that multimodality approaches combining local ablation with radiotherapy, vertebroplasty, or iodine-125 seed implantation can improve pain control in patients with bone metastases [24-26]. These findings suggest that co-ablation could potentially be incorporated into multimodal palliative strategies in future practice and investigation. In our study, we also observed that the dosage of analgesic drugs was significantly reduced in some patients following the procedure. This finding is consistent with previous reports indicating a reduction in analgesic requirements after cryoablation [19,20,24,27].

Following treatment with the co-ablation system, patients demonstrated a gradual improvement in functional status, as reflected by increasing KPS scores over the follow-up period. Although no statistically significant improvement was observed at 1 week after treatment, KPS scores improved significantly from 4 weeks onward and remained stable through 12 weeks. These findings are in line with the MOTION (Multicenter Study of Cryoablation for Palliation of Painful Bone Metastases) study, which reported that cryoablation provided rapid pain relief while functional status was maintained and quality of life improved during follow-up [28].

This study is limited by its small sample size and retrospective, single-arm design; these factors reduce statistical power, limit generalizability, and preclude adjustment for pain-related confounders. Patient heterogeneity in primary tumors and metastatic sites, as well as the absence of a comparative control group, prevents assessment of relative efficacy vs established ablation modalities. The 12-week follow-up period was sufficient for short-term pain assessment but inadequate for evaluating the durability of response, pain recurrence, or delayed AEs. Pain from untreated disease sites may have influenced NRS and KPS scores, and the use of KPS without patient-reported outcome measures provides limited insight into overall quality of life.

In this small retrospective exploratory study, the co-ablation system was technically feasible and well tolerated in patients with painful bone metastases. Pain reduction was observed during short-term follow-up, and functional performance was maintained or gradually improved, with no procedure-related AEs recorded. These preliminary findings support further evaluation of the co-ablation system in larger, prospective controlled studies.

Funding

This research was funded by the National Key Research and Development Program of China (grants 2023YFC2414006 and 2023YFC2414002; June 2023-June 2026), the Henan Province Outstanding Foreign Expert Workstation Project (grant GZS2022020; January 2023-December 2025), the Foreign Expert Project of the Ministry of Science and Technology of China (grant G2023026016L; January 2023-December 2024), the Medical Education Research Project of Henan Province (grant Wjlx2021334; January 2022-December 2024), and the General Project of the Natural Science Foundation of Henan Province (grant 212300410403; January 2023-December 2024).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

HY: writing—original draft preparation, formal analysis, investigation, and data curation
WLX: writing—review and editing
LZ: funding acquisition, resources, visualization, and writing—review and editing
WJF: funding acquisition and writing—review and editing
HYS: funding acquisition and writing—review and editing
HTH: supervision, funding acquisition, writing—review and editing, conceptualization, methodology, and validation
All authors critically reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Smith H, Mohsin I. Painful boney metastases. Korean J Pain 2013 Jul;26(3):223-241 [FREE Full text] [doi: 10.3344/kjp.2013.26.3.223] [Medline: 23861996]
2. Gbané M, Soglo JC, Diomandé M, Kengni GL, Ouattara B, Kouassi JM, et al. Etiologies of bone metastases at the rheumatology department (UTH) of Abidjan. Open J Rheumatol Autoimmune Dis 2020;10:1-7. [doi: 10.4236/ojra.2020.101001]

XSL•FO
RenderX

3. Filippiadis DK, Tutton S, Kelekis A. Percutaneous bone lesion ablation. Radiol Med 2014 Jul 4;119(7):462-469. [doi: 10.1007/s11547-014-0418-8] [Medline: 24894924]
4. Massie MJ, Holland JC. The cancer patient with pain: psychiatric complications and their management. J Pain Symptom Manage 1992 Feb;7(2):99-109 [FREE Full text] [doi: 10.1016/0885-3924(92)90121-w] [Medline: 1573292]
5. Spiegel D, Sands S, Koopman C. Pain and depression in patients with cancer. Cancer 1994 Nov 01;74(9):2570-2578. [doi: 10.1002/1097-0142(19941101)74:9<2570::aid-cncr2820740927>3.0.co;2-3] [Medline: 7923013]
6. Callstrom MR, Atwell TD, Charboneau JW, Farrell MA, Goetz MP, Rubin J, et al. Painful metastases involving bone: percutaneous image-guided cryoablation--prospective trial interim analysis. Radiology 2006 Nov;241(2):572-580. [doi: 10.1148/radiol.2412051247] [Medline: 17057075]
7. Li F, Wang W, Li L, Chang Y, Su D, Guo G, et al. An effective therapy to painful bone metastases: cryoablation combined with zoledronic acid. Pathol Oncol Res 2014 Oct 30;20(4):885-891 [FREE Full text] [doi: 10.1007/s12253-014-9769-7] [Medline: 24781761]
8. Zeng L, Chow E, Bedard G, Zhang L, Fairchild A, Vassiliou V, et al. Quality of life after palliative radiation therapy for patients with painful bone metastases: results of an international study validating the EORTC QLQ-BM22. Int J Radiat Oncol Biol Phys 2012 Nov 01;84(3):e337-e342. [doi: 10.1016/j.ijrobp.2012.05.028] [Medline: 22763028]
9. Khanmohammadi S, Noroozi A, Yekaninejad MS, Rezaei N. Cryoablation for the palliation of painful bone metastasis: a systematic review. Cardiovasc Intervent Radiol 2023 Nov 11;46(11):1469-1482. [doi: 10.1007/s00270-022-03356-z] [Medline: 36631660]
10. Gennaro N, Sconfienza LM, Ambrogi F, Boveri S, Lanza E. Thermal ablation to relieve pain from metastatic bone disease: a systematic review. Skeletal Radiol 2019 Aug 09;48(8):1161-1169. [doi: 10.1007/s00256-018-3140-0] [Medline: 30627778]
11. Primavesi F, Swierczynski S, Klieser E, Kiesslich T, Jäger T, Urbas R, et al. Thermographic real-time-monitoring of surgical radiofrequency and microwave ablation in a perfused porcine liver model. Oncol Lett 2018 Mar 18;15(3):2913-2920 [FREE Full text] [doi: 10.3892/ol.2017.7634] [Medline: 29435018]
12. Thurlow PC, Azhideh A, Ho CK, Stratchko LM, Pooyan A, Alipour E, et al. Thermal protection techniques for image-guided musculoskeletal ablation. Radiographics 2025 Apr 01;45(4):e240078. [doi: 10.1148/rg.240078] [Medline: 40048387]
13. Yang W, An Y, Li Q, Liu C, Zhu B, Huang Q, et al. Co-ablation versus cryoablation for the treatment of stage III-IV non-small cell lung cancer: a prospective, noninferiority, randomized, controlled trial (RCT). Thorac Cancer 2021 Feb 14;12(4):475-483 [FREE Full text] [doi: 10.1111/1759-7714.13779] [Medline: 33319493]
14. Liu J, Zhou Y, Yu T, Gui L, Deng Z, Lv Y. Minimally invasive probe system capable of performing both cryosurgery and hyperthermia treatment on target tumor in deep tissues. Minim Invasive Ther Allied Technol 2004 Feb 10;13(1):47-57. [doi: 10.1080/13645700310022691] [Medline: 16754125]
15. Farrar JT, Young JPJ, LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001 Nov;94(2):149-158. [doi: 10.1016/S0304-3959(01)00349-9] [Medline: 11690728]
16. Baerlocher MO, Nikolic B, Sze DY. Adverse event classification: clarification and validation of the society of interventional radiology specialty-specific system. J Vasc Interv Radiol 2023 Jan;34(1):1-3. [doi: 10.1016/j.jvir.2022.10.011] [Medline: 36244632]
17. Hovenier R, Huele EH, Bindels BJ, van Oss A, Kuijten RH, Verlaan JJ, et al. Pain response of patients treated with radiotherapy for painful bone metastases between 2013 and 2024 - an analysis of the prospective real-world PRESENT cohort. Radiother Oncol 2026 Jan;214:111263 [FREE Full text] [doi: 10.1016/j.radonc.2025.111263] [Medline: 41187792]
18. Ma Y, Wallace AN, Waqar SN, Morgensztern D, Madaelil TP, Tomasian A, et al. Percutaneous image-guided ablation in the treatment of osseous metastases from non-small cell lung cancer. Cardiovasc Intervent Radiol 2018 May 4;41(5):726-733. [doi: 10.1007/s00270-017-1843-6] [Medline: 29204695]
19. Yang Y, Li Y, Wu Y, Qiu S, Liu C, Wang Q, et al. Retrospective analysis of CT-guided percutaneous cryoablation for treatment of painful osteolytic bone metastasis. Cryobiology 2020 Feb 01;92:203-207. [doi: 10.1016/j.cryobiol.2020.01.012] [Medline: 31958427]
20. Tomasian A, Wallace A, Northrup B, Hillen TJ, Jennings JW. Spine cryoablation: pain palliation and local tumor control for vertebral metastases. AJNR Am J Neuroradiol 2015 Oct 01;37(1):189-195. [doi: 10.3174/ajnr.a4521]
21. Scipione R, Anzidei M, Bazzocchi A, Gagliardo C, Catalano C, Napoli A. HIFU for bone metastases and other musculoskeletal applications. Semin Intervent Radiol 2018 Oct 05;35(4):261-267 [FREE Full text] [doi: 10.1055/s-0038-1673363] [Medline: 30402009]
22. Bongiovanni A, Foca F, Oboldi D, Diano D, Bazzocchi A, Fabbri L, et al. 3-T magnetic resonance-guided high-intensity focused ultrasound (3 T-MR-HIFU) for the treatment of pain from bone metastases of solid tumors. Support Care Cancer 2022 Jul 24;30(7):5737-5745. [doi: 10.1007/s00520-022-06990-y] [Medline: 35332373]
23. Napoli A, Anzidei M, Marincola BC, Brachetti G, Ciolina F, Cartocci G, et al. Primary pain palliation and local tumor control in bone metastases treated with magnetic resonance-guided focused ultrasound. Invest Radiol 2013;48(6):351-358. [doi: 10.1097/rli.0b013e318285bbab]

XSL•FO
RenderX

24. Di Staso M, Gravina GL, Zugaro L, Bonfili P, Gregori L, Franzese P, et al. Treatment of solitary painful osseous metastases with radiotherapy, cryoablation or combined therapy: propensity matching analysis in 175 patients. PLoS One 2015 Jun 23;10(6):e0129021 [FREE Full text] [doi: 10.1371/journal.pone.0129021] [Medline: 26103516]
25. David E, Kaduri S, Yee A, Chow E, Sahgal A, Chan S, et al. Initial single center experience: radiofrequency ablation assisted vertebroplasty and osteoplasty using a bipolar device in the palliation of bone metastases. Ann Palliat Med 2017 Apr;6(2):118-124 [FREE Full text] [doi: 10.21037/apm.2016.12.02] [Medline: 28249543]
26. Zhang Q, Zhang K, Xie B, Ren Y, Li G, Zhang L, et al. Analysis of curative effect of I125 implantation combined with radiofrequency ablation in treating bone metastases. J Bone Oncol 2018 Jun;11:23-26 [FREE Full text] [doi: 10.1016/j.jbo.2018.01.001] [Medline: 29892521]
27. Zugaro L, DI Staso M, Gravina GL, Bonfili P, Gregori L, Franzese P, et al. Treatment of osteolytic solitary painful osseous metastases with radiofrequency ablation or cryoablation: a retrospective study by propensity analysis. Oncol Lett 2016 Mar;11(3):1948-1954 [FREE Full text] [doi: 10.3892/ol.2016.4108] [Medline: 26998106]
28. Jennings JW, Prologo JD, Garnon J, Gangi A, Buy X, Palussière J, et al. Cryoablation for palliation of painful bone metastases: the MOTION multicenter study. Radiol Imaging Cancer 2021 Mar 01;3(2):e200101 [FREE Full text] [doi: 10.1148/rycan.2021200101] [Medline: 33817650]

Abbreviations

AE: adverse event
CT: computed tomography
KPS: Karnofsky Performance Status
MOTION: Multicenter Study of Cryoablation for Palliation of Painful Bone Metastases
MR-HIFU: magnetic resonance–guided high-intensity focused ultrasound
NRS: numerical rating scale

Edited by A Schwartz, M Balcarras; submitted 22.Oct.2025; peer-reviewed by KW Liang, S Vanni; comments to author 13.Jan.2026; revised version received 04.Feb.2026; accepted 04.Feb.2026; published 19.Feb.2026.

Please cite as:
Yuan H, Xia WL, Song HY, Zheng L, Fan WJ, Hu HT
Co-Ablation System for Pain Management in Bone Metastases: Retrospective Exploratory Study
JMIR Cancer 2026;12:e86301
URL: https://cancer.jmir.org/2026/1/e86301
doi:10.2196/86301
PMID:

©Hang Yuan, Wei-Li Xia, Ho-Young Song, Lin Zheng, Wei-Jun Fan, Hong-Tao Hu. Originally published in JMIR Cancer (https://cancer.jmir.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Quality of Life, Symptom Burden, and Associated Factors Among Patients With Lung Cancer in Sub-Saharan Africa: Cross-Sectional Study

Mpho Ratshikana[1,2,3], MPhil; Kija Malale[4,5], PhD; Abdul-Rauf Sayed[6], MSc; Oluwatosin Ayeni[7,8], PhD; Peedi Mathobela[2], MSc; Rofhiwa Mathiba[2,3], MMed; Charmaine Blanchard[8], PhD; Merika Tsitsi[1,2,3], MBBCh; Daniel Osei-Fofie[9], MBChB; Anita Graham[3,10], MBBCh; Herry Owuor Dhudha[5], MSc; Nestory Masalu[5,11], MMed

[1]Department of Internal Medicine, Chris Hani Baragwanath Academic Hospital, Johannesburg, Gauteng, South Africa

[2]Soweto Comprehensive Cancer Center, Chris Hani Baragwanath Hospital, Johannesburg, Gauteng, South Africa

[3]Department of Internal Medicine, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, Gauteng, South Africa

[4]Department of Clinical Nursing, Catholic University of Health and Allied Sciences, Mwanza, United Republic of Tanzania

[5]Department of Oncology, Bugando Medical Centre, Mwanza, United Republic of Tanzania

[6]Bristol-Myers Squibb Foundation (BMSF), Johannesburg, Gauteng, South Africa

[7]Department of Radiation Oncology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, Gauteng, South Africa

[8]Strengthening Oncology Services Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, Gauteng, South Africa

[9]Department of Oncology, Robert Mangaliso Sobukwe Hospital, Kimberly, Northern Cape, South Africa

[10]Division of Medical Oncology, University of Witwatersrand Faculty of Health Sciences, Johannesburg, Gauteng, South Africa

[11]Department of Oncology, Catholic University of Health and Allied Sciences, Mwanza, United Republic of Tanzania

Corresponding Author:
Kija Malale, PhD
Department of Clinical Nursing
Catholic University of Health and Allied Sciences
Wurzburg Road 35
Mwanza, 33109
United Republic of Tanzania
Phone: 255 754731576
Email: kmalale53@gmail.com

Abstract

Background: Lung cancer remains a major contributor to cancer mortality in sub-Saharan Africa (SSA), where late diagnosis, driven by low awareness, sociocultural barriers, and health system constraints, limits effective treatment. Despite the growing burden, evidence on patients' quality of life (QoL) and symptom experience in SSA is limited.

Objective: This study aimed to describe the common symptoms and QoL of patients with lung cancer treated at 2 hospitals in SSA, and to investigate the association of demographics, clinical characteristics, and symptom burden with QoL.

Methods: This was a cross-sectional study that consecutively recruited patients with lung cancer from 2 teaching hospitals in SSA: Bugando Medical Centre (BMC) in Tanzania and the University of the Witwatersrand Centre of Respiratory Excellence (WITS-CORE) in South Africa. Data collected included demographics, clinical information, and performance status using the Eastern Cooperative Oncological Group Performance Scale (ECOG-PS). Health-related QoL was assessed using the 30-item European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30). The study followed all ethical procedures, and data were analyzed using both descriptive and inferential statistics in Stata 18. A *P* value of <.05 was considered statistically significant.

Results: A total of 174 patients with lung cancer were enrolled across the 2 sites. The score on the EORTC QLQ-C30 global health status/QoL subscale was low, with a median of 41.67 (IQR 33.33-41.67), and it varied by site. Patients from WITS-CORE demonstrated higher social functioning scores, while those from BMC reported greater financial difficulties. A low global health status/QoL score was independently associated with the BMC site (adjusted odds ratio [aOR] 3.5, 95% CI 1.3-9.3) and poor performance status (ECOG-PS 3-4; aOR 3.4, 95% CI 1.2-6.6). Furthermore, symptoms such as nausea and vomiting, pain, dyspnea, insomnia, appetite loss, diarrhea, and financial struggles were all associated with a low global health status/QoL score.

XSL•FO
RenderX

Conclusions: QoL among patients with lung cancer in SSA is poor. Low QoL is strongly associated with the Multinational Lung Cancer Control Program study site, poor performance status, and a range of symptoms and financial difficulties. Addressing these factors may help to improve patient outcomes and well-being in SSA.

(JMIR Cancer 2026;12:e87232) doi:10.2196/87232

KEYWORDS

cancer; lung; quality of life; QoL; symptoms

Introduction

GLOBOCAN estimates indicated that more than 19 million new cancer cases occurred worldwide in 2022, with lung cancer accounting for 12.4%, making it the most diagnosed cancer [1]. Although the incidence of lung cancer in Africa is lower than in high-income countries, it remains among the top 5 cancers in both sexes, and it is the fourth leading cause of cancer-related mortality, with southern Africa bearing the highest burden on the continent [1]. In sub-Saharan Africa (SSA), most patients present at advanced stages, partly due to limited community awareness, alongside the influence of sociocultural factors, health system constraints, and gaps in health care provider knowledge [2-5]. These challenges contribute to delayed diagnosis, limited treatment options, and consequently high mortality rates [6].

Lung cancer is associated with a substantial symptom burden, including fatigue, dyspnea, cough, pain, appetite loss, and hemoptysis. However, in SSA, late presentation often results in more severe complications, such as brain metastases, pleural effusion, spinal cord compression, and superior vena cava obstruction syndrome [7,8]. As a result, treatment is frequently limited to palliative approaches, including chemotherapy and radiotherapy, aimed primarily at improving quality of life (QoL).

QoL is a critical outcome in cancer care, particularly for patients with advanced disease [9-11]. Lung cancer significantly compromises physical, emotional, social, and spiritual well-being, with strong evidence linking QoL to symptom burden, clinical characteristics, and sociodemographic factors [12,13]. Importantly, better QoL is associated with improved health outcomes and survival [9,14]. Interventions such as chemotherapy, targeted therapy, immunotherapy, radiotherapy, and palliative care have demonstrated benefits in symptom control, functional status, and overall well-being [10,15]. However, most evidence on symptom burden and QoL originates from high-income settings, with limited data from SSA.

This study aimed to describe the common symptoms and QoL of patients with lung cancer receiving care at 2 tertiary institutions in SSA: the University of the Witwatersrand (WITS) Centre for Respiratory Excellence (CORE) in South Africa and Bugando Medical Centre (BMC) in Tanzania. The study also examined the association between sociodemographic and clinical characteristics, symptom burden, and QoL. We hypothesized that patients with lung cancer in SSA would experience a high symptom burden and poor QoL, with significant associations between these outcomes and patient-related factors.

Methods

Ethical Considerations

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki, including respect for individuals, their right to make informed decisions, and the protection of their health and rights. Ethics approval was obtained from the Human Research Ethics Committee (Medical), WITS (M180436). In Tanzania, the ethics approval was obtained from the Joint Catholic University of Health and Allied Sciences and BMC Research Ethics and Review Committee (CREC/278/2018) and the National Institute for Medical Research (MR/53/100/598). Informed consent was obtained from all participants before participating in this study. Participants' privacy and confidentiality were strictly maintained throughout the study by anonymizing data and securely handling all research information.

Setting

The study was conducted at 2 teaching hospitals participating in the Multinational Lung Cancer Control Program (MLCCP) in 2 SSA countries: Tanzania and South Africa. WITS-CORE is a specialized respiratory unit providing comprehensive lung cancer services to patients in Johannesburg, South Africa. It is based at Helen Joseph Hospital, a teaching hospital affiliated with the WITS. BMC is a consultant-teaching and referral hospital for the Lake and Western zone of the United Republic of Tanzania affiliated with the Catholic University of Health and Allied Sciences in Mwanza, Tanzania. The hospital's Oncology department is the only cancer center in the Lake and Western zone of United Republic of Tanzania that provides comprehensive cancer care with outreach services to urban and rural areas.

Design and Participants

This cross-sectional study consecutively recruited patients with lung cancer from 2 teaching hospitals in SSA. The inclusion criteria were patients aged 18 years or older, a confirmed primary lung cancer diagnosis, and being physically and mentally able to participate in the study. A similar protocol was applied across the study sites; once the diagnosis was confirmed and before treatment began, patients were invited to participate in the study. Those who consented were enrolled, and sociodemographic and clinical information were collected. Thereafter, the questionnaire was administered in English (WITS-CORE) or Swahili (BMC) and, where needed, translated into the vernacular. Data collection was conducted for 1 year, from August 2020 to August 2021.

XSL•FO
RenderX

Study Measures

Trained nurses and interviewers conducted interviews following the standardized protocol to obtain demographic information (age, gender, marital status, education level, employment status, smoking, and mining history) and clinical information (histology, clinical staging, Eastern Cooperative Oncological Group Performance Scale [ECOG-PS], weight loss, active or current tuberculosis, and other comorbidities). The ECOG-PS was used to determine the performance status of patients [16].

QoL, functional status, symptom burden, and financial impact were assessed using the 30-item European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30) [17]. The EORTC QLQ-C30 is composed of 3 multi-item or single-item subscales: the global health status/QoL scale, 5 functional scales (physical, role, social, emotional, and cognitive functioning), 8 symptom scales (fatigue, pain, nausea and vomiting, dyspnea, insomnia, appetite loss, constipation, and diarrhea), and a financial impact scale. Each item has four possible response alternatives: (1) “not at all,” (2) “a little,” (3) “quite a bit,” and (4) “very much.” The responses to the scale items refer to “last week,” with the exception of the patient’s physical performance scale, where the timeframe is the present. The scores of each scale were calculated in accordance with the EORTC QLQ-C 30 scoring manual [11]. The sum of items in each category was added, and the total was divided by the number of questions in the category. A linear transformation was then undertaken to convert the score to a percentage scale. All the scales and single-item measures range from 0 to 100. Higher scores on the functional and QoL scales translated to better health-related QoL, whereas higher scores on the symptom scales translated to a higher level of symptoms or problems.

Data Analysis

Study data were collected and managed using REDCap tools hosted at WITS. Patients’ demographics, clinical characteristics, and health-related QoL information were extracted from the REDCap database and exported to Stata (version 18; StataCorp) for statistical analysis [18,19]. Categorical variables were presented as frequency tables, and continuous variables were presented as descriptive measures, expressed as medians and interquartile ranges (IQRs). The nonparametric Wilcoxon rank-sum test was used to compare the QoL scores between study sites. Univariate and multivariate binary logistic regression analysis were performed to evaluate the association between QoL and sociodemographic factors, clinical characteristics, and symptom burden. Consistent with previous studies, QoL was dichotomized using the median (50th percentile) as the cut-off point: a score greater than or equal to 50 was defined as “above-average QoL,” while a score less than 50 was defined as “below-average QoL” [20]. Odds ratios (ORs) were used to test the association between variables, and 95% CI that did not span unity were considered as thresholds of statistical significance. Adjusted ORs (aORs) were used in multivariate analysis.

Results

Sociodemographic Characteristics

A total of 174 patients with lung cancer consented and were enrolled across the 2 sites: 35 (20.1%) at BMC, and 139 (79.9%) at WITS-CORE. Table 1 shows the sociodemographic characteristics of the participants. Most patients were male (n=115, 66.1%), unemployed or retired (n=135, 80%), with a mean age of 61.6 (SD 11.6) years. There were significant differences in the demographic characteristics between the patients at the 2 MLCCP sites. More WITS-CORE patients were male (97/139, 69.8%) than those at BMC (18/35, 51.4%; P=.04). More than two-thirds of the patients with lung cancer at WITS-CORE (98/139, 70.5%) had a high school or higher level of education compared with participants at BMC (8/35, 22.9%; P<.001). A significant proportion of WITS-CORE patients (108/139, 77.7%) were current or ex-smokers compared with patients at BMC (8/35, 22.9%; P<.001).

XSL•FO
RenderX

Table 1. Sociodemographic characteristics of the participants.

Sociodemographic variables	BMC[a] (n=35)	WITS-CORE[b] (n=139)	Total (n=174)	*P* value
Age group (y), n (%)				
<40	4 (11.4)	5 (3.6)	9 (5.2)	—[c]
40-49	4 (11.4)	10 (7.2)	14 (8)	—
50-59	10 (28.6)	32 (23)	42 (24.1)	—
60-69	11 (31.4)	56 (40.3)	67 (38.5)	—
≥70	6 (17.1)	36 (25.9)	42 (24.1)	—
Age (y), mean (SD)	58.0 (13.4)	62.5 (12.0)	61.6 (11.6)	.04
Sex, n (%)				
Female	17 (48.6)	42 (30.2)	59 (33.9)	—
Male	18 (51.4)	97 (69.8)	115 (66.1)	.04
Educational level, n (%)				
≤Primary education	27 (77.1)	41 (29.5)	68 (39.1)	—
≥High school	8 (22.9)	98 (70.5)	106 (60.9)	<.001
Marital status, n (%)				
Married or in partnership	29 (82.9)	79 (58.5)	108 (62.1)	—
Single, divorced, or widowed	6 (17.1)	60 (41.5)	66 (37.9)	<.001
Occupation, n (%)				
Employed	10 (28.6)	29 (20.9)	39 (20)	—
Unemployed or retired	25 (71.4)	110 (79.1)	135 (80)	.33
Smoking, n (%)				
Never smoked	27 (77.1)	31 (22.3)	58 (33.3)	—
Current or ex-smoker	8 (22.9)	108 (77.7)	116 (66.7)	<.001
Ever a miner, n (%)				
Yes	0 (0.0)	19 (13.7)	19 (10.9)	—
No	35 (100)	120 (86.3)	155 (89.1)	.02

[a]BMC: Bugando Medical Centre.

[b]WITS-CORE: University of the Witwatersrand Center of Respiratory Excellence.

[c]Not applicable.

Clinical Characteristics

Most patients (157/174, 90.2%) were diagnosed with non–small cell lung cancer (NSCLC). Most presented with late-stage (III-IV) diseases, including 96.7% (148/153) of the patients with NSCLC and 85.7% (18/21) of those with small cell lung cancer. Only 4% (5/126) of the WITS-CORE patients with NSCLC presented with early stage (I-II) disease (Table 2). Most patients at BMC (30/35, 85.7%) presented with weight loss compared with patients at WITS-CORE (94/139, 67.6%; *P*=.04). Similarly, a significant proportion of BMC patients had current tuberculosis (12/35, 34.3%) compared with those at WITS-CORE (5/139, 3.6%; *P*<.001). A higher proportion of WITS-CORE patients presented with one or more comorbidities (81/139, 58.3%) compared with patients at BMC (14/35, 40%; *P*<.001).

XSL•FO
RenderX

Table 2. Clinical characteristics of participants.

Clinical variables	BMC[a] (n=35), n (%)	WITS-CORE[b] (n=139), n (%)	Total (n=174), n (%)	*P* value
Clinical staging				
NSCLC[c] stage I-II[d]	0 (0)	5 (4)	5 (3.3)	—[e]
NSCLC stage III-IV	27 (100)	121 (96)	148 (96.7)	.29
SCLC[f] (limited)[g]	0 (0.0)	3 (23.1)	3 (14.3)	—
SCLC (extensive)	8 (100)	10 (76.9)	18 (85.7)	.14
ECOG[h] performance status				
0-2	23 (65.7)	100 (72.7)	123 (70.7)	—
3-4	12 (34.3)	39 (27.3)	51 (29.3)	.47
Weight loss				
No	5 (14.3)	45 (32.4)	50 (28.7)	—
Yes	30 (85.7)	94 (67.6)	124 (71.3)	.04
Current tuberculosis				
No	23 (65.7)	134 (96.4)	157 (90.2)	—
Yes	12 (34.3)	5 (3.6)	17 (9.8)	<.001
HIV status				
HIV negative	35 (100)	117 (84.2)	152 (87.4)	—
HIV positive	0 (0)	22 (15.8)	22 (12.6)	.01
Comorbidity				
No	21 (60)	58 (41.7)	79 (45.4)	—
Yes	14 (40)	81 (58.3)	95 (54.6)	.05

[a]BMC: Bugando Medical Centre.

[b]WITS-CORE: University of the Witwatersrand Center of Respiratory Excellence.

[c]NSCLC: non–small cell lung cancer.

[d]For NSCLC staging, denominators were 27 (BMC), 126 (WITS-CORE), and 153 (total).

[e]Not applicable.

[f]SCLC: small cell lung cancer.

[g]For SCLC staging, denominators were 8 (BMC), 13 (WITS-CORE), and 21 (total).

[h]ECOG: Eastern Cooperative Oncological Group.

QoL, Functional Status, and Symptom Burden

Overall, the EORTC QLQ-C30 global health status/QoL subscale score was low (Table 3); the median score was 41.7 (IQR 33.33-41.67). WITS-CORE patients had higher QoL (*P*=.004) and social functioning scores (*P*<.001) compared with patients at BMC. Financial difficulty was significantly higher among BMC patients compared with WITS-CORE (*P*<.001). Overall, the highest symptom scores reported were pain (66.67, IQR 33.33-66.67) and fatigue (55.56, IQR 44.44-55.56). Pain and insomnia scores were significantly higher among patients at BMC, with median scores of 83.33 (IQR 50.00-83.33; *P*<.001) and 66.67 (IQR 33.33-66.67; *P*<.001), respectively, compared with patients at WITS-CORE.

XSL•FO
RenderX

Table 3. Quality of life, functional status, and symptom burden of patients with lung cancer at 2 teaching hospitals across sub-Saharan Africa.

Assessment aspect	Total (n=174), median (IQR)	BMC[a] (n=35), median (IQR)	WITS-CORE[b] (n=139), median (IQR)	*P* value[c]
Global health status/quality of life	41.67 (33.33-41.67)	33.33 (16.67-33.33)	41.67 (33.33-41.67)	.004
Functional scales				
Physical functioning	46.67 (26.67-46.67)	66.67 (6.67-66.67)	46.67 (33.33-46.67)	.58
Role functioning	33.33 (16.67-33.33)	50.00 (0.00-50.00)	33.33 (16.67-33.33)	.45
Emotional functioning	66.67 (50.00-66.67)	75.00 (41.67-75.00)	66.67 (50.00-66.67)	.76
Cognitive functioning	83.33 (50.00-83.33)	83.33 (50.00-83.33)	83.33 (50.00-83.33)	.30
Social functioning	33.33 (0.00-33.33)	0.00 (0.00-0.00)	50.00 (33.33-50.00)	<.001
Symptoms scale				
Fatigue	55.56 (44.44-55.56)	55.56 (33.33-55.56)	55.56 (44.44-55.56)	.80
Nausea and vomiting	0.00 (0.00-0.00)	16.67 (0.00-16.67)	0.00 (0.00-0.00)	<.001
Pain	66.67 (33.33-66.67)	83.33 (50.00-83.33)	50.00 (33.33-50.00)	<.001
Dyspnea	33.33 (33.33-33.33)	33.33 (0.00-33.33)	33.33 (33.33-33.33)	.36
Insomnia	33.33 (0.00-33.33)	66.67 (33.33-66.67)	33.33 (0.00-33.33)	<.001
Appetite loss	33.33 (0.00-33.33)	33.33 (0.00-33.33)	33.33 (0.00-33.33)	.09
Constipation	0.00 (0.00-0.00)	0.00 (0.00-0.00)	0.00 (0.00-0.00)	<.001[d]
Diarrhea	0.00 (0.00-0.00)	0.00 (0.00-0.00)	0.00 (0.00-0.00)	.01[d]
Financial impact scale				
Financial difficulties	66.67 (33.33-66.67)	100.00 (100.00-100.00)	66.67 (33.33-66.67)	<.001

[a]BMC: Bugando Medical Centre.

[b]WITS-CORE: University of the Witwatersrand Centre of Respiratory Excellence.

[c]Significant at *P*<.003 (Bonferroni corrected).

[d]Despite identical medians and IQRs in both groups for constipation and diarrhea, the statistical test revealed a significant difference (*P*<.001) due to the presence of small but systematic differences in the distribution of values that are not captured by median or IQR.

Associations Between Sociodemographics, Clinical Characteristics, and QoL

The results of the multiple logistic regression analysis using the dependent variable (QoL) as a binary outcome based on the 50th percentile or median (median QoL score <50 and median QoL score ≥50) are shown in Table 4. A significantly higher proportion of patients with lung cancer at BMC had below-average QoL compared with patients at WITS-CORE (aOR 3.5, 95% CI 1.3-9.3). Poor ECOG-PS score (3-4) was associated with poorer QoL score (aOR 3.4, 95% CI 1.2-6.6).

XSL•FO
RenderX

Table 4. Associations between sociodemographics, clinical characteristics, and quality of life (QoL) of patients with lung cancer at 2 teaching hospitals across sub-Saharan Africa.

Variables	Below-average QoL (<50; n=97), n (%)	Above-average QoL (≥50; n=77), n (%)	OR[a] (95% CI)	*P* value	aOR[b] (95% CI)[c]
Age group (y)					
<60	38 (39.2)	27 (35.1)	Reference	—[d]	Reference
≥60	59 (60.8)	50 (64.9)	1.1 (0.6-1.9)	.84	1.1 (0.5-2.1)
Sex					
Male	64 (66)	51 (66.2)	Reference	—	Reference
Female	33 (34)	26 (33.8)	1.1 (0.6-1.9)	.99	1.0 (0.5-2.0)
Education					
≤Primary education	45 (46.4)	23 (29.9)	1.8 (0.9-3.5)	.08	1.8 (1.0-3.1)
≥Secondary education	52 (53.6)	54 (70.1)	Reference	—	Reference
Occupation					
Employed	18 (18.6)	21 (27.3)	Reference	—	Reference
Unemployed or retired	79 (81.4)	56 (72.7)	1.4 (0.7-2.8)	.22	0.6 (0.3-1.4)
Smoking					
Never smoked	35 (36.1)	23 (29.9)	Reference	—	Reference
Current or ex-smoker	62 (63.9)	54 (70.1)	0.9 (0.5-4.5)	.06	2.0 (0.9-4.3)
MLCCP[e] study sites					
WITS-CORE[f]	9 (9.3)	10 (13)	Reference	—	Reference
BMC[g]	88 (90.7)	67 (87)	2.8 (1.3-6.3)	.01	3.5 (1.3-9.3)[h]
Clinical staging					
I-II	3 (3.6)	2 (2.9)	Reference	—	Reference
III-IV	81 (96.4)	67 (97.1)	1.1 (0.3-4.5)	.62	1.6 (0.2-11.2)
ECOG[i] performance status					
0-2	60 (61.9)	63 (81.8)	Reference	—	Reference
3-4	37 (38.1)	14 (18.2)	2.8 (1.4-6.7)	.01	3.4 (1.2-6.6)[h]
Comorbidity					
No	44 (45.4)	35 (45.5)	Reference	—	Reference
Yes	53 (55.6)	42 (54.5)	1.3 (0.7-2.2)	.94	0.9 (0.5-1.9)

[a]OR: odds ratio.

[b]aOR: adjusted odds ratio.

[c]Adjusted for age group, gender, education, occupation, smoking, Multinational Lung Cancer Control Program study site, clinical staging, Eastern Cooperative Oncological Group performance status, and comorbidity.

[d]Not applicable.

[e]MLCCP: Multinational Lung Cancer Control Program.

[f]WITS-CORE: University of the Witwatersrand Centre of Respiratory Excellence.

[g]BMC: Bugando Medical Centre.

[h]Statistically significant.

[i]ECOG: Eastern Cooperative Oncological Group.

Association Between Symptom Burden and QoL

Symptom scores were significantly higher among patients with below-average QoL (<50) for fatigue, pain, dyspnea, insomnia, and appetite loss when compared with above-average QoL (≥50; Table 5). Higher symptom scores were associated with poor QoL.

XSL•FO
RenderX

Table 5. Associations between symptom burden and quality of life (QoL) of patients with lung cancer at 2 teaching hospitals across sub-Saharan Africa.

	Below-average QoL (<50; n=97), median (IQR)	Above-average QoL (≥50; n=77), median (IQR)	*P* value
Fatigue	66.67 (44.44-66.67)	44.44 (33.33-44.44)	<.001
Nausea and vomiting	0.00 (0.00-0.00)	0.00 (0.00-0.00)	.01[a]
Pain	66.67 (50.00-100.00)	50.00 (33.33-50.00)	<.001
Dyspnea	66.67 (33.33-66.67)	33.33 (0.00-33.33)	.004
Insomnia	66.67 (33.33-66.67)	33.33 (0.00-33.33)	<.001
Appetite loss	33.33 (0.00-66.67)	33.33 (0.00-33.33)	.001[a]
Constipation	0.00 (0.00-0.00)	0.00 (0.00-0.00)	.17
Diarrhea	0.00 (0.00-0.00)	0.00 (0.00-0.00)	.05
Financial difficulties	66.67 (33.33-66.67)	66.67 (33.33-66.67)	.01[a]

[a]Despite identical medians and IQRs in both groups (for nausea and vomiting, appetite loss, and financial difficulties), the statistical test revealed a significant difference ($P<.005$) due to the presence of small but systematic differences in the distribution of values—particularly in the frequency or magnitude of nonzero observations—that are not captured by median or IQR.

Discussion

Principal Findings

This study provides important insights into the QoL, symptom burden, and associated factors among patients with lung cancer across 2 sites in SSA. Overall, the findings demonstrated that QoL among participants was markedly low (median 41.67, IQR 33.33-41.67), substantially below the EORTC reference population, and varied between BMC and WITS-CORE [21]. Consistent with existing literature, a higher symptom burden, particularly pain, fatigue, dyspnea, insomnia, and appetite loss, was strongly associated with poorer QoL. Significant intersite differences were also observed in sociodemographic characteristics, clinical profiles, functional status, symptom experience, and financial challenges.

Regarding QoL domains, physical and role functioning were notably impaired, reflecting patients' reduced capacity to perform daily activities, maintain independence, and fulfill social and family responsibilities. These findings are consistent with prior studies indicating that disease progression significantly compromises functional status. As lung cancer advances, patients often require assistance with basic activities, such as mobility, personal hygiene, and household responsibilities, which in turn affects their social roles and overall well-being [22]. Interestingly, cognitive functioning in this study appeared relatively preserved despite the high symptom burden. This aligns with some previous studies suggesting that cognitive domains may remain intact compared with physical and emotional domains [11,23-26]. However, this finding contrasts with evidence indicating that symptoms such as insomnia can negatively affect cognitive performance, while better cognitive functioning may help patients cope more effectively with stress and symptom distress [27]. This discrepancy highlights the need for more nuanced and longitudinal assessments of cognitive outcomes in this population.

The comparison between the 2 sites revealed both similarities and differences. Participants at both BMC and WITS-CORE were predominantly male, unemployed, and commonly presented without active tuberculosis, with many diagnosed at earlier stages (stage I-II). However, WITS-CORE participants generally had higher levels of education, were more likely to be smokers, and had a higher prevalence of comorbidities. These differences may reflect broader contextual factors, such as urban-rural disparities, health care accessibility, and diagnostic capacity. For example, the urban setting of WITS-CORE may facilitate earlier detection and access to specialized services, including advanced imaging and oncology expertise. In contrast, BMC serves a larger rural population, where geographical barriers, limited resources, and delayed health-seeking behaviors may influence disease presentation and outcomes. Previous studies have highlighted distance to health care facilities as a major barrier to timely cancer diagnosis and care in SSA [28].

Among sociodemographic factors, education level emerged as a significant determinant of QoL, while other variables, such as age, sex, and economic status, showed inconsistent associations, reflecting mixed findings in the literature [11,21,29,30]. Higher education may enhance health literacy, enabling patients to better understand their condition, seek timely care, adhere to treatment, and engage in self-management [31]. The lower education levels observed among BMC participants, who largely came from rural communities, underscore the need for targeted health education interventions delivered in local languages to improve awareness, early detection, and ultimately QoL outcomes.

Symptom burden in this study was substantial, particularly among patients with advanced disease, and aligned with global evidence identifying fatigue, pain, dyspnea, insomnia, and appetite loss as the most prevalent and distressing symptoms in lung cancer [7,11,22,32,33]. Notably, pain and insomnia were more severe among patients at BMC, possibly reflecting disparities in access to pain management and supportive care services. In many settings in SSA, effective pain control remains a challenge due to regulatory restrictions on opioid availability, limited health care provider training, and financial constraints faced by patients [34]. Addressing these barriers through policy reforms, task sharing in opioid prescribing, and capacity building

XSL•FO
RenderX

among health care providers is essential. Given the strong association between symptom burden and poor QoL observed in this study, integrating palliative care into routine cancer services is critical to ensure effective symptom management and improve patient outcomes [35-37].

Financial hardship was widely reported, particularly among BMC participants, and is consistent with the high unemployment rate observed in the study population. Although financial difficulties were not statistically associated with QoL in this analysis, extensive evidence indicates that economic burden significantly affects access to care, treatment adherence, and overall well-being [21,25,34,38,39]. Financial challenges may include direct health care costs as well as indirect costs, such as transportation, nutrition, and loss of income. Differences in social protection systems, such as disability grants available in South Africa, may partly explain variations between sites. Expanding financial support mechanisms for patients with life-limiting illnesses across SSA could play a crucial role in improving QoL.

Limitations

This study has several limitations. The relatively small sample size at BMC may limit generalizability and statistical power. Although a standardized tool (EORTC QLQ-C30) was used, differences in interview languages across sites may have influenced data consistency. Variability in diagnostic and staging capabilities between facilities may also affect the accuracy of clinical classification. Additionally, some relevant factors such as cough, treatment-related effects, nutritional status, use of alternative therapies, and polypharmacy were not assessed but may influence QoL. The cross-sectional design further limits the ability to establish causal relationships or assess changes over time. Future prospective or longitudinal studies are therefore recommended to better understand causal pathways and temporal dynamics of QoL and symptom burden. Despite these limitations, this study represents one of the first multicountry analyses of QoL and symptom burden among patients with lung cancer in SSA. The findings underscore the urgent need for integrated, patient-centered cancer care that prioritizes symptom management, strengthens health system capacity, addresses socioeconomic barriers, and incorporates palliative care as a core component of oncology services.

Conclusions

In consistency with previous studies, this study revealed disparities in QoL and symptom burden among patients with lung cancer in SSA. Overall, patients experienced low QoL, functional impairment, and financial impact, with prevalent symptoms such as pain, fatigue, dyspnea, insomnia, and appetite loss. Variations in education levels, health care access, and site-specific factors further underscored inequalities in cancer care across the region. These findings highlight the pressing need for context-sensitive interventions that strengthen palliative care services, enhance cancer health literacy through community-based education, and expand social and financial support for patients. Collectively, the study offers valuable regional evidence to guide policies and practices aimed at improving the QoL and clinical outcomes of patients with lung cancer in SSA.

Acknowledgments

The authors acknowledge the support received from the clinical teams at Bugando Medical Centre (Tanzania) and University of the Witwatersrand Centre for Palliative Care (South Africa), as well as patients and caregivers who participated in the study. ChatGPT (GPT-4; OpenAI) was used solely to support English language editing of the manuscript to enhance clarity and readability.

Funding

The research was supported through grants from the Bristol Myers Squibb Foundation. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

MR and KM contributed to the conception and design of the study, acquisition, analysis, and interpretation of data and drafted and revised the manuscript. AG, NM, CB, and HOD contributed to the acquisition of funding, design of the study, overall supervision of the project, interpretation of data, and review of the manuscript. ARS, PM, and OA contributed to data management, analysis, interpretation of the data, and revision of the manuscript. RM, MT, DO-F, NM, and HOD contributed to the analysis, interpretation of data, and revision of the manuscript. All authors approved the final manuscript and agreed to be personally accountable for their contributions.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263 [FREE Full text] [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Dlamini SB, Sartorius B, Ginindza TG. Knowledge, attitudes and practices towards lung cancer among adults in KwaZulu-Natal, South Africa: a cross-sectional survey. J Public Health Afr 2022 Oct 04;13(3):2111 [FREE Full text] [doi: 10.4081/jphia.2022.2111] [Medline: 36313926]
3. Lubuzo B, Ginindza T, Hlongwana K. The barriers to initiating lung cancer care in low-and middle-income countries. Pan Afr Med J 2020 Feb 12;35:38 [FREE Full text] [doi: 10.11604/pamj.2020.35.38.17333] [Medline: 32499854]
4. Cassim S, Chepulis L, Keenan R, Kidd J, Firth M, Lawrenson R. Patient and carer perceived barriers to early presentation and diagnosis of lung cancer: a systematic review. BMC Cancer 2019 Jan 08;19(1):25 [FREE Full text] [doi: 10.1186/s12885-018-5169-9] [Medline: 30621616]
5. Mapanga W, Norris SA, Chen WC, Blanchard C, Graham A, Baldwin-Ragaven L, et al. Consensus study on the health system and patient-related barriers for lung cancer management in South Africa. PLoS One 2021 Feb 11;16(2):e0246716 [FREE Full text] [doi: 10.1371/journal.pone.0246716] [Medline: 33571312]
6. Tsai CH, Kung PT, Kuo WY, Tsai WC. Effect of time interval from diagnosis to treatment for non-small cell lung cancer on survival: a national cohort study in Taiwan. BMJ Open 2020 Apr 22;10(4):e034351 [FREE Full text] [doi: 10.1136/bmjopen-2019-034351] [Medline: 32327476]
7. Iyer S, Taylor-Stokes G, Roughley A. Symptom burden and quality of life in advanced non-small cell lung cancer patients in France and Germany. Lung Cancer 2013 Aug;81(2):288-293. [doi: 10.1016/j.lungcan.2013.03.008] [Medline: 23561304]
8. Iyer S, Roughley A, Rider A, Taylor-Stokes G. The symptom burden of non-small cell lung cancer in the USA: a real-world cross-sectional study. Support Care Cancer 2014 Jan;22(1):181-187. [doi: 10.1007/s00520-013-1959-4] [Medline: 24026981]
9. Montazeri A. Quality of life data as prognostic indicators of survival in cancer patients: an overview of the literature from 1982 to 2008. Health Qual Life Outcomes 2009 Dec 23;7:102 [FREE Full text] [doi: 10.1186/1477-7525-7-102] [Medline: 20030832]
10. Braun DP, Gupta D, Staren ED. Quality of life assessment as a predictor of survival in non-small cell lung cancer. BMC Cancer 2011 Aug 15;11:353 [FREE Full text] [doi: 10.1186/1471-2407-11-353] [Medline: 21843358]
11. Larsson M, Ljung L, Johansson BB. Health-related quality of life in advanced non-small cell lung cancer: correlates and comparisons to normative data. Eur J Cancer Care (Engl) 2012 Sep;21(5):642-649. [doi: 10.1111/j.1365-2354.2012.01346.x] [Medline: 22519874]
12. Avelino CU, Cardoso RM, de Aguiar SS, da Silva MJ. Assessment of quality of life in patients with advanced non-small cell lung carcinoma treated with a combination of carboplatin and paclitaxel. J Bras Pneumol 2015;41(2):133-142 [FREE Full text] [doi: 10.1590/S1806-37132015000004367] [Medline: 25972967]
13. Polanski J, Jankowska-Polanska B, Rosinczuk J, Chabowski M, Szymanska-Chabowska A. Quality of life of patients with lung cancer. Onco Targets Ther 2016 Feb 29;9:1023-1028 [FREE Full text] [doi: 10.2147/OTT.S100685] [Medline: 27013895]
14. Thronicke A, von Trott P, Kröz M, Grah C, Matthes B, Schad F. Health-related quality of life in patients with lung cancer applying integrative oncology concepts in a certified cancer centre. Evid Based Complement Alternat Med 2020 May 10;2020:5917382 [FREE Full text] [doi: 10.1155/2020/5917382] [Medline: 32454866]
15. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998 Jan;16(1):139-144. [doi: 10.1200/JCO.1998.16.1.139] [Medline: 9440735]
16. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol 1982 Dec;5(6):649-656. [doi: 10.1097/00000421-198212000-00014]
17. Fayers P, Bottomley A, EORTC Quality of Life Group, Quality of Life Unit. Quality of life research within the EORTC-the EORTC QLQ-C30. European Organisation for Research and Treatment of Cancer. Eur J Cancer 2002 Mar;38 Suppl 4:S125-S133. [doi: 10.1016/s0959-8049(01)00448-8] [Medline: 11858978]
18. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
19. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
20. Nguyen TX, Nguyen AH, Nguyen HT, Nguyen TT, Nguyen HL, Goldberg RJ, et al. Health-related quality of life among older adults with dementia living in Vietnamese nursing homes. Int J Environ Res Public Health 2024 Jan 25;21(2):135 [FREE Full text] [doi: 10.3390/ijerph21020135] [Medline: 38397626]
21. Scott NW, Fayers P, Aaronson NK, Bottomley A, de Graeff A, Groenvold M, EORTC Quality of Life Group. EORTC QLQ-C30 Reference Values Manual 2nd Edition. Brussels, Belgium: EORTC Quality of Life Group; 2008.
22. Togas C, Alexias G, Anagnostopoulos F. Evaluation of health-related quality of life in lung cancer patients in Greece and associations with demographic and medical characteristics. J Biosci Med 2019 Oct;07(10):17-31. [doi: 10.4236/jbm.2019.710002]

XSL•FO
RenderX

23. Hung HY, Wu LM, Chen KP. Determinants of quality of life in lung cancer patients. J Nurs Scholarsh 2018 May;50(3):257-264. [doi: 10.1111/jnu.12376] [Medline: 29569334]
24. Dai YL, Yang CT, Chen KH, Tang ST. Changes in and determinants of quality of life in patients with advanced non-small-cell lung cancer undergoing initial chemotherapy. J Nurs Res 2017 Jun;25(3):203-215. [doi: 10.1097/JNR.0000000000000148] [Medline: 28481816]
25. Koch M, Hjermstad MJ, Tomaszewski K, Tomaszewska I, Hornslien K, Harle A, et al. Gender effects on quality of life and symptom burden in patients with lung cancer: results from a prospective, cross-cultural, multi-center study. J Thorac Dis 2020 Aug;12(8):4253-4261 [FREE Full text] [doi: 10.21037/jtd-20-1054] [Medline: 32944337]
26. Chen ML, Yu CT, Yang CH. Sleep disturbances and quality of life in lung cancer patients undergoing chemotherapy. Lung Cancer 2008 Dec;62(3):391-400. [doi: 10.1016/j.lungcan.2008.03.016] [Medline: 18468718]
27. Porter LS, Keefe FJ, Garst J, McBride CM, Baucom D. Self-efficacy for managing pain, symptoms, and function in patients with lung cancer and their informal caregivers: associations with symptoms and distress. Pain 2008 Jul 15;137(2):306-315 [FREE Full text] [doi: 10.1016/j.pain.2007.09.010] [Medline: 17942229]
28. Tegegne TK, Chojenta C, Loxton D, Smith R, Kibret KT. The impact of geographic access on institutional delivery care use in low and middle-income countries: systematic review and meta-analysis. PLoS One 2018 Aug 30;13(8):e0203130 [FREE Full text] [doi: 10.1371/journal.pone.0203130] [Medline: 30161201]
29. Derogar M, van der Schaaf M, Lagergren P. Reference values for the EORTC QLQ-C30 Quality of Life Questionnaire in a random sample of the Swedish population. Acta Oncol 2012 Jan;51(1):10-16. [doi: 10.3109/0284186X.2011.614636] [Medline: 21961499]
30. Pierzynski JA, Ye Y, Lippman SM, Rodriguez MA, Wu X, Hildebrandt MA. Socio-demographic, clinical, and genetic determinants of quality of life in lung cancer patients. Sci Rep 2018 Jul 13;8(1):10640 [FREE Full text] [doi: 10.1038/s41598-018-25712-1] [Medline: 30006595]
31. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
32. Bülbül Y, Ozlu T, Arinc S, Ozyurek BA, Gunbatar H, Senturk A, et al. Assessment of palliative care in lung cancer in Turkey. Med Princ Pract 2017;26(1):50-56 [FREE Full text] [doi: 10.1159/000452801] [Medline: 27780164]
33. Iwase S, Kawaguchi T, Tokoro A, Yamada K, Kanai Y, Matsuda Y, et al. Assessment of cancer-related fatigue, pain, and quality of life in cancer patients at palliative care team referral: a multicenter observational study (JORTC PAL-09). PLoS One 2015 Aug 05;10(8):e0134022 [FREE Full text] [doi: 10.1371/journal.pone.0134022] [Medline: 26244975]
34. Namisango E, Allsop MJ, Powell RA, Friedrichsdorf SJ, Luyirika EB, Kiyange F, et al. Investigation of the practices, legislation, supply chain, and regulation of opioids for clinical pain management in Southern Africa: a multi-sectoral, cross-national, mixed methods study. J Pain Symptom Manage 2018 Mar;55(3):851-863 [FREE Full text] [doi: 10.1016/j.jpainsymman.2017.11.010] [Medline: 29155288]
35. Ferrell BR, Temel JS, Temin S, Alesi ER, Balboni TA, Basch EM, et al. Integration of palliative care into standard oncology care: American Society of Clinical Oncology Clinical Practice Guideline update. J Clin Oncol 2017 Jan;35(1):96-112. [doi: 10.1200/JCO.2016.70.1474] [Medline: 28034065]
36. Dans M, Smith T, Back A, Baker JN, Bauman JR, Beck AC, et al. NCCN guidelines insights: palliative care, version 2.2017. J Natl Compr Canc Netw 2017 Aug;15(8):989-997. [doi: 10.6004/jnccn.2017.0132] [Medline: 28784860]
37. Smith CB, Phillips T, Smith TJ. Using the new ASCO clinical practice guideline for palliative care concurrent with oncology care using the TEAM approach. Am Soc Clin Oncol Educ Book 2017;37:714-723 [FREE Full text] [doi: 10.1200/EDBK_175474] [Medline: 28561696]
38. Lathan CS, Cronin A, Tucker-Seeley R, Zafar SY, Ayanian JZ, Schrag D. Association of financial strain with symptom burden and quality of life for patients with lung or colorectal cancer. J Clin Oncol 2016 May 20;34(15):1732-1740 [FREE Full text] [doi: 10.1200/JCO.2015.63.2232] [Medline: 26926678]
39. Abegaz TM, Ayele AA, Gebresillassie BM. Health related quality of life of cancer patients in Ethiopia. J Oncol 2018 Apr 15;2018:1467595 [FREE Full text] [doi: 10.1155/2018/1467595] [Medline: 29849628]

Abbreviations

aOR: adjusted odds ratio
BMC: Bugando Medical Centre
CORE: Centre of Respiratory Excellence
ECOG-PS: Eastern Cooperative Oncological Group Performance Scale
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
MLCCP: Multinational Lung Cancer Control Program
NSCLC: non–small cell lung cancer
OR: odds ratio

XSL•FO
RenderX

QoL: quality of life
SSA: sub-Saharan Africa
WITS: University of the Witwatersrand

Edited by M Balcarras; submitted 06.Nov.2025; peer-reviewed by E Coz, X Dlamini; comments to author 27.Feb.2026; revised version received 20.Apr.2026; accepted 20.Apr.2026; published 16.Jul.2026.

Please cite as:
Ratshikana M, Malale K, Sayed AR, Ayeni O, Mathobela P, Mathiba R, Blanchard C, Tsitsi M, Osei-Fofie D, Graham A, Dhudha HO, Masalu N
Quality of Life, Symptom Burden, and Associated Factors Among Patients With Lung Cancer in Sub-Saharan Africa: Cross-Sectional Study
JMIR Cancer 2026;12:e87232
URL: https://cancer.jmir.org/2026/1/e87232
doi:10.2196/87232
PMID:

©Mpho Ratshikana, Kija Malale, Abdul-Rauf Sayed, Oluwatosin Ayeni, Peedi Mathobela, Rofhiwa Mathiba, Charmaine Blanchard, Merika Tsitsi, Daniel Osei-Fofie, Anita Graham, Herry Owuor Dhudha, Nestory Masalu. Originally published in JMIR Cancer (https://cancer.jmir.org), 16.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

An Information Tool Incorporating Real-World Outcome Data for Women With Metastatic Breast Cancer Eligible for Treatment With a CDK4/6 Inhibitor: Development and Evaluation

Ellen G Engelhardt[1*], PhD; Mariska Q N Hackert[1*], PhD; Anne Vogelaar[1], MSc; Afke F de Jong[2], MSc, MD; Cristina Guerrero Paez[3], MSc; Annette W G van der Velden[4], MD; Christa Putker[5], MSc; Mariette J Agterof[5], MD, PhD; Ewoudt M W van de Garde[6,7], MD, PhD; Cornelia F van Uden-Kraan[1], PhD

[1]Santeon (Netherlands), Herculesplein 102, Utrecht, The Netherlands

[2]Stichting Kijksluiter, Amsterdam, The Netherlands

[3]Borstkankervereniging Nederland, Utrecht, The Netherlands

[4]Department of Internal Medicine, Martini Ziekenhuis, Groningen, The Netherlands

[5]Department of Internal Medicine, St. Antonius Hospital, Utrecht/Nieuwegein, The Netherlands

[6]Department of Pharmacy, St. Antonius Hospital, Utrecht/Nieuwegein, The Netherlands

[7]Department of Pharmaceutical Sciences, Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht University, Utrecht, The Netherlands

[*]these authors contributed equally

Corresponding Author:
Cornelia F van Uden-Kraan, PhD
Santeon (Netherlands), Herculesplein 102, Utrecht, The Netherlands

Abstract

Background: Accelerating the transition to value-based health care (VBHC) is essential to ensure sustainable care delivery. VBHC maximizes patient value by optimizing outcomes, controlling costs, and leveraging data to improve quality and patient-doctor communication. Integrating real-world data in health care is important to achieve informed decisions, especially in palliative care, where choices are complex.

Objective: This study described the development and pilot-testing of an information tool that incorporates real-world outcome data for women with metastatic breast cancer who are initiating CDK4/6 inhibitor treatment. Real-world insights are needed to better inform these patients, in particular because real-world outcomes are known to differ from trial results because of larger heterogeneity in patient characteristics, and differences in frequency of check-ups and handling side effects.

Methods: We developed an information tool together with patient representatives and clinicians using a participatory development approach that consisted of five key steps: (1) establishment of a multidisciplinary steering group (n=10 steering group members), (2) mapping of the patient journey and patients' needs through focus groups (n=9) and semistructured interviews (n=8), (3) extraction of real-world outcome data from electronic health records systems of 229 patients, (4) prototyping of the tool (n=10), and (5) pilot evaluation with the targeted patient population using semistructured interviews (n=38). We used qualitative analysis methods to analyze the focus group and interview data.

Results: We developed a tool consisting of (1) a communication aid for use during doctor-patient consultations (ie, KIJKgesprek [Stichting Kijksluiter]) and (2) a 2-component companion app with informational videos for use at home, both incorporating real-world outcome data (ie, KIJKbericht and KIJKsluiter [Stichting Kijksluiter]). Participants valued the tool for its clarity and structured design, reporting that the outcome data reinforced their experiences and facilitated the setting of realistic expectations. However, some participants described the outcome data as overwhelming, underscoring the importance of careful framing and delivery. Preferences regarding the type, level of detail, and timing of information presentation varied among participants, highlighting the necessity of individualizing information tools to meet diverse informational needs.

Conclusions: Most patients valued the inclusion of real-world outcome data in the information tool, although many found it challenging to process. Preferences for the type and presentation of information varied widely among individuals. Information tools incorporating outcome data have the potential to enhance patient understanding and support informed decision-making about care that they value most. However, these tools must be designed to allow for customization, ensuring they address individual informational needs and preferences effectively.

(JMIR Cancer 2026;12:e73156) doi:10.2196/73156

XSL•FO
RenderX

KEYWORDS

information tool; real world outcome data; CDK4/6 inhibitor; advanced breast cancer; value-based health care

Introduction

Given recent health care challenges, accelerating the transition towards value-based health care (VBHC) has become urgent [1]. VBHC aims to maximize value for patients by optimizing patient-relevant care outcomes, while preferably lowering or maintaining the same costs. Outcome data play a major role in doing so. On a group level, outcome data are increasingly used for quality improvement through benchmarking, that is, comparisons over time, between treatments, and across institutions. On an individual patient level, outcome data should be considered an essential component of patient information and communication. However, such use of outcome data is still not standard in everyday clinical practice [2-5]. In particular, real-world outcome data, collected in daily clinical practice, can be used to support shared decision-making (SDM). These data help patients weigh the pros and cons of treatment options available to them, in line with their own values and expectations. This likely increases patients' satisfaction and treatment compliance [5,6].

SDM is widely recognized for supporting preference-sensitive decisions, particularly when multiple treatment options offer similar benefits or when opting out of treatment is viable, but outcomes remain uncertain [7]. In palliative care, patients have complex needs [8]; however, their desired involvement in decision-making is often not achieved, and their information needs, including their need for prognostic information, are not met [9-11]. They are often not informed about expected survival benefit or the option to refrain from treatment, while the survival benefit is often small and uncertain. Simultaneously, the burden of treatment is likely high in terms of side effects and frequent check-ups [12]. Furthermore, drug performance in clinical trials often differs significantly from results observed in real-world settings [2,13-15], which likely complicates decision-making and raises questions about the applicability of trial findings to broader real-world populations whose characteristics may differ from those of trial participants.

We aimed to deliver a proof-of-concept of including real-world outcome data in patient information and communication. Therefore, we conducted a case study in women with hormone receptor–positive (HR_{pos}), human epidermal growth factor receptor 2–negative ($HER2_{neg}$) metastatic breast cancer (BC). This type accounts for 65% of all BC cases worldwide [16], of which 3% - 12% are projected to have metastatic disease [17] with a 5-year survival rate of approximately 37% [16]. From this target group, we included women eligible to receive the CDK4/6 inhibitor palbociclib combined with endocrine therapy. This targeted therapy aims for disease control by slowing tumor growth, prolonging patients' progression-free survival (PFS). In clinical practice, CDK4/6 inhibitors are most often offered in second- or later-line, following progression on previous endocrine therapy. By targeting cell cycle regulation, CDK4/6 inhibitors reverse or delay resistance to endocrine treatment, and delay the need for more burdensome chemotherapy. Clinical trials have reported promising results (eg, improved time until disease progression with 4 - 10 months [18,19]); however, there are also data that outcomes differ in real-world settings. For example, how modifications are handled in case of neutropenia [20,21]. Based on these considerations, we prioritized the use of real-world data in this research.

In this study, we aimed to (1) systematically develop a patient information tool incorporating real-world outcome data to support patients in making informed decisions about treatment initiation and managing their care throughout the treatment process, and (2) evaluate the tool in a pilot study with the target population. We used a participatory development approach [22] to center this study on patients and health care professionals (HCPs), deepening our understanding of their needs and care experiences and enhancing the usefulness of our patient information tool for supporting SDM. In line with the VBHC emphasis on collaboration across patients, clinicians, organizations, and other stakeholders, participatory design brings these groups together in co-design. This helps surface challenges earlier and leads to more effective and user-friendly solutions.

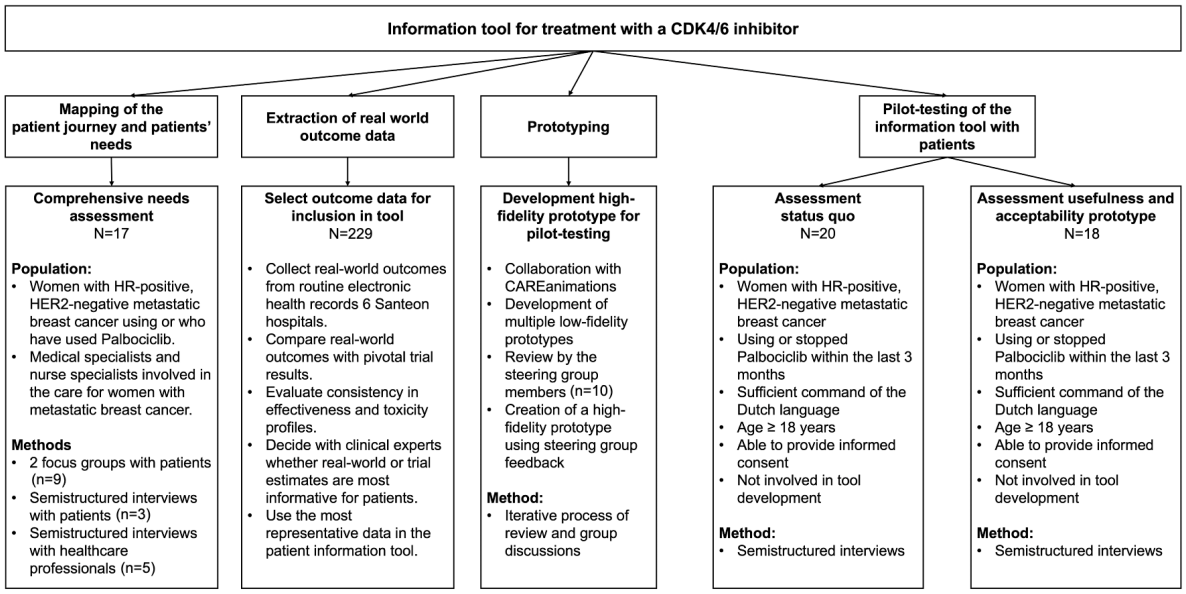
Methods

Overview

We used qualitative methods to develop the CDK4/6 inhibitor real-world patient information tool. The development process consisted of five steps: (1) establishment of a multidisciplinary steering group, (2) mapping of the patient journey and patients' needs through focus groups and semistructured interviews, (3) extraction of real-world outcome data from electronic health records systems, (4) prototyping of the tool, and (5) pilot evaluation with the targeted patient population using semistructured interviews (Figure 1).

XSL•FO
RenderX

Figure 1. Flow chart development process components. HER2: human epidermal growth factor receptor 2; HR: hormone receptor.



Ethical Considerations

The study protocol was approved by the Medical Research Ethics Committees United (Nieuwegein, the Netherlands; reference SDB2021-010). Ethics and research governance approval were obtained from the local medical ethics committees of the participating hospitals. Participating patients provided written informed consent.

Establishment of a Multidisciplinary Steering Group

The development of the information tool was initiated by Santeon, a Dutch hospital group in which 7 top clinical hospitals openly collaborate to improve medical care [21]. The project initiators (postdoctoral researcher MQNH, program manager CFvUK, and hospital pharmacist EMWvdG) established a multidisciplinary steering group for the development of the tool.

Mapping of the Patient Journey and Patients' Needs

The patient journey mapping (ie, inventory of patients' experiences from diagnosis until disease progression) served as the starting point in the development of the information tool. The goal was to obtain insight into patients' experiences from diverse patients to help identify information and support needs that needed to be addressed in the patient information tool. We collaborated with the Patient Journey Lab, a company specialized in patient journey mapping, to develop the patient journey. First, to roughly outline the patient journey, 2 focus groups were conducted by MQNH and an independent research consultant (experienced in conducting focus groups) with 5 and 4 patients, respectively, that lasted about an hour. They focused on the following topics: (1) information provision, (2) decision-making to start the CDK4/6 inhibitor, (3) unmet information needs, and (4) perspectives on the use of real-world outcome data. Second, to verify and fill out the patient journey in greater detail, semistructured interviews, with an expected duration of 30 minutes, were conducted by MQNH with 3 patients and the 2 oncologists and 3 nurse specialists who participated in the steering group. All recordings were transcribed and summarized by MQNH. Manual summative content analysis [23] was used to infer meaning from recurring keywords and messages. Based on the topic list, keywords were selected to highlight in the text to count their frequencies before defining and contextualizing them in our final summary.

Extraction of Real-World Outcome Data

To inform new patients about the treatment, the information tool includes outcome data of 229 patients treated between 2016 and 2020 in 6 Santeon hospitals (St Antonius Hospital Utrecht/Nieuwegein, Martini Hospital Groningen, Canisius-Wilhelmina Hospital Nijmegen, Catharina Hospital Eindhoven, OLVG Amsterdam, and Maasstad Hospital Rotterdam). More detailed information on the real-world cohort and the data collection has been reported previously and is available elsewhere [24,25]. Briefly, the data collected included PFS, the reason for ending treatment, the severity, frequency, and timing of all side effects and toxicities (with a special focus on neutropenia), and the timing and frequency of all treatment modification strategies. They were derived from routinely collected electronic health records, with 1 hospital excluded due to incomplete documentation of palbociclib use. To ensure the robustness of the real-world data to be included in the tool, we compared in this dataset outcomes in a cohort of 229 patients with HR_{pos}, $HER2_{neg}$ metastatic BC treated with palbociclib plus fulvestrant in second- or later-line therapy to those reported in the pivotal *PALOMA-3* trial that formed the basis for regulatory approval of palbociclib in this setting. On the basis of the comparison, the expert work group concluded that given the greater clinical heterogeneity and the mismatch in toxicity profiles compared with trial participants, real-world outcome estimates were more representative for informing patients and were therefore used in the information tool. [24,25]

Prototyping

In consultation with the steering group and based on the patient journey mapping, it was decided to develop a digital information tool to allow future efforts to link the tool to patients' electronic health records. Such a linkage provides two main benefits: (1) it gives patients access to all medical information in one place,

XSL•FO
RenderX

and (2) it allows the tool to display up-to-date real-world outcome data based on automatic updates retrieved from personal data stored in electronic health record systems. To examine how the tool could best be integrated in the digital landscape of the Santeon hospitals, the Patient Journey Lab inventoried the software potential per hospital, including available health applications. As CAREanimations offers accessible information suitable for most people, including those with low (health) literacy, tailored to each step in the patient journey, it was selected as the development and implementation partner [26]. In collaboration with CAREanimations, the initiators developed several low-fidelity prototypes, which were reviewed by the steering group to arrive at a high-fidelity prototype for pilot-testing.

Pilot-Testing of the Information Tool Among the Targeted Population

We carried out a pilot evaluation in 2 steps. First, before implementing the information tool, we assessed the status quo, that is, patients' experiences with treatment and their met and unmet information and support needs. We did this to validate the findings of our previous unmet need assessment. Second, we assessed patients' views on the usefulness and acceptability of the high-fidelity prototype of the information tool and inventoried suggestions for improvement.

In the first step, we conducted in-depth semistructured interviews with women with $HR_{pos}HER2_{neg}$ metastatic BC who (1) were using or had stopped using the CDK4/6 inhibitor Palbociclib within the last 3 months, (2) had sufficient command of the Dutch language, (3) were 18 years or older, (4) did not have (severe) cognitive impairment that would preclude them from providing informed consent, and (5) had not been involved in the development process. We obtained insights into patients' (1) experiences with information provision about their diagnosis, prognosis, and treatment with CDK4/6 inhibitor; (2) experiences with treatment (eg, side effects and impact on daily functioning); (3) experiences with treatment decision-making; and (4) preferences regarding receipt of outcome data. In addition, we obtained information about sociodemographic characteristics (age and educational level) and medical history (date of primary BC diagnosis, treatment history; refer to Multimedia Appendix 1 for the full interview guide). We recruited participants from 5 Santeon hospitals (ie, St Antonius Hospital [Utrecht/Nieuwegein], Martini Hospital [Groningen], Canisius-Wilhelmina Hospital [Nijmegen], Catharina Hospital [Eindhoven], and Maasstad Hospital [Rotterdam]). Patients were invited to participate by their oncologist or nurse specialist. The interviews were conducted by EGE (researcher with interviewing experience).

In the second step, we evaluated the prototype of the information tool through individual semistructured interviews with a new cohort of women not previously involved in the development or pilot-testing, meeting the same inclusion criteria as described in step 1. Participants were recruited at the same 5 Santeon hospitals and via the Dutch Breast Cancer Society. Recruitment through the Dutch Breast Cancer Society broadened the target population, as participants were also treated at other hospitals, which accelerated the recruitment process. Two weeks before the interview, participants were provided with access to all the animation videos contained in the information tool and the graphical information intended for use during the consultation with their oncologist. They were asked to view all the materials and take notes of any suggestions for possible improvements. After they reviewed the information, a telephone interview was conducted by MQNH or EGE. The interview guide to assess patients' experiences with treatment and their met and unmet information and support needs was used and extended to also obtain insight into patients' views on the completeness and clarity of the CDK4/6 inhibitor information tool and to collect suggestions for improvement (refer to Multimedia Appendix 1 for the full interview guide).

All interviews (in the first and second steps) were audio-recorded with the participant's permission and transcribed. The transcripts were analyzed using a combined technique of deductive and inductive thematic analysis [27,28]. For the deductive approach, we constructed an analytic framework based on predefined themes in the interview guide. In addition, we conducted open coding (inductive) on the transcripts to allow the inclusion of other relevant themes. We generated categories and abstractions through detailed review, quantifying keywords and phrases, and comparing findings. As educational level was considered to be an important predictor of patients' comprehension of and satisfaction with the information tool, we explored differences in participants' responses based on their educational background. All transcripts were coded by 2 researchers (MQNH and EGE) independently. Discrepancies were resolved through consensus. Qualitative data analyses were performed using Atlas.ti (version 9; Lumivero) software. Quotes are provided to illustrate our findings.

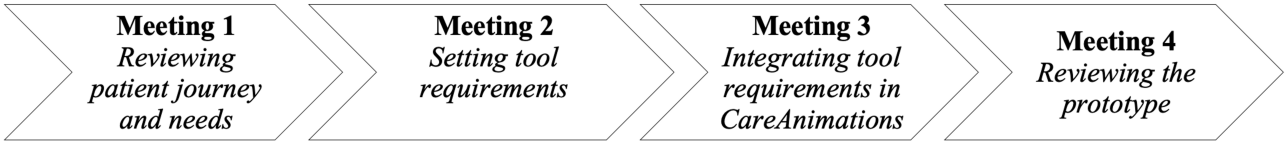
Results

Establishment of a Multidisciplinary Steering Group

The steering group consisted of patient representatives or advocates from the Dutch Breast Cancer Society (n=3), oncologists (n=2), specialist nurses (n=3), an information architect (n=1), and a health information consultant (n=1; for more details, refer to Multimedia Appendix 2). The steering group convened for 4 cocreation sessions. Figure 2 provides an overview of the focus of each session. The steering group was involved throughout the development process.

XSL•FO
RenderX

Figure 2. Focus of the cocreative steering group sessions.



Mapping the Patient Journey and Patients' Needs

The findings from interviews and focus groups with patients (n=12) and clinicians (n=5) were used to map the patient journey of women undergoing treatment with a CDK4/6 inhibitor, from the start of treatment to its discontinuation (Table 1). A detailed overview of the patient journey is provided in Figure 3. Findings from the interviews with clinicians about the requirements for the information tool were in line with the findings from the patient interviews. All participants valued real-world outcome data in addition to clinical trial results, preferably displayed in a leaflet or app (connected to the electronic medical record). Information needed to be tailored to each step of the patient journey. Repetition and gradually offering more information were believed to be key. Even though dealing with a wide range and likely negative outcomes may be complex, participants wished insight into the timing, frequency, and severity of side effects, as well as the timing and reasons for ending treatment, so they could compare their outcomes to those of peers. Most patients wished to be better informed about possible medication alterations. Information on what type of treatment CDK4/6 inhibitor is, as well as what to do when experiencing side effects, was also valued. A summary of the key findings from the focus groups and interviews conducted to map the patient journey can be found in Table 2.

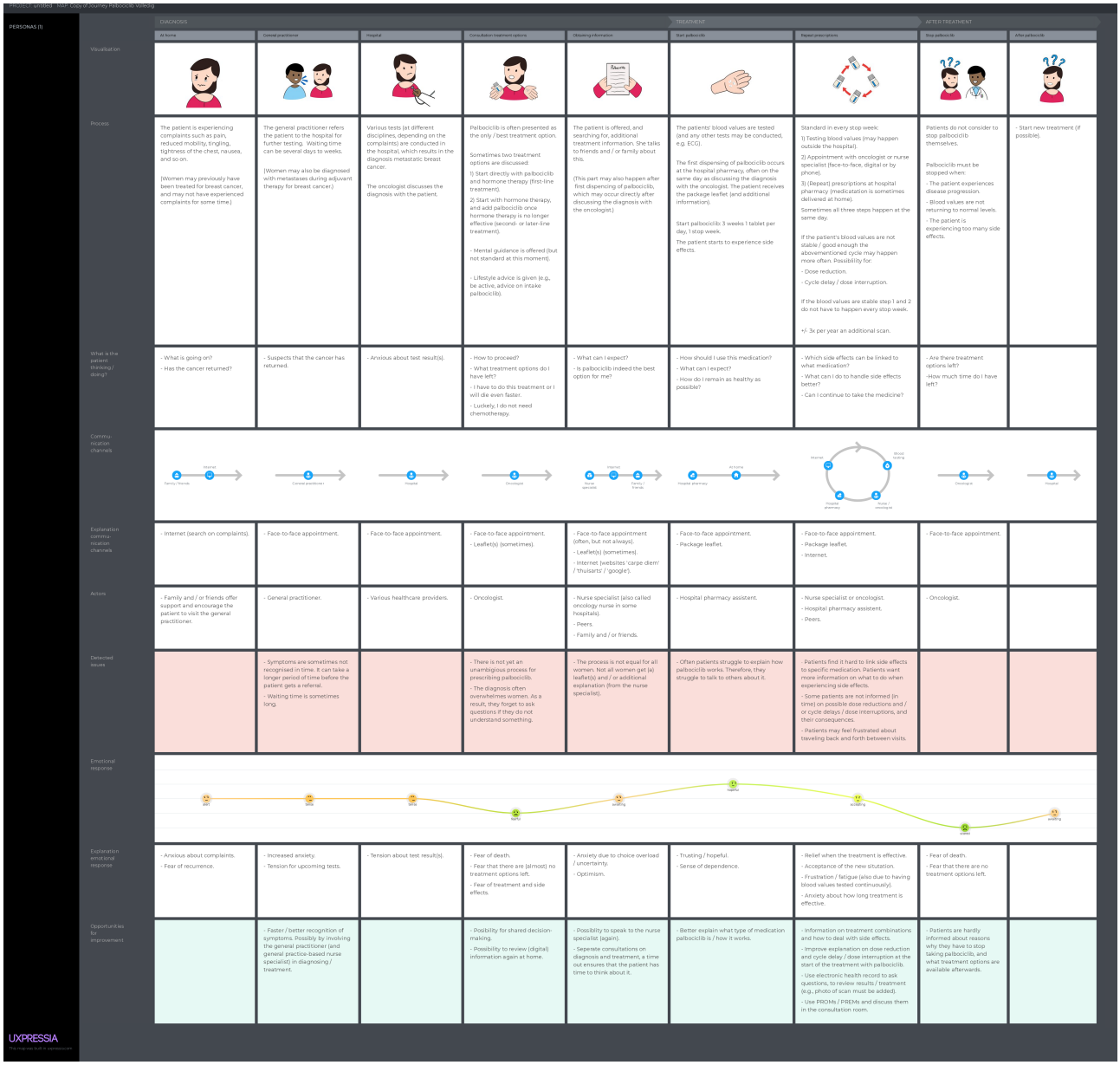
Table . Descriptive statistics of patients participating in focus groups and semistructured interviews (n=12).

Characteristics[a]	Patients
Age (y), range	38 - 66
Missing, n	3
Education, n	
Low	1
Middle	4
High	5
Missing	2
Country of residence, n	
The Netherlands	10
Belgium	1
France	1
Breast cancer, n	
Metastatic recurrence	5
Immediate metastasis	6
Missing	1
Therapy, n	
First-line: palbociclib and letrozole or anastrozole	5
Second- or later-line: palbociclib and fulvestrant	6
Missing	1
Duration usage palbociclib, n	
Half a year or longer	12

[a]In a small number of interviews, certain sociodemographic questions were not asked or probed further because the conversation centered on highly emotional and sensitive topics, and the interviewer prioritized participant comfort and rapport.

XSL•FO
RenderX

Figure 3. Detailed overview of the patient journey.



XSL•FO
RenderX

Table . Summary of findings per theme and subtheme of the focus groups and interviews conducted to map the patient journey.

Theme and subtheme	Description
Information resources	
Health care professionals	Patients were informed about CDK4/6 inhibitor by their medical oncologist. After the initial consultation with their medical oncologist, some, but not all, patients received more detailed information from a nurse specialist or oncology nurse. Patients who did not have a dedicated nurse specialist or oncology nurse experienced a lack of guidance or support.
Websites and information leaflets	Some patients were referred to websites or information leaflets or were given an information leaflet at the hospital pharmacy. Patients liked receiving these additional information resources as the information received from health care providers was a lot to take in all at once, and these additional resources could serve as a reference for later.
Medication leaflet	All patients received the medication leaflet, which they handled differently. Some patients read it partly or in its entirety (multiple times) before starting with the treatment. Others consulted the medication leaflet upon experiencing complaints to avoid forgetting information, getting (unnecessarily) worried, or imagining symptoms that were not there. Some patients consulted peers on the internet.
Experiences with information provision about treatment	
Knowledge gap	To many patients it was not clear what type of treatment CDK4/6 inhibitor is: it was often mistaken for chemo- or immunotherapy. Patients used the internet to search for more detailed information on side effects and how to deal with them.
Variation in information provision and unmet information needs	Participants had mixed experiences with information provision about medication alterations (eg, lowering of the dosage or extension of the time interval between treatments). Some participants were informed about possible medication alterations before treatment started, some when it became relevant during treatment, whilst others were not consulted when an alteration was needed. Most patients wished to be more informed about their options.
Outcome data	
Preferences for receiving outcome data	Patients indicated they would value receiving real-world outcome data in addition to clinical trial results. Real-world outcome data, tailored to their medication combination, could help them compare their outcomes to those of peers, even though dealing with a wide range, and likely negative, outcomes may be complex.
Role in decision-making	
Oncologists	Upon diagnosis, the oncologist often immediately decided to start treatment with CDK4/6 inhibitor, as they believed it to be the best treatment option.
Patients	Some patients liked the oncologist to take the lead as they felt there was a need to act swiftly after their diagnosis and trusted their oncologist. Some patients were more involved in decision making, by proposing treatment with CDK4/6 inhibitor themselves, or by discussing alternatives (eg, chemotherapy) with their oncologist. Overall, conservative treatment was not considered an option by patients.
Requirements for the information tool	
Important content	A description of the type of treatment and what to do when experiencing side effects.
Preferences for incorporating outcome data in the information tool	Also, real-world outcome data related to the timing, frequency, and severity of side effects and medication alterations was deemed important to disclose in the tool, as well as the timing and reasons for ending treatment.
Type of tool and specifications	Patients liked to receive real-life outcome data in a leaflet, handed out by the nurse specialist or oncology nurse, but an app (connected to the electronic medical record) was also mentioned. In terms of information presentation, it was deemed important that the information in the tool was tailored to each step of the patient journey, and that repetition and gradually offering more detailed information is key.

XSL•FO
RenderX

Extraction and Assessment of Real-World Outcome Data

Real-world outcome data were extracted from a cohort of patients treated with palbociclib in combination with fulvestrant in routine clinical practice. In this cohort, the median PFS was 11.6 (95% CI 10.2-13.9) months, which was longer than the 9.5 (95% CI 9.2-11.0) months reported for the PALOMA-3 trial population (more detailed information is available in the study by Hackert et al [24]). This difference may partly reflect differences in outcome assessment between clinical trials and routine clinical practice. When the real-world cohort was stratified according to PALOMA-3 eligibility criteria, differences in outcomes were observed. Patients who would have met the PALOMA-3 eligibility criteria had a median PFS of 14.1 (95% CI 11.5-17.0) months, whereas patients who would not have met these criteria had a median PFS of 10.2 (95% CI 8.8-12.3) months.

The occurrence of hematological toxicities and nonhematological side effects associated with palbociclib was generally similar in real-world settings compared with the clinical trial data. However, patients in routine clinical practice reported certain side effects more frequently, particularly fatigue and decreased appetite. In clinical practice, 54% (124/229) of patients experienced at least 1 cycle delay, 39% (89/229) had at least 1 dose reduction, and 26% (59/229) had at least 1 dose interruption, compared with 36% (124/345), 34% (117/345), and 54% (186/345), respectively, in the PALOMA-3 trial. Detailed findings on patient-reported experiences are described in the study by Vaisson et al [22].

Choices in Data Presentation in the Information Tool

To support the development of the information tool, the extracted real-world outcomes were summarized using lay terminology and reviewed together with the steering group. Based on the extracted outcome data, decisions were made in collaboration with the steering group regarding which outcomes to include in the tool and how these should be presented to patients. Particular attention was given to ensuring that the translation of clinical outcomes into patient-friendly information accurately reflected the limitations of real-world data, such as potential selection bias, missing or incomplete data, and variability in outcome measurement and reporting. For example, rather than presenting outcome distributions based solely on real-world data, outcomes from clinical trial data and real-world data were compared, and presentation choices were made based on this comparison. This approach aimed to provide balanced and transparent information for patients using the information tool.

Prototyping

Based on the requirements set by the steering group (using the results from the patient journey mapping and the patient and clinician needs assessment), prototypes for the information tool were developed. Figures 4 and 5 provide a summary of the patient journey, including components of the information tool, to support information provision and treatment decision-making at each phase of the patient journey. The information tool consists of three main components: (1) KIJKgesprek (Stichting Kijksluiter), (2) KIJKsluiter (Stichting Kijksluiter), and (3) KIJKbericht. The 3 components are described in Figure 6. The steering group reviewed the prototypes, and in an iterative process, the prototypes were refined to a version ready for pilot-testing—the high-fidelity prototype.

XSL•FO
RenderX

Figure 4. Examples of the presentation of the probabilistic and biomedical information in the tool. (A) Positive and negative framing of probability of experiencing neutropenia. (B) Supporting biomedical information about neutropenia before presentation of outcome data.

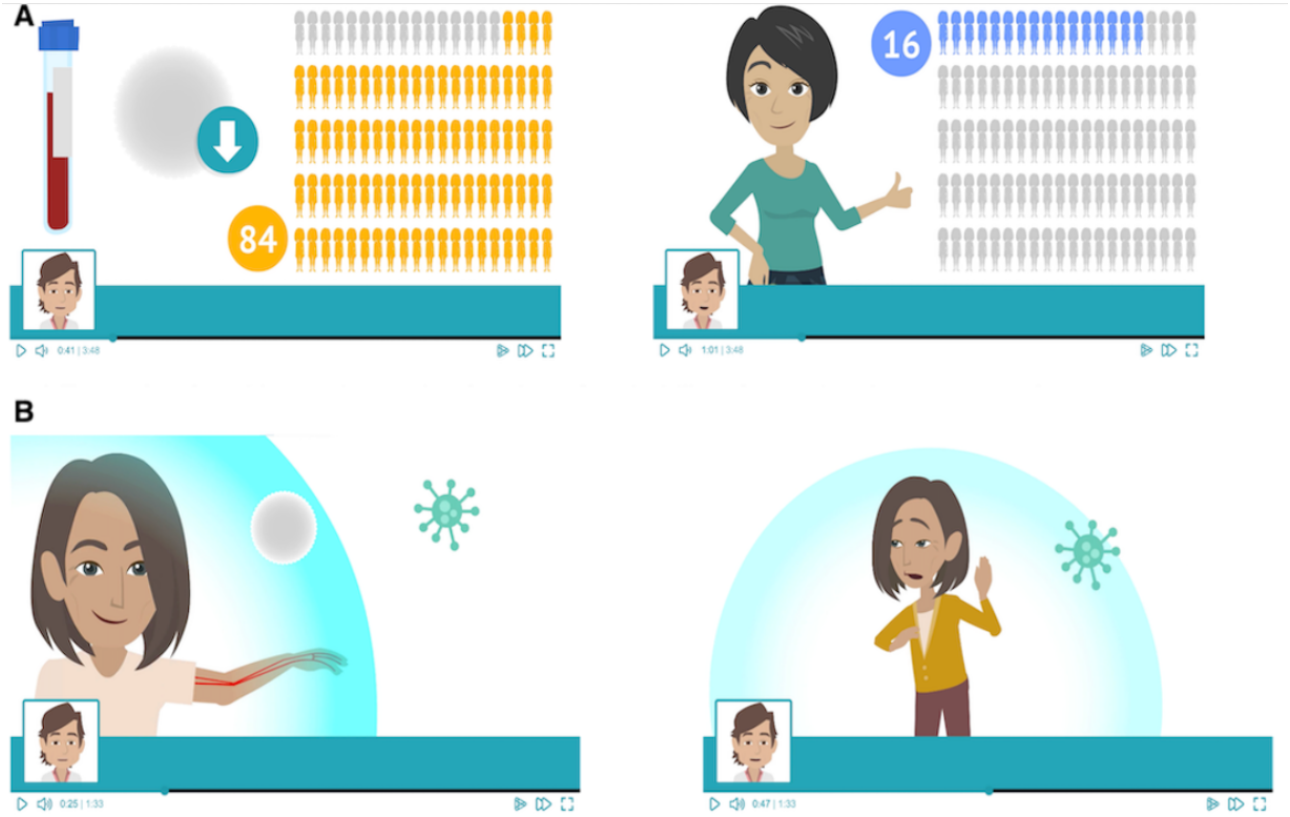
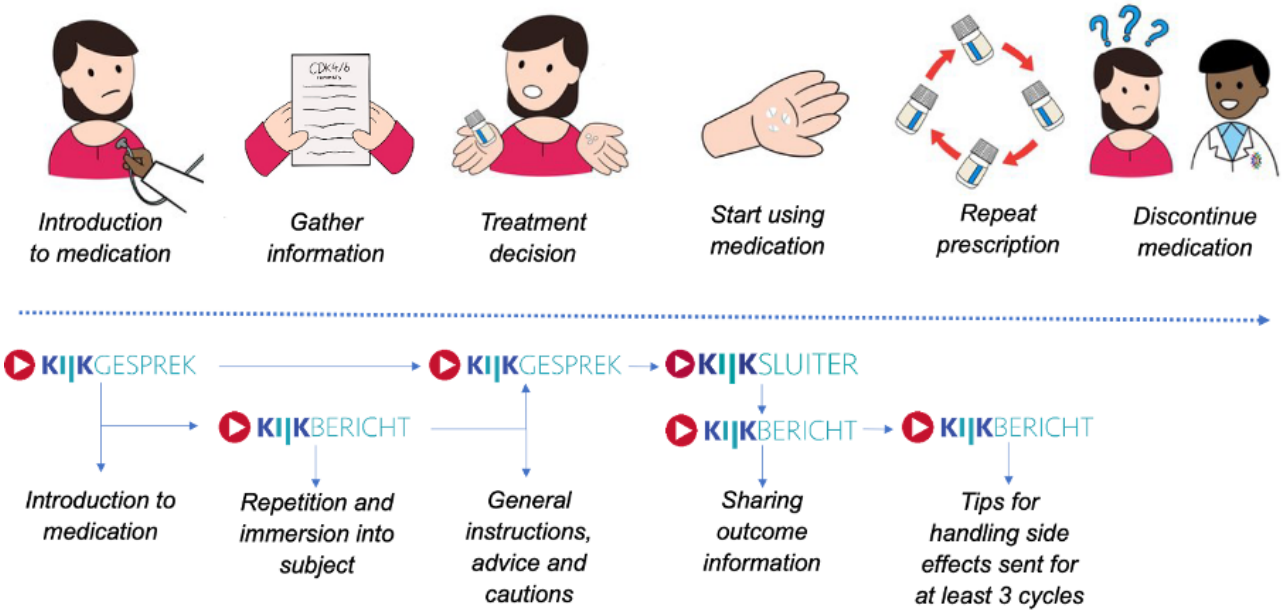
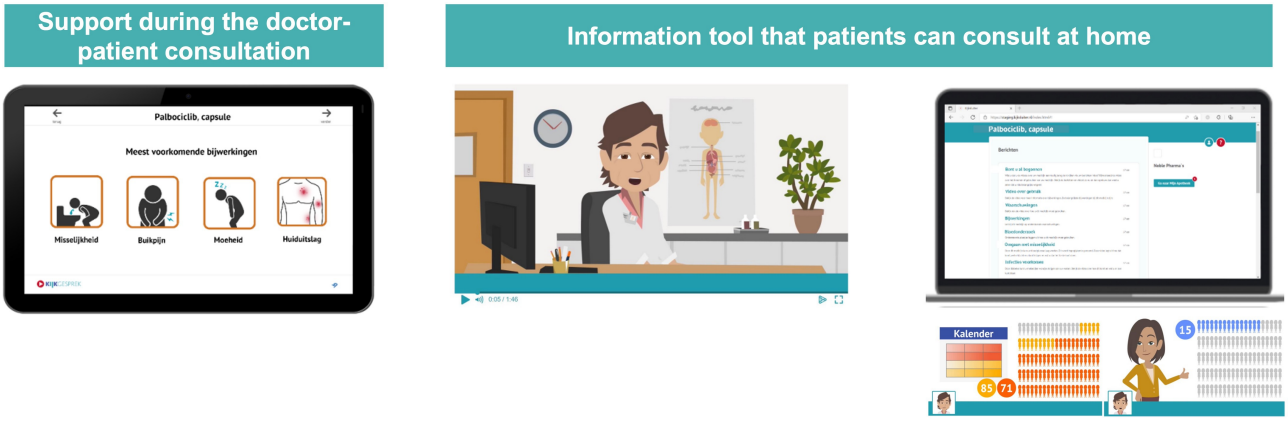


Figure 5. Summary of patient journey including the solutions developed to support information provision at each phase of the patient journey.



XSL•FO
RenderX

Figure 6. Components of the CDK4/6 inhibitor information tool.



Pilot-Testing of the Information Tool Among the Targeted Population

Step 1: Status Quo Prior to Implementing the Information Tool

Overview

To obtain insight into patients' experiences with information provision, treatment decision-making, and the treatment burden before the CDK4/6 inhibitor information tool became available, we conducted 20 interviews. Table 3 provides an overview of the participants' characteristics. We identified 5 key themes and summarized the findings per theme.

XSL•FO
RenderX

Table . Sociodemographic characteristics of women interviewed during pilot-testing of the information tool.

Characteristics[a]	Step 1: before implementation (n=20), n (%)	Step 2: prototype testing (n=18), n (%)
Marital status		
Married or in a relationship	9 (47)	12 (80)
Single	4 (21)	2 (13)
Widowed	6 (32)	1 (7)
Missing	1	5
Education		
Lower	7 (41)	3 (20)
Intermediate	4 (24)	5 (33)
Higher	6 (35)	7 (47)
Missing	3	5
Occupation		
Paid job	2 (11)	2 (14)
Volunteering	1 (5)	1 (7)
Retired	7 (37)	3 (21)
Unemployed	9 (47)	8 (57)
Missing	1	6
Diagnosis		
Prior breast cancer	13 (93)	10 (67)
Immediate metastasis	1 (7)	5 (33)
Missing	6	5
CDK4/6 inhibitor as:		
First-line	9 (56)	8 (57)
Second- or later-line	7 (44)	6 (43)
Missing	4	6

[a]In a small number of interviews, certain sociodemographic questions were not asked or probed further because the conversation centered on highly emotional and sensitive topics, and the interviewer prioritized participant comfort and rapport.

Patients' Treatment Knowledge

Most patients could not describe what kind of treatment a CDK4/6 inhibitor is. Many patients mistook the CDK4/6 inhibitor for chemotherapy:

> [I understood that the drug is a kind of chemotherapy that fights cancer cells. ID03]

Patients did seem to understand that a CDK 4‐6 inhibitor helps to slow or stop the growth of cancer cells. Side effects that most patients knew of before the start of treatment were alopecia, gastrointestinal complaints, fatigue, and hematological changes (eg, neutropenia). Many patients linked blood check-ups to assess how well they were tolerating the treatment and whether they could start a new treatment cycle. There seemed to be variation in information provision; some patients were informed about treatment modification strategies (such as stopping for an additional week before starting the next treatment cycle or lowering the dosage), while others were not.

Information Resources Use

Study participants discussed various information resources they consulted, including HCPs, written materials, medication leaflets, and the internet. Patients consulted medication leaflets before or during treatment, finding them helpful but overwhelming due to the extensive list of potential side effects. Many patients preferred information about side effects in smaller doses, and timing of information was also deemed important (eg, getting information when it became relevant for them instead of hearing everything in advance). Some patients sought additional information online, while others avoided it due to concerns about accuracy and biased narratives (eg, overrepresentation of "horror stories"). Although most participants were satisfied with the information they received from HCPs, some felt it was insufficient, particularly regarding the management of side effects. Also, participants felt that the amount of time available to consult with their oncologist was too limited. Overall, participants appreciated that they could contact HCPs for further clarification outside of their scheduled consultations if the need should arise.

XSL•FO
RenderX

Patients' Preferences Regarding Outcome Data

Many participants were receptive to receiving information about potential side effects, with preferences varying in terms of detail. Some preferred information limited to the most common side effects based on real-world data, without probabilities, while others wanted a comprehensive overview including probabilities. Many trusted their oncologist to provide relevant information as needed, expressing no strong desire to receive all information upfront or to use a tool containing such data. For some, not knowing everything served as a coping mechanism to maintain hope and reduce anxiety about factors beyond their control (eg, response to treatment). Despite this, a few participants acknowledged an internal struggle, balancing the desire to understand what could lie ahead with the comfort of remaining uninformed. Additionally, many patients expressed reluctance to receive outcome data related to prognosis, feeling such statistics might not apply to them (eg, "I do not like statistics or prognoses at all, because I always fall outside their scope" [ID01] or "You have to wait and see how things are going to turn out in your situation" [ID02]) and could be too confronting or overly influential on their life choices (eg, "I do not want to think I only have a year left" [ID03]). Some preferred outcome data to be available only upon request or in urgent situations, such as disease progression requiring decision-making. This ambivalence highlights the complex and individualized nature of patients' preferences for receiving prognostic information.

Experienced Treatment Burden

All patients experienced treatment side effects, with about a quarter encountering cycle delay, dose reductions, or interruptions due to, for example, severe side effects. Participants used several coping strategies. Many spoke of acceptance being important to cope with their situation, and acknowledging the futility of questioning what could have been:

> [You can start wondering, could I have prevented this metastatic disease? Could it have been different? But I know from experience that those kinds of questions don't get answered anyway, and that's no use to you. So, you have to make do with what you have at that moment. ID04]

Some found comfort in hoping to defy the odds:

> [They clinicians did say right away, we cannot cure you, but we can make the quality of your life better and prolong your life. But they also said that bit prolonging life. So, I've had the hope that maybe I would be cured after all... ID05]

Denial and trivialization of their symptoms and situation were also common coping mechanisms. Despite many patients rating their quality of life high, many simultaneously acknowledged significant negative impacts on their daily lives due to the disease and/or treatment side effects, and ways they needed to adapt.

Treatment Decision-Making

Patients expressed that they felt a sense of urgency to start treatment with a CDK4/6 inhibitor, as it is perceived to be the best available option at this stage of their disease:

> [I had already taken letrozole for four years and it wasn't working anymore … so you have to switch to something else. For me it feels like I don't really have a choice … they know which option is best for me. ID02]

A few preferred a CDK4/6 inhibitor due to its perceived lower burden compared with other treatments, such as chemotherapy. Many patients indicated their only other option would be chemotherapy and expressed reluctance to start this, given its potential side effects and impact on their quality of life:

> [For me the choice was chemotherapy and its side effects, or this treatment and its side effects. So, I didn't really focus on what the side effects of this treatment were, because compared with chemotherapy it's obviously much less severe. ID18]

Given that chemotherapy was the alternative, being eligible for treatment with a CDK4/6 inhibitor was seen as positive news. While some patients felt involved in the decision-making process, most patients indicated that their oncologist took the lead in prescribing treatment, often without discussing alternatives. Trust in oncologists played a significant role in patients' acceptance of treatment decisions. However, some oncologists did offer patients the opportunity to decide or take time to consider their options. Patients commonly cited the desire to extend life, slow cancer cell growth, reduce symptoms, and improve quality of life as motivation for starting CDK4/6 inhibitor treatment.

Step 2: Prototype Testing

Overview

A total of 18 patients pilot-tested the information tool (Table 3). Their experiences with treatment, information provision, knowledge about treatment, and coping strategies aligned with the findings from the women who did not have access to the tool (step 1). Therefore, we only summarize the prototype testing participants' views and recommendations for each component of the information tool below.

Views on the Visualized Key Information for Use During the Clinical Encounter (KIJKgesprek)

Participants generally thought that the so-called KIJKgesprek used during a consultation (a visualization using animated pictograms to help patients understand and remember essential medical information from the consultation, eg, medication intake, common side effects, and contraindications) was well-structured and comprehensive in covering key points. However, opinions varied regarding its added value. While some women believed it could enhance doctor-patient communication by aiding information processing, others doubted its added benefit or expressed concerns about it being potentially distracting to patients (eg, it can be too much to have to look at a screen as well as process what the HCP is saying):

> [I didn't feel it added much value for me, although it might be useful for the oncologist to support the explanation …. Personally, I didn't really think it would have added anything to the conversation with my oncologist. Everything was already clear to me.

XSL•FO
RenderX

> What did stand out to me, though, was that I hadn't really received the warnings mentioned, such as avoiding St John's wort, grapefruit, alcohol, and being aware of possible liver problems. That might have been a useful addition. ID12]

One participant acknowledged the KIJKgesprek's potential utility as a guide for HCPs to ensure comprehensive information delivery.

Views on the Animated Information Videos

Most participants found the videos informative, and the language use was deemed clear and accessible:

> [I found it information tool very helpful. I think it can really be a valuable addition when you start with it treatment … although also midway through treatment. I found it clear. ID4]

Participants indicated that it would have been of added value to them if they had had access to this tool at the time they started with treatment. While participants believed all relevant topics were covered in the videos, opinions varied about the level of detail. More highly educated participants desired more information and a greater level of detail. The tips for coping with side effects provided in the videos were highly valued by all. Some provided suggestions for more tips that could be added. The videos' length was deemed appropriate, and participants appreciated the ability to revisit specific topics. Participants saw the tool as a valuable reference, especially for processing information after receiving difficult news. However, a few preferred to avoid excessive information unless necessary, citing concerns about losing hope.

Views on the Incorporation of Outcome Data in the Information Tool

Visualization of real-world outcome data in the information tool introduced a novel aspect, but opinions among participants varied. Overall, participants thought that the visualization of the outcome estimates was clear. Some expressed an internal conflict, acknowledging the importance of knowing realistic probabilities of experiencing outcomes while also feeling overwhelmed or wanting to avoid such information:

> [It outcome data is reality, of course. It's just that sometimes it depends, it's good to know, but at times you also don't want to know. Right? ID07]

Many participants believed it was useful to provide patients with realistic outcome data, particularly regarding PFS and side effects. Some participants indicated that they had explicitly requested information about prognosis from their oncologists, and not all had received numerical information. However, many expressed concerns that providing patients with outcome data could diminish hope or discourage treatment initiation (in line with the findings described in the Step 1: Status Quo Prior to Implementing the Information Tool section). Some worried particularly about visualizing the average time the medication is effective in preventing disease progression without adequately acknowledging the variability in treatment response:

> [If I had seen before starting with treatment that on average women only had four months before the disease got worse, I might not have started. For me, that hasn't been the case. I'm lucky to be able to do well on this CDK4/6 inhibitor for years now. … It would be good to explicitly state that the four months is an average, but there are exceptions. ID08]

Opinions also diverged on whether to provide outcome data about side effects, with some finding it overwhelming and others valuing the insights and validation of their own experiences:

> [Seeing that others other women using a CDK4/6 inhibitor also experience the same complaints is reassuring, it's not just me. ID09]

Linking Baseline Patient Experiences to Prototype Evaluation

Overall, step 1 showed substantial gaps in patients' understanding of CDK4/6 inhibitor therapy, wide variation in information needs and preferences, and challenges in coping with treatment burden and decision-making. These findings underscore the need for structured, accessible, and patient-tailored support materials that incorporate realistic expectations of treatment and side effects. Step 2 demonstrated that the prototype information tool addressed many of these needs. Even though preferences for receiving outcome data varied between patients, many participants did find this information useful. Together, these insights support the incorporation of real-world outcome estimates into the tool.

Discussion

We present the development and pilot testing of a novel online patient information tool for women with $HR_{pos}HER2_{neg}$ metastatic BC undergoing treatment with a CDK4/6 inhibitor. Designed based on patient and HCP needs assessments and input from a multidisciplinary steering group of experts, the tool delivers numerical data on outcomes such as PFS, side effects, and treatment modifications in an accessible format. It also features visualizations to support doctor-patient communication during consultations and includes an app with animation videos and push notifications for home use. By facilitating discussions about treatment outcomes, the tool can help patients form realistic expectations about the benefits and risks of their treatment. This, in turn, is aimed at supporting better-informed SDM and fostering more meaningful patient engagement in their care. Despite its benefits in informing patients and supporting treatment decisions, the use of real-world outcome data in clinical practice is still novel and may be challenging for patients to understand and apply. As its integration becomes more common, it is important to keep in mind that patients may need time to adapt.

Many participants in our pilot test demonstrated significant knowledge gaps regarding CDK4/6 inhibitors. Although participants generally understood that these compounds slow cancer cell growth, awareness of treatment modification strategies and potential side effects was limited. These findings highlight the critical role of patient information tools, like the tool we developed, in promoting consistent and clear communication. Many patients reported that symptom burden significantly impacts their quality of life, underscoring the need

XSL•FO
RenderX

for resources with information about strategies to address this challenge [29-31].

Furthermore, challenges in communicating about palliative care often stem from HCPs' desire to offer hope, which can result in incomplete discussions about treatment options and expected outcomes [32]. The visualizations for use during the consultation incorporated in our tool are designed to support clinicians in delivering complex information more effectively to patients. This approach could be particularly helpful in ensuring patients are better informed, fostering engagement, and setting realistic expectations about their treatment journey. Tailored explanations of treatment mechanisms and side effects, whether detailed or concise, based on individual preferences, can help address misconceptions, reduce concerns, and provide a balanced sense of hope that aligns with patients' needs and preferences.

Preferences for outcome data among participants in the pilot test were highly individualized. While some patients valued receiving numerical outcome data about side effects and PFS, others preferred to avoid such information to preserve hope or reduce anxiety. Many participants in our pilot test expressed a preference for outcome data to be offered only on request or in specific scenarios, such as if there is disease progression, emphasizing the importance of tailoring and sensitivity in presenting this type of data. These findings align with research showing that patients often prefer to be asked whether they want to receive outcome data, rather than having it provided automatically, and that these preferences may change over time, necessitating periodic reassessment [11]. In this evaluation study, all participants were exposed to the full set of outcome information. However, this does not reflect the intended use of the tool in practice, where information would be shown only at clinically relevant time points. Our findings also support adding advanced notifications (eg, pop-ups), alerting patients that outcome estimates are available. This allows patients to decide whether they wish to view the information, thereby supporting autonomy in how and when they receive potentially sensitive data. Finally, while the tool acknowledges that outcomes are highly individual, we were unable to personalize estimates due to limited available data. Future development could further explore personalized presentations of expected benefits and risks as more real-world data become available.

Participants in the pilot test expressed concerns about the potential for numerical outcome data to diminish hope, particularly when averages were presented without sufficient context regarding variability and uncertainty about outcomes for individual patients. While many patients reviewed the visualizations of outcomes as useful for understanding PFS and side effects, others felt overwhelmed or worried that such information might have deterred them from initiating treatment. This aligns with findings from a recent multicenter survey, which reported that a subset of patients with advanced cancer actively prefers not to receive prognostic information to preserve hope and reduce anxiety [33]. Communicating variability in outcomes using outcome ranges instead of only point estimates and explicitly framing numerical data as averages is especially important in palliative care, where maintaining hope is a central aspect of patient engagement. Balancing honesty about prognosis with empathy is critical for fostering trust and minimizing distress [11]. Striking this balance is essential to align treatment decisions with patients' values and priorities. There is a need for future research on how negative emotions in SDM may be minimized [34,35].

Some study limitations need to be considered when interpreting the results of this study. Women reviewed all the components of the tool at once, instead of getting the push notifications in line with the phase of treatment they were in, as intended. Consequently, the tool was less attuned to meet women's diverse informational needs. The benefits of gradual information display and repetition were lost, which may have caused women to feel overwhelmed by (outcome) information. Moreover, the prototype testing phase included more highly educated patients, which may have influenced study outcomes. It requires further testing to determine whether the tools meet the needs of patients with low numeracy and health literacy. Also, the findings of our pilot testing do not provide insight into how the tool will fit into the clinical workflow and potential barriers to implementation. Identifying any potential barriers to clinical use, as well as identifying criteria for when to use and when not to use, is particularly important for the tool component intended for use during consultations (ie, Kijkgesprek). An important strength of this study is the use of real-world outcome data in a patient information tool, specifically to get more realistic data regarding the patient's experience with side effects and their clinical management. The breadth of patient experiences with treatment side effects captured in the real-world information is more likely to reflect the experiences of patients than information obtained from the highly selected populations in clinical trials. An important strength of the tool design process is the participatory development approach we used, involving both patients and HCPs to ensure the resulting information tool matched their needs and preferences. Finally, given the complexity of communicating and understanding outcome data, great effort has been put into presenting the information in an accessible manner to ensure that patients with low (health) literacy can also understand the information.

In conclusion, we used the context of treatment with a CDK4/6 inhibitor in combination with endocrine therapy for $HR_{pos}HER2_{neg}$ metastatic BC as a case study to assess the value and usability of real-world outcome data in informing patients about their treatment and outcomes. Women in the pilot test rated our tool highly, appreciating its clear and structured information. However, preferences for how to receive (different types of) outcome data varied. To accommodate patients' varied preferences for outcome data, we recommend that tool developers provide options for patients to choose whether they want to see additional information (eg, survival probabilities) and in how much detail. For instance, using pop-up warnings or expandable text would be beneficial. Moreover, we highly encourage future studies to evaluate the usage of our tool in clinical practice. In our study, we identified the facilitators and barriers that need to be addressed to enable the successful (nationwide) implementation of our information tool. Finally, our study can serve as a template for others who want to develop patient information tools incorporating real-world data. As most patients are open to receiving outcome data, concerted efforts should be made to set up the infrastructure to collect the

XSL•FO
RenderX

necessary data and facilitate its effective use in patient information tools.

Acknowledgments

We sincerely thank the patients who participated in this study and the health care professionals for their invaluable assistance in recruiting participants. We also extend our gratitude to the Dutch Breast Cancer Society for their valued partnership and support in facilitating participant recruitment for this project.

Funding

This work was supported by Pfizer (grant 62785943) and ZonMw as part of the "Experiment Uitkomstindicatoren Santeon."

Conflicts of Interest

AFdJ declares a potential conflict of interest in relation to the tool evaluated in this study: she is chair of the foundation that created the tool's functionality. The remaining authors declare no conflicts of interest.

Multimedia Appendix 1
Interview guide.
[DOCX File, 20 KB - cancer_v12i1e73156_app1.docx]

Multimedia Appendix 2
Steering group members.
[DOCX File, 16 KB - cancer_v12i1e73156_app2.docx]

References

1. Ministerie van Volksgezondheid Welzijn en Sport. Integraal zorgakkoord: "samen werken aan gezonde zorg". Rijksoverheid. 2022. URL: https://www.rijksoverheid.nl/documenten/rapporten/2022/09/16/integraal-zorgakkoord-samen-werken-aan-gezonde-zorg [accessed 2024-12-16]
2. van der Nat PB. The new strategic agenda for value transformation. Health Serv Manage Res 2022 Aug;35(3):189-193. [doi: 10.1177/09514848211011739] [Medline: 33900128]
3. Westerink HJ, Bresser CC, Garvelink MM, et al. The use of outcome data in patient consultations from the healthcare professionals' and patients' perspectives: a mixed methods study. Patient Educ Couns 2024 Jan;118:108043. [doi: 10.1016/j.pec.2023.108043] [Medline: 37925975]
4. Hackert MQN, Ankersmid JW, Engels N, et al. Effectiveness and implementation of SHared decision-making supported by OUTcome information among patients with breast cancer, stroke and advanced kidney disease: SHOUT study protocol of multiple interrupted time series. BMJ Open 2022 Aug 1;12(8):e055324. [doi: 10.1136/bmjopen-2021-055324] [Medline: 35914919]
5. Damman OC, Jani A, de Jong BA, et al. The use of PROMs and shared decision-making in medical encounters with patients: an opportunity to deliver value-based health care to patients. J Eval Clin Pract 2020 Apr;26(2):524-540. [doi: 10.1111/jep.13321] [Medline: 31840346]
6. Stiggelbout AM, Pieterse AH, De Haes J. Shared decision making: concepts, evidence, and practice. Patient Educ Couns 2015 Oct;98(10):1172-1179. [doi: 10.1016/j.pec.2015.06.022] [Medline: 26215573]
7. van der Horst DEM, Garvelink MM, Bos WJW, Stiggelbout AM, Pieterse AH. For which decisions is shared decision making considered appropriate? - a systematic review. Patient Educ Couns 2023 Jan;106:3-16. [doi: 10.1016/j.pec.2022.09.015] [Medline: 36220675]
8. Cherny NI, Paluch-Shimon S, Berner-Wygoda Y. Palliative care: needs of advanced breast cancer patients. Breast Cancer (Dove Med Press) 2018;10:231-243. [doi: 10.2147/BCTT.S160462] [Medline: 30584354]
9. Bélanger E, Rodríguez C, Groleau D. Shared decision-making in palliative care: a systematic mixed studies review using narrative synthesis. Palliat Med 2011 Apr;25(3):242-261. [doi: 10.1177/0269216310389348] [Medline: 21273220]
10. Rabben J, Vivat B, Fossum M, Rohde GE. Shared decision-making in palliative cancer care: a systematic review and metasynthesis. Palliat Med 2024 Apr;38(4):406-422. [doi: 10.1177/02692163241238384] [Medline: 38481012]
11. van der Velden NCA, Meijers MC, Han PKJ, van Laarhoven HWM, Smets EMA, Henselmans I. The effect of prognostic communication on patient outcomes in palliative cancer care: a systematic review. Curr Treat Options Oncol 2020 Apr 23;21(5):40. [doi: 10.1007/s11864-020-00742-y] [Medline: 32328821]
12. Henselmans I, van Laarhoven HWM, van Maarschalkerweerd P, et al. Effect of a skills training for oncologists and a patient communication aid on shared decision making about palliative systemic treatment: a randomized clinical trial. Oncologist 2020 Mar;25(3):e578-e588. [doi: 10.1634/theoncologist.2019-0453] [Medline: 32162796]

XSL•FO
RenderX

13. Cramer-van der Welle CM, Peters BJM, Schramel F, et al. Systematic evaluation of the efficacy-effectiveness gap of systemic treatments in metastatic nonsmall cell lung cancer. Eur Respir J 2018 Dec;52(6):1801100. [doi: 10.1183/13993003.01100-2018] [Medline: 30487206]
14. Phillips CM, Parmar A, Guo H, et al. Assessing the efficacy-effectiveness gap for cancer therapies: a comparison of overall survival and toxicity between clinical trial and population-based, real-world data for contemporary parenteral cancer therapeutics. Cancer 2020 Apr 15;126(8):1717-1726. [doi: 10.1002/cncr.32697] [Medline: 31913522]
15. Engels N, van der Nat PB, Ankersmid JW, et al. Development of an online patient decision aid for kidney failure treatment modality decisions. BMC Nephrol 2022 Jul 6;23(1):236. [doi: 10.1186/s12882-022-02853-0] [Medline: 35794539]
16. Cijfers over kankerCijfers en inzichten uit de Nederlandse Kankerregistratie. Integraal kankercentrum Nederland. URL: http://www.cijfersoverkanker.nl/ [accessed 2026-07-05]
17. Benitez Fuentes JD, Morgan E, de Luna Aguilar A, et al. Global stage distribution of breast cancer at diagnosis: a systematic review and meta-analysis. JAMA Oncol 2024 Jan 1;10(1):71-78. [doi: 10.1001/jamaoncol.2023.4837] [Medline: 37943547]
18. Cristofanilli M, Turner NC, Bondarenko I, et al. Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA-3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. Lancet Oncol 2016 Apr;17(4):425-439. [doi: 10.1016/S1470-2045(15)00613-0] [Medline: 26947331]
19. Finn RS, Martin M, Rugo HS, et al. Palbociclib and letrozole in advanced breast cancer. N Engl J Med 2016 Nov 17;375(20):1925-1936. [doi: 10.1056/NEJMoa1607303] [Medline: 27959613]
20. Verma S, Bartlett CH, Schnell P, et al. Palbociclib in combination with fulvestrant in women with hormone receptor-positive/HER2-negative advanced metastatic breast cancer: detailed safety analysis from a multicenter, randomized, placebo-controlled, phase III study (PALOMA-3). Oncologist 2016 Oct;21(10):1165-1175. [doi: 10.1634/theoncologist.2016-0097] [Medline: 27368881]
21. Engels N, Bos WJW, de Bruijn A, et al. Santeon's lessons from a decade of implementing value-based health care. NEJM Catalyst 2023 Dec 20;5(1). [doi: 10.1056/CAT.23.0232]
22. Vaisson G, Provencher T, Dugas M, et al. User involvement in the design and development of patient decision aids and other personal health tools: a systematic review. Med Decis Making 2021 Apr;41(3):261-274. [doi: 10.1177/0272989X20984134] [Medline: 33655791]
23. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
24. Hackert MQN, van Uden-Kraan CF, Agterof MJ, et al. Real-world palbociclib effectiveness in patients with metastatic breast cancer: focus on neutropenia-related treatment modification strategies and clinical outcomes. Cancer Treat Res Commun 2023;35:100691. [doi: 10.1016/j.ctarc.2023.100691] [Medline: 36913809]
25. Bui TBV, Burgers DM, Agterof MJ, van de Garde EM. Real-world effectiveness of palbociclib versus clinical trial results in patients with advanced/metastatic breast cancer that progressed on previous endocrine therapy. Breast Cancer (Auckl) 2019;13:1178223418823238. [doi: 10.1177/1178223418823238] [Medline: 30675102]
26. te Paske R, Vervloet M, Rademakers J, van Dijk L. Implementatie van Kijksluiter in de apotheek: een vragenlijstonderzoek onder apothekers. Nivel. URL: https://www.nivel.nl/nl/publicatie/implementatie-van-kijksluiter-de-apotheek-een-vragenlijstonderzoek-onder-apothekers [accessed 2025-12-16]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. Willig C, Rogers WS. The SAGE Handbook of Qualitative Research in Psychology 2017. [doi: 10.4135/9781526405555]
29. Stephenson JJ, Gable JC, Zincavage R, et al. Treatment experiences with CDK4&6 inhibitors among women with metastatic breast cancer: a qualitative study. Patient Prefer Adherence 2021;15:2417-2429. [doi: 10.2147/PPA.S319239] [Medline: 34764640]
30. Conley CC, McIntyre M, Pensak NA, et al. Barriers and facilitators to taking CDK4/6 inhibitors among patients with metastatic breast cancer: a qualitative study. Breast Cancer Res Treat 2022 Apr;192(2):385-399. [doi: 10.1007/s10549-022-06518-2] [Medline: 34997364]
31. Kalloger SE, Ho C, Mitton C, Regier DA. Mapping the experiences of people with advanced cancer across multiple cancer types-a scoping review. J Cancer Surviv 2024 Apr;18(2):318-324. [doi: 10.1007/s11764-022-01263-3] [Medline: 36180762]
32. Buiting HM, Rurup ML, Wijsbek H, van Zuylen L, den Hartogh G. Understanding provision of chemotherapy to patients with end stage cancer: qualitative interview study. BMJ 2011 Apr 4;342:d1933. [doi: 10.1136/bmj.d1933] [Medline: 21464103]
33. van der Velden NCA, van Laarhoven HWM, Burgers SA, et al. Characteristics of patients with advanced cancer preferring not to know prognosis: a multicenter survey study. BMC Cancer 2022 Sep 1;22(1):941. [doi: 10.1186/s12885-022-09911-8] [Medline: 36050628]
34. van de Water LF, Bos-van den Hoek DW, Kuijper SC, et al. Potential adverse outcomes of shared decision making about palliative cancer treatment: a secondary analysis of a randomized trial. Med Decis Making 2024 Jan;44(1):89-101. [doi: 10.1177/0272989X231208448] [Medline: 37953598]

XSL•FO
RenderX

35. Henselmans I, Smets EMA, Han PKJ, de Haes H, Laarhoven HV. How long do I have? Observational study on communication about life expectancy with advanced cancer patients. Patient Educ Couns 2017 Oct;100(10):1820-1827. [doi: 10.1016/j.pec.2017.05.012] [Medline: 28511804]

Abbreviations

BC: breast cancer
HCP: health care professional
$HER2_{neg}$: human epidermal growth factor receptor 2–negative
HR_{pos}: hormone receptor–positive
PFS: progression-free survival
SDM: shared decision-making
VBHC: value-based health care

Edited by J Bender; submitted 26.Feb.2025; peer-reviewed by MT Knobf, R Bhatnagar, Z Davey; revised version received 17.Mar.2026; accepted 17.May.2026; published 30.Jul.2026.

Please cite as:
Engelhardt EG, Hackert MQN, Vogelaar A, de Jong AF, Guerrero Paez C, van der Velden AWG, Putker C, Agterof MJ, van de Garde EMW, van Uden-Kraan CF
An Information Tool Incorporating Real-World Outcome Data for Women With Metastatic Breast Cancer Eligible for Treatment With a CDK4/6 Inhibitor: Development and Evaluation
JMIR Cancer 2026;12:e73156
URL: https://cancer.jmir.org/2026/1/e73156
doi:10.2196/73156

© Ellen G Engelhardt, Mariska Q N Hackert, Anne Vogelaar, Afke F de Jong, Cristina Guerrero Paez, Annette W G van der Velden, Christa Putker, Mariette J Agterof, Ewoudt M W van de Garde, Cornelia F van Uden-Kraan. Originally published in JMIR Cancer (https://cancer.jmir.org), 30.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinician Perspectives on Integrating Mobile Sensor Data Into Cancer Care: Mixed Methods Study

Christianna Bartel[1], BS; Leeann Chen[1], MPS; Krina C Durica[1], MA; Jennifer Fedor[1], MS; Akiera Palm[2], MSc; Tommy Selwood[1]; Pari Lakhia[1]; Roby Thomas[3], MD; Heidi Donovan[1], PhD; Carissa A Low[1], PhD

[1]University of Pittsburgh, 5051 Centre Avenue, Suite 5002, Pittsburgh, PA, United States

[2]Morehouse School of Medicine, Atlanta, GA, United States

[3]University of Pittsburgh Medical Center, Pittsburgh, PA, United States

Corresponding Author:
Carissa A Low, PhD
University of Pittsburgh, 5051 Centre Avenue, Suite 5002, Pittsburgh, PA, United States

Abstract

Background: Wearable devices are becoming more ubiquitous and are capable of capturing health-relevant information that patients may be interested in sharing with their providers. However, limited research has been conducted on oncology provider perspectives on how these data could be used to inform cancer care.

Objective: The goal of this study was to understand oncology clinicians' preferences about which data would be most clinically valuable and in what clinical scenarios, the benefits and barriers to integrating wearable device data into cancer care, and perspectives on how wearable device data could impact decision-making using 3 clinical vignettes.

Methods: A total of 13 oncology care clinicians completed an online questionnaire to assess the perceived value of different types of wearable device data in different clinical scenarios and participated in semistructured interviews to gather preferences around integrating these data into clinical workflows. During the interviews, providers were also presented with 3 clinical vignettes and asked for clinical recommendations both before and after seeing the patient's wearable device data. Descriptive statistics were calculated to summarize quantitative data from questionnaires and structured interview questions, and interviews were transcribed and coded using an iterative thematic analysis approach.

Results: Survey responses indicated that providers were most interested in tracking vital sign metrics, followed by data related to falls and functioning, and then by data on sleep and activity. Clinicians thought that wearable device data might be especially useful for remotely monitoring patients at high-risk moments in their care trajectory, such as after an acute hospitalization or after starting a new outpatient treatment. Four main themes were discussed by providers in the interviews: (1) corroborating reports, (2) identifying new issues, (3) coordinating care, and (4) patient-provider communication. Although there were no statistically significant differences in clinical recommendations before and after viewing wearable device data for any vignette (all $P>.25$), all clinicians reported that the wearable device data impacted their decision-making confidence, and most rated the wearable device data as helpful.

Conclusions: Oncology providers highlighted the potential clinical value of vital sign and physical functioning data from wearable devices, particularly when outpatients might be at risk for readmissions or other acute deteriorations between clinic visits. Providers noted that the objective data captured by consumer wearable devices can be helpful complements to patient and caregiver subjective reports and that this information could improve patient-provider communication and care coordination. During the interviews, most providers found wearable device data to be helpful when making decisions. While there are challenges to address on how to integrate this information into the clinical workflow and communicate alerts with patients, there is cautious enthusiasm among clinicians about how these data could inform and improve cancer care.

(JMIR Cancer 2026;12:e86412) doi:10.2196/86412

KEYWORDS

wearable devices; oncology providers; cancer; symptom management; doctor-patient communication; remote monitoring

Introduction

Background

Data from wearable devices can capture objective patterns of physical activity and physiology that may be useful for remotely monitoring patients during chemotherapy and other outpatient cancer treatments [1]. Work by our team and others suggests that it is feasible to collect wearable device data from patients during cancer treatment [2-4], that patients are willing to share these data with clinicians [5], and that these data are correlated

XSL•FO
RenderX

with clinical outcomes among patients with cancer [6,7]. As these devices become more ubiquitous and powerful, research is needed to understand how best to integrate these devices and data into clinical cancer care.

In previous work, noncancer physicians reported that wearable device data may improve patient-provider communication by efficiently providing objective information about patient health, informing decision-making by monitoring response to treatment and symptom progression outside of the clinic, and helping patients feel like more active participants in their care [8,9]. Physicians also noted potential barriers to the clinical implementation of wearable device monitoring, including concerns about the digital health literacy of older patients, the potential cost of devices widening digital divides, device data increasing anxiety for some patients, and increased strain on clinical workloads [10].

In the area of oncology care, several recent studies have assessed the preferences and perspectives of patients with cancer on wearable health monitors [11,12]. Barriers noted by patients with cancer included fears that wearable devices may serve as a constant reminder of their illness and concerns that technology may replace human interactions if, for example, patients grow to rely on devices to communicate their health status to clinicians and refrain from reaching out directly. Patients also reported benefits, noting that wearable devices can be motivating, can provide new insights into their health and illness, and that they viewed the wearable device as a part of their treatment relationship and wanted their cancer treatment teams to have access to their wearable device data.

Objective

To date, limited work has focused on cancer care clinicians' perspectives on how wearable device data could inform cancer care. The goal of this questionnaire and interview study was to understand oncology clinicians' preferences about which data would be most clinically valuable and in what clinical scenarios, the benefits and barriers to integrating wearable device data into cancer care, and perspectives on how wearable device data could impact decision-making using 3 clinical vignettes.

Methods

Recruitment

Oncology care clinicians (N=13) were recruited through purposive sampling to capture different health professionals involved in providing cancer care. Providers were eligible if they were (1) working as clinical providers caring for patients receiving cancer treatment, (2) aged over 18 years, and (3) fluent in English.

Study Design

Participants first completed an online questionnaire to collect demographic and clinical background information, assess which sensor information they would be interested in receiving and in what scenarios, and explore potential resources that might be helpful to them in integrating wearable device data into their clinics (Multimedia Appendix 1). Semistructured interview guides were developed by the research team with a focus on preferences and needs around how technology could support care management and fit into clinical workflows (Multimedia Appendix 2).

Based on a similar study in pediatric surgery [13], participants were also presented with 3 clinical vignettes using wearable data collected as part of the remote oncology symptom assessment research protocol that followed patients receiving cytotoxic chemotherapy for solid tumors [4]. Patient participants in this study wore a Fitbit Inspire device for 90 days, and the data of 3 participants who had experienced complications during the study were presented at the end of semistructured interviews to determine whether the availability of wearable device data impacted clinician decision-making. The three vignettes presented were as follows: (1) a male patient with stage IV pancreatic cancer, aged 75 years, who received chemotherapy 2 days prior and reported worsening abdominal pain, nausea, vomiting, anorexia, and weakness over the past few days; (2) a male patient with stage IV pancreatic cancer, aged 58 years, who had last received chemotherapy 10 days prior and who reported dizziness as well as intermittent fever and chills with some abdominal pain over the past 4 to 5 days; and (3) a female patient with stage IV breast cancer, aged 56 years, whose treatment plan changed from chemotherapy to targeted therapy 2 weeks prior, following an unplanned hospitalization for cellulitis, and who now reported worsening fatigue, nausea, and decreased appetite.

The 3 clinical vignettes were presented to each clinician participant in 2 formats: without and with wearable data. First, the scenario was presented without wearable data, and participants were asked to triage the patient and determine the urgency of follow-up care. Clinicians were then asked to rate their likelihood of recommending that the patient present to the emergency department (ED) immediately using a 10-point Likert scale, with 0 representing "not at all likely" and 10 representing "very likely."

The scenario was then augmented with wearable data. Participants were shown visualizations of wearable device data (including heart rate, gait speed, and step count data collected around the time of the phone call or visit) and were asked to rate the likelihood of recommending ED presentation using the same 10-point Likert scale. They were also asked if the wearable data increased their confidence in their recommendation and to rate how helpful the wearable device data was to their understanding of the situation on a 10-point Likert scale with 0 representing "not at all helpful" and 10 representing "extremely helpful."

Interviews lasted between 20 and 40 minutes and were conducted remotely over Zoom (Zoom Communications Inc) by members of the research team (LC and CB). Interviews were digitally recorded, transcribed, deidentified, and entered into Dovetail (Dovetail Research Pty Ltd), a qualitative analysis software. Study team members conducted interviews until thematic saturation was achieved.

Analysis

Interview transcripts were analyzed using an iterative thematic analysis approach [14]. A coding team, comprised of 4 authors,

XSL•FO
RenderX

read all the transcripts and developed preliminary codes. A few selected transcripts were coded by 2 research team members to ensure consistency in coding and to group codes into larger thematic classifications. One team member coded all the transcripts. Participant quotes were lightly edited for clarity.

Descriptive statistics, including frequencies, medians, and IQR, were calculated to summarize quantitative data from questionnaires and structured interview questions. Due to the small size of our sample, nonparametric Wilcoxon signed-rank tests were performed to compare clinicians' recommendations for ED presentation without wearable device data to their recommendations after viewing wearable device data for each vignette. An α level of .05 was used as a cutoff for statistical significance. Analyses were performed in Python version 3.10 (Python Software Foundation).

Ethical Considerations

This study was conducted at a National Cancer Institute–Designated Comprehensive Cancer Center. The institutional review board of the University of Pittsburgh approved all study activities (STUDY23120104). Informed consent was obtained from all individual participants included in the study. Participant data were identified only by anonymized study ID numbers and stored in secure locations. Data collection was conducted between March 2024 and June 2024. Participants received US $100 in compensation for their participation.

Results

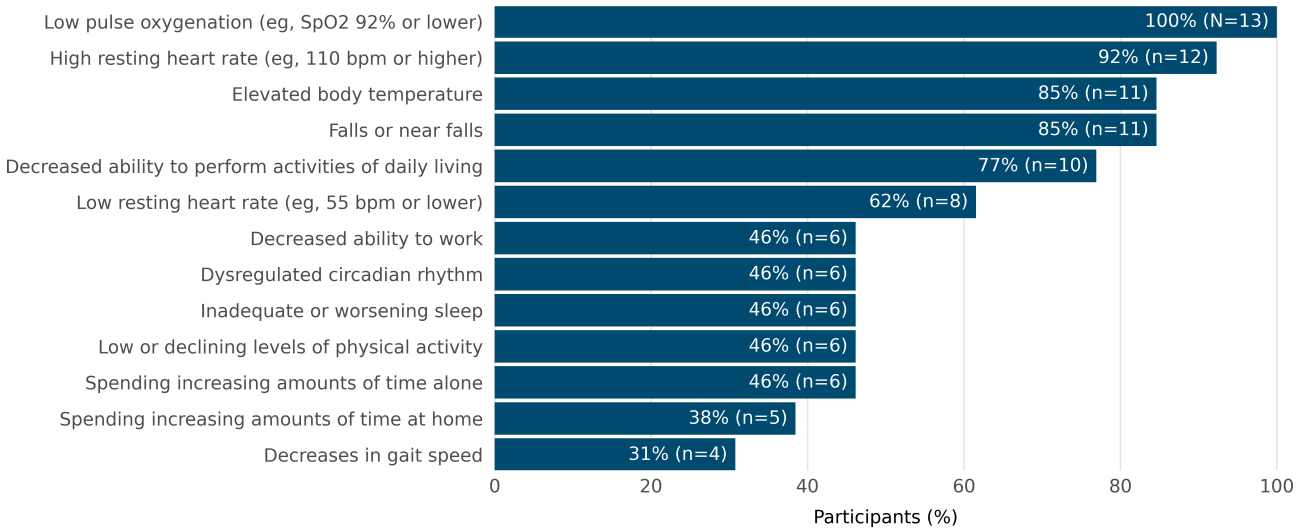
Participant Characteristics

Participants included physicians (n=4), physician assistants (n=5), nurses (n=3), and a certified registered nurse practitioner (n=1) who had a mean age of 42 (SD 12.60; range 25 - 61) years, were primarily White (12/13, 92%), and had been treating patients with cancer for an average of 9.7 (SD 6.42) years with a range from 5 months to 22 years.

Survey

The majority of providers (11/13, 85%) indicated that they wanted to know if their patients experienced low pulse oxygenation, high resting heart rate, elevated body temperature, and falls or near falls (Figure 1). All providers considered wearable device data to be moderately to very important when monitoring a patient after an acute hospitalization, managing toxicities during outpatient chemotherapy, and tracking the impact of a new treatment on a patient's quality of life and function (Figure 2). When asked what would allow them to integrate wearable device data into their care, all providers responded that information about how to monitor and interpret wearable device data; having an additional member to monitor, respond to, and notify the team as appropriate; and integration of the wearable device data into the medical record system would be moderately to very helpful (Figure 3).

Figure 1. Number of participants who indicated they would want to know the following information about their patients.



XSL•FO
RenderX

Figure 2. Potential clinical scenarios in which wearable devices might be valuable.

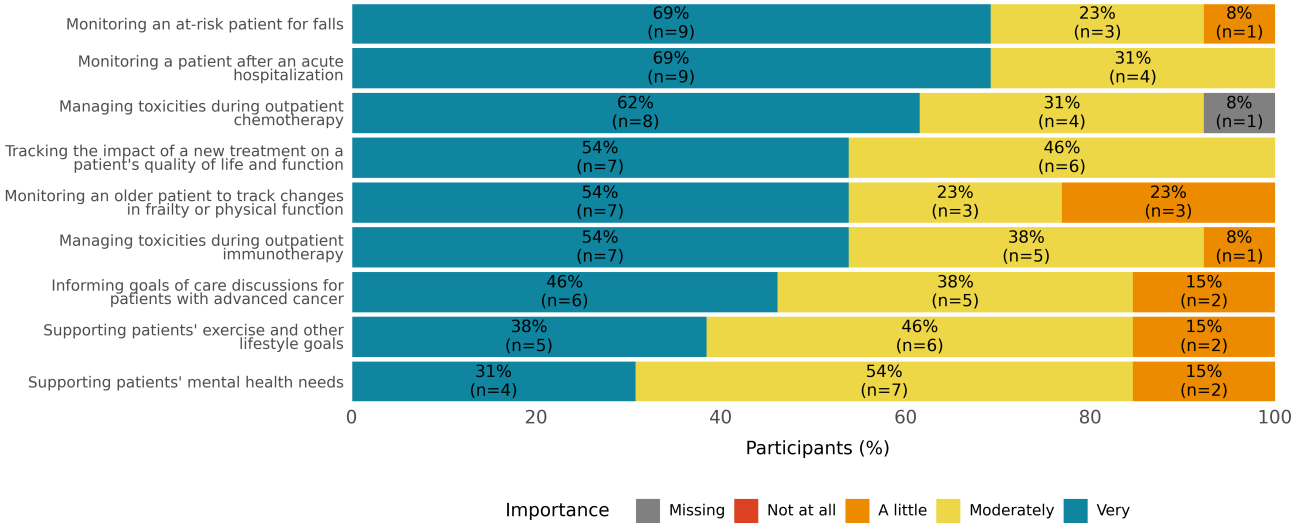
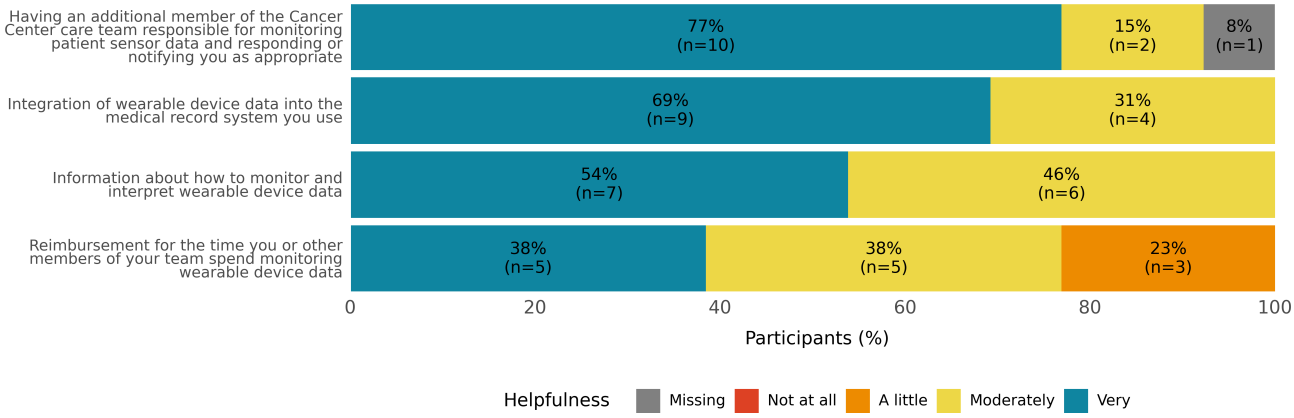


Figure 3. Perceived helpfulness of various supports for integrating wearable device data in the clinic.



Themes

During the semistructured interviews, four primary themes related to the benefits and barriers of integrating wearable device data into care were discussed: (1) corroborating reports, (2) identifying new issues, (3) coordinating care, and (4) patient-provider communication.

Theme 1: Corroborating Reports

Providers discussed how objective wearable device data could provide oncology clinics with a better understanding of patients' concerns and symptoms between clinic visits.

> *It would be nice if somebody's calling me with questions, and I can just pull it up and see what's really going on...It's easier for people to kind of not understand the full grasp of things when they're just talking. When you could just actually see the graph, it would probably be nice. But, also, I think it's important for the providers to see it, because whenever they come in and they're being seen by the providers, it would be good for them to be able to quantify what's going on between visits.* [P014, RN]

In particular, several providers mentioned that wearable device data could be used to help resolve conflicting reports they receive from patients and their support persons.

> *Especially if maybe I was getting one story from the patient and a different story from their caregiver, or I was surprised by what I was hearing.* [P007, MD]

> *Sometimes I feel like patients don't want to be dose reduced and they'll lie. And then I try to usually get the family to rat them out like, "Okay, so are they really grocery shopping and cutting the lawn?" They're like, "No, they're sitting on the couch all day."* [P003, MD]

Theme 2: Identifying New Issues

Some providers discussed the benefits of having wearable device data for identifying new issues that patients may not be reporting. One provider gave an example of a common scenario in which symptoms are typically experienced early in the chemotherapy cycle, but patients do not have another office visit for several weeks. This can lead to symptoms not being detected by the clinic or reported by the patient.

> *But I feel like that would be really helpful, because when we see the patients for chemotherapy, we're generally seeing them at the beginning of a cycle, right? And I feel like we have these people who rally and come in the office and look great because it's been 3 weeks or 4 weeks, or something since their last chemo, and sometimes the data that we would have to kind of say, "Hey, you know, we might have*

XSL•FO
RenderX

> *to dose reduce you, because this whole first 10 day span, here was what was happening from the data that we got back, even if you're not telling us that." So, I feel like that would be a valuable tool.* [P016, PA]

Theme 3: Coordinating Care

Oncology patients often have a number of doctors with whom they must coordinate their care. A few providers additionally discussed how integrated wearable device data could improve patient care coordination across various clinicians and specialties.

> *[T]here's a lot of collaboration between teams. And I think, sometimes the timing of different therapies, and how people are responding, can kind of be indicative. So potentially in multi-disciplinary collaboration between, like, "Hey is a patient, you know, how do we know that they're recovered from their chemo to get surgery? How do we know that they're recovered from their surgery to get chemo? How do we know that they're not having these toxicities from radiation that would make us prevent them from going on?" So, I think there'd be a role to kind of integrate it from a medical oncology perspective, but also kind of have interplay between the groups, to kind of look at what the best avenue of approach of treatment might be. And that's kind of outside of just, individual patient and clinician reporting, and that kind of thing. Just getting other teams involved, utilizing the data.* [P016, PA]

Theme 4: Patient-Provider Communication

Finally, when discussing whether patients should be notified about significant changes in their wearable device data, most providers thought it would be beneficial for promoting patient advocacy.

> *Yeah, I think their kind of awareness can only help the situation, even if it's having a family member call in and say, "Yeah, we did notice this, this pulse ox was really low," or "Yeah, she is having more trouble." Even if they make the contact and it's not necessarily just automatic from the data. If the patients alerted too, I think advocacy for themselves would be helpful.* [P008, PA]

> *Yes, so if they...became febrile or, their pulse ox was low, I would want them to be alerted. That way they could either call us if we weren't getting back to them, or assess how they're feeling. If your temp's 100.0 but you feel okay, that's not necessarily something we need to do anything about. But [it] would allow them, I think, to say, "Hey, this was my temp, but I felt fine," or the opposite. And I think that would be helpful to correlate those two things together.* [P015, PA]

However, several providers raised concerns about how these notifications would be triggered and suggested that careful wording should be used to prevent increasing patient anxiety.

> *That's a little bit harder, because I think it's like, what does that mean? Right? What if it's natural variation for them. I don't want to alarm them. I don't want them to think that there's something critically wrong with them, if they otherwise feel fine, right. If your heart rate bounced up a little bit, but then it comes back down again. They probably already know if they're not sleeping well, right. They probably already know if they feel bad. If it's something that's like a critical alert, then yes. Right. But that's also a little bit nuanced. What's critical? You have to set a value? But I don't know that they need feedback on every little variation.* [P005, MD]

> *I feel like, in a sense, that could be good. But then, in another sense, it could just stress the patient out further if it's a weekend, and somebody's not getting back to them ASAP that might send them into a spiral.* [P004, RN]

Vignettes

Results related to the wearable device data vignettes are presented in Table 1. Across all 3 vignettes, the median difference in the likelihood of recommending ED presentation before and after viewing wearable device data was 0 (Q1 range: −0.5 to 0; Q3 range: 0 to 0.875). There were no statistically significant differences in the likelihood of recommending ED presentation before and after viewing wearable device data for any vignette (all $P>.25$). All clinicians reported that the wearable device data impacted their decision-making or confidence, and clinicians generally rated the wearable device data as helpful.

XSL•FO
RenderX

Table . Vignette responses (N=13).

Metric	Vignette 1	Vignette 2	Vignette 3
Likelihood of recommending ED[a,b], median (IQR)			
Without wearable data	3 (3 to 5)	9 (8 to 9.5)	2 (0.75 to 2.625)
With wearable data	5.25 (2.875 to 6.25)	9 (9 to 10)	2 (1 to 2.5)
Difference[c]	0 (−0.5 to 0.875)[d]	0 (0 to 0)[e]	0 (−0.5 to 0)[f]
Change in recommendation with wearable data, n (%)			
More likely (difference>0)	4 (31)	3 (23)	2 (15)
No change (difference=0)	4 (31)	10 (77)	6 (46)
Less likely (difference<0)	4 (31)	0 (0)	4 (31)
Missing	1 (8)	0 (0)	1 (8)
Wearable data impacted decision making or confidence, n (%)	10 (77)	10 (77)	9 (69)
Helpfulness of wearable data[b], median (IQR)	7 (4 to 8)	8 (5 to 9)	7 (6 to 8)

[a]ED: emergency department.

[b]Rating on 0 - 10 scale.

[c]Wilcoxon signed-rank test (W=test statistic; *P*=*P* value; *r*=effect size).

[d]Vignette 1: W=13.5, *P*=.64, *r*=0.12.

[e]Vignette 2: W=0.0, *P*=.25, *r*=0.32.

[f]Vignette 3: W=7.0, *P*=.56, *r*=0.15.

Discussion

Principal Results

The goal of this mixed methods study was to examine oncology clinicians' perspectives on integrating wearable device data into clinical decision-making and care. Overall, providers noted many potential benefits of collecting wearable device data during outpatient treatment for patients with cancer and sharing these data with clinical care teams. Survey responses helped pinpoint which wearable device data providers would most like to be notified about, with clinicians indicating that they were most interested in tracking vital sign metrics, followed by data related to falls and functioning, and then by data on sleep and activity. Clinicians thought that wearable device data might be especially useful for remotely monitoring patients at high-risk moments in their care trajectory, such as after an acute hospitalization or after starting a new outpatient treatment.

Themes from interviews highlight that wearable device data could serve several useful functions in the context of cancer care, including providing objective data to corroborate, clarify, or reconcile contradictory subjective reports; helping to identify new issues that emerge when the patient is outside of the clinic; and providing objective baseline and historical data that might serve as a foundation for the coordination of multidisciplinary cancer care. Clinicians further noted both positive and negative implications for patient-provider communication. While wearable device data could help patients with cancer take a more active role in monitoring and managing their health, clinicians also noted that the patient-facing interfaces and notifications must be thoughtfully implemented to prevent undue patient anxiety and distress. Providing information on how to interpret wearable device data, adding an additional team member to monitor notifications about changes in these data, and integrating this information into the electronic medical record were judged as being potentially moderately to very helpful for future clinical integration.

Clinicians' responses to vignettes in which clinical scenarios were presented with or without visualizations of wearable device data collected from patients further demonstrated the potential clinical value of this technology, with clinicians rating the helpfulness of the data visualizations highly and noting that the data visualizations increased confidence in their clinical recommendations even though these recommendations were not changed by the data. Even if recommendations do not shift, increased clinician confidence in decisions and recommendations may be clinically significant, with benefits that could include enhanced patient trust in clinicians, reduced time spent on clinical decision-making, and decreased clinician burnout.

Comparison With Prior Work

These results are consistent with findings from previous studies in which noncancer physicians reported that wearable device data could improve communication and symptom management [8,9]. Similarly, oncology physicians were also concerned about whether alerts might cause undue anxiety for their patients [10]. Beyond the previously documented benefits and disadvantages, providers in this study indicated which specific wearable device data would be most relevant to their clinical needs and identified the support necessary to integrate this information into their clinical workflow. Furthermore, although providers' ED recommendations did not change significantly before and after

XSL•FO
RenderX

viewing the wearable device data, they reported increases in confidence comparable to those observed in the pediatric surgery vignette study that informed this aspect of the present study [13].

Limitations

This study used multiple methods, including online surveys, semistructured interviews, and clinical vignettes augmented with visualizations of oncology research patients' wearable device data, to begin to probe cancer clinicians' perspectives on wearable devices in cancer care. Limitations of this study include the small sample of clinicians from a single academic cancer center who were interested in participating in a study about wearable devices and who may be favorably biased toward digital health. The small sample also constrained our ability to examine different perspectives across disciplines (eg, among physicians vs nurses), disease specialties, or different clinics or care facilities, and exploration of how perspectives differ across these roles and settings should be a focus of future research. The survey and interview questions were not adapted from validated instruments or pilot-tested and suffered from various conceptual and methodological weaknesses. Additionally, the vignettes were based on a convenience selection of available wearable device data from our research group and therefore focused on patients with stage IV cancer who used a Fitbit Inspire device [7]. Findings may be different if vignettes focused on patients with early-stage cancer undergoing curative treatment or if data from a different wearable device were presented. Because these vignettes were selected retrospectively from a research study, they may not reflect real-world clinical scenarios where this information would be best applied.

Despite these limitations, this work contributes to a growing body of literature on the value of wearable and other mobile sensors in clinical oncology and highlights cautious enthusiasm among clinicians about how these data could inform and improve cancer care.

Future Directions

Findings suggest that oncology providers recognize the clinical value of integrating wearable device data into their workflow and that such integration may enhance their confidence in decision-making. More work is needed to identify the threshold at which wearable device data signals increasing risk as well as the best way to communicate this risk to providers and patients to avoid straining clinical resources or increasing patient anxiety or clinician alert fatigue. Research has suggested that early comprehensive training and additional support within the care team can mitigate burnout, but further investigation is needed into how this will be structured within oncology clinics to minimize cost and burden [15]. Additionally, future studies should examine how providing clinicians with access to wearable device data impacts clinical decision-making, patient-provider communication, and patient outcomes.

Future studies exploring wearable device data integration into oncology care should also acknowledge issues related to cost and access. Disparities exist along socioeconomic lines as well as between rural and urban areas, which, if left unchecked, could result in further inequity and widen the digital divide. Those with lower incomes, less educational achievement, and those living in rural areas are less likely to own wearable devices [16,17]. Simply supplying these devices to patients may not be sufficient either, as previous work has demonstrated that demographic and clinical factors, such as non-White race and cognitive impairments, have been linked to lower adherence to wearable devices [4]. Further investigation is needed into ways to improve digital health literacy in patients with cancer, in addition to improving access, to ensure wearable device data integration does not continue to advantage only certain populations.

In conclusion, oncology providers view patient-generated wearable device data as potentially valuable in the cancer care setting. They indicated that access to this information could improve communication between providers and patients, as well as between collaborating clinicians. Displaying wearable data alongside symptom data may also be helpful in contextualizing patient and caregiver reports. While the wearable device data visualizations did not change hypothetical clinical recommendations, providers reported that the data increased confidence in their decisions. As wearable devices become more common, it becomes increasingly important to have systems in place that allow these new real-world data streams to be integrated into the oncology clinic workflow.

Acknowledgments

The authors gratefully acknowledge the providers who participated in this research.

Funding

This project was supported in part by the National Cancer Institute (R37CA242545, R37CA242545-S1, and P30CA047904) and by the CAncer Research Equity-Gleitsman Scholars Program.

Data Availability

The deidentified datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: CB, CAL, HD, JF, KCD, LC, RT

XSL•FO
RenderX

Data curation: AP, CB, CAL, JF, KCD, PL, TS
Formal analysis: CB, JF
Funding acquisition: CAL
Investigation: CB, LC
Methodology: CB, CAL, JF, KCD, LC
Project administration: CB, CL, KCD, LC
Resources: CB, CAL, KCD, LC
Supervision: CAL
Visualization: CAL, JF
Writing – original draft: CB, CAL, JF, KCD, TS
Writing – review and editing: CB, CAL, JF, KCD, TS

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interviews for remote oncology symptom assessment 2 provider questionnaire.
[PDF File, 51 KB - cancer_v12i1e86412_app1.pdf]

Multimedia Appendix 2
Interviews for remote oncology symptom assessment 2 provider interview guide.
[PDF File, 260 KB - cancer_v12i1e86412_app2.pdf]

References

1. Hawley ST, Haymart MR. The potential of wearable devices in cancer care delivery. JAMA Oncol 2024 May 1;10(5):573-574. [doi: 10.1001/jamaoncol.2024.0001] [Medline: 38546699]
2. Beauchamp UL, Pappot H, Holländer-Mieritz C. The use of wearables in clinical trials during cancer treatment: systematic review. JMIR mHealth uHealth 2020 Nov 11;8(11):e22006. [doi: 10.2196/22006] [Medline: 33174852]
3. Huang Y, Upadhyay U, Dhar E, Kuo LJ, Syed-Abdul S. A scoping review to assess adherence to and clinical outcomes of wearable devices in the cancer population. Cancers (Basel) 2022 Sep 13;14(18):4437. [doi: 10.3390/cancers14184437] [Medline: 36139602]
4. McClaine S, Fedor J, Bartel C, Chen L, Durica KC, Low CA. Engagement with daily symptom reporting, passive smartphone sensing, and wearable device data collection during chemotherapy: longitudinal observational study. JMIR Cancer 2024 Dec 10;10(1):e57347. [doi: 10.2196/57347] [Medline: 39656513]
5. Rising CJ, Gaysynsky A, Blake KD, Jensen RE, Oh A. Willingness to share data from wearable health and activity trackers: analysis of the 2019 Health Information National Trends Survey data. JMIR mHealth uHealth 2021 Dec 13;9(12):e29190. [doi: 10.2196/29190] [Medline: 34898448]
6. Kos M, Brouwer CG, van Laarhoven HWM, Hopman MTE, van Oijen MGH, Buffart LM. The association between wearable device metrics and clinical outcomes in oncology: a systematic review with evidence synthesis and meta-analysis. Crit Rev Oncol Hematol 2023 May;185:103979. [doi: 10.1016/j.critrevonc.2023.103979] [Medline: 37001837]
7. Low CA, Bartel C, Durica K, Donovan HS, Thomas R, Fedor J. Consumer wearable device measures of gait cadence and activity fragmentation as predictors of survival among patients undergoing chemotherapy. JCO Clin Cancer Inform 2025 Jul;9:e2500111. [doi: 10.1200/CCI-25-00111] [Medline: 40644640]
8. Alpert JM, Manini T, Roberts M, et al. Secondary care provider attitudes towards patient generated health data from smartwatches. NPJ Digit Med 2020;3(1):27. [doi: 10.1038/s41746-020-0236-4] [Medline: 32140569]
9. Loos JR, Davidson EJ. Wearable health monitors and physician-patient communication: the physician's perspective. Presented at: 2016 49th Hawaii International Conference on System Sciences (HICSS); Jan 5-8, 2016; Koloa, HI, USA p. 3389-3399. [doi: 10.1109/HICSS.2016.422]
10. Volpato E, Toniolo S, Pagnini F, Banfi P. The relationship between anxiety, depression and treatment adherence in chronic obstructive pulmonary disease: a systematic review. Int J Chron Obstruct Pulmon Dis 2021;16:2001-2021. [doi: 10.2147/COPD.S313841] [Medline: 34262270]
11. Stuijt DG, van Doeveren EEM, Kos M, et al. Remote patient monitoring using mobile health technology in cancer care and research: patients' views and preferences. JCO Clin Cancer Inform 2024 Nov;8:e2400092. [doi: 10.1200/CCI.24.00092] [Medline: 39531613]
12. Collinson S, Ingram-Walpole S, Jackson C, et al. Patient experiences of using wearable health monitors during cancer treatment: a qualitative study. Clin Oncol (R Coll Radiol) 2025 Jan;37:103664. [doi: 10.1016/j.clon.2024.10.036] [Medline: 39550935]

XSL•FO
RenderX

13. Carter M, Linton SC, Zeineddin S, et al. Impact of consumer wearables data on pediatric surgery clinicians' management: multi-institutional scenario-based usability study. JMIR Perioper Med 2024 Nov 12;7:e58663. [doi: 10.2196/58663] [Medline: 39531288]
14. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
15. Thomas Craig KJ, Willis VC, Gruen D, Rhee K, Jackson GP. The burden of the digital environment: a systematic review on organization-directed workplace interventions to mitigate physician burnout. J Am Med Inform Assoc 2021 Apr 23;28(5):985-997. [doi: 10.1093/jamia/ocaa301] [Medline: 33463680]
16. Adepoju O, Dang P, Nguyen H, Mertz J. Equity in digital health: assessing access and utilization of remote patient monitoring, medical apps, and wearables in underserved communities. Inquiry 2024;61:469580241271137. [doi: 10.1177/00469580241271137] [Medline: 39323052]
17. Nagappan A, Krasniansky A, Knowles M. Patterns of ownership and usage of wearable devices in the United States, 2020-2022: survey study. J Med Internet Res 2024 Jul 26;26:e56504. [doi: 10.2196/56504] [Medline: 39058548]

Abbreviations

ED: emergency department

Edited by M Balcarras; submitted 24.Oct.2025; peer-reviewed by J Johnson, N Goyal; revised version received 18.Mar.2026; accepted 19.Mar.2026; published 08.May.2026.

Please cite as:
Bartel C, Chen L, Durica KC, Fedor J, Palm A, Selwood T, Lakhia P, Thomas R, Donovan H, Low CA
Clinician Perspectives on Integrating Mobile Sensor Data Into Cancer Care: Mixed Methods Study
JMIR Cancer 2026;12:e86412
URL: https://cancer.jmir.org/2026/1/e86412
doi:10.2196/86412

© Christianna Bartel, Leeann Chen, Krina C Durica, Jennifer Fedor, Akiera Palm, Tommy Selwood, Pari Lakhia, Roby Thomas, Heidi Donovan, Carissa A Low. Originally published in JMIR Cancer (https://cancer.jmir.org), 8.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of Experiences With Ecological Momentary Assessment Among Women With Metastatic Breast Cancer: Qualitative Study

Emily P Taylor[1], BS; Brenna Mossman[1], PhD; Mikaela A Velazquez-Sosa[1], BA; Zaynab Rashid[1]; Natalie Kloster[1], BS; Maureen Schwarz[1]; Hannah Kang[1], BS; Jennifer D Rodriguez[1], BA; Elizabeth L Addington[2], PhD; Judith Tedlie Moskowitz[2], MPH, PhD; Lesley Glenn[3], BS; Shontè Drakeford[4], MSN, CRNP, AGNP-C; Roxana Guerra[5]; Claudine Isaacs[1], MD; Ami Chitalia[6], MD; Christopher Gallagher[6], MD; Deena Graham[7], MD; Suzanne C O'Neill[1], PhD; Claire C Conley[1], PhD

[1]Lombardi Comprehensive Cancer Center, Georgetown University, 2115 Wisconsin Ave NW, Suite 300, Washington, DC, United States

[2]Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[3]Patient Author, Project Life MBC, Calabasas, CA, United States

[4]Patient Author, Upper Marlboro, MD, United States

[5]Patient Author, Washington, DC, United States

[6]MedStar Washington Hospital Center, Washington, DC, United States

[7]John Theurer Cancer Center, Hackensack Meridian Health, Hackensack, NJ, United States

Corresponding Author:
Claire C Conley, PhD
Lombardi Comprehensive Cancer Center, Georgetown University, 2115 Wisconsin Ave NW, Suite 300, Washington, DC, United States

Abstract

Background: Patients with metastatic breast cancer (MBC) experience significant quality-of-life decrements, but there are few supportive care interventions specifically designed for this group that significantly improve quality of life. Ecological momentary assessment (EMA) and related ecological momentary interventions (EMIs) may be particularly beneficial for patients with MBC. However, no studies have previously examined the use of EMIs in the context of metastatic cancer.

Objective: The purpose of this study was to qualitatively assess experiences with EMA and preferences for intervention content and mode among patients living with MBC, with an emphasis on EMIs.

Methods: Women with MBC (n=29) were recruited from a longitudinal, observational study of quality of life using an EMA design. In-depth qualitative interviews assessed participants' perspectives on the EMA design, including its feasibility, acceptability, and relevance to participants' MBC experiences. Guided by participants' EMA data, the interviews also examined the impact of self-monitoring and possible interventions when patients reported high symptom burden and/or low quality of life. Interviews were audio-recorded, transcribed verbatim, and analyzed using open coding, axial coding, and selective coding. The codebook was developed by reviewing a randomly selected subset of transcripts (n=5) and creating inductive, data-driven codes using the raw data, which were deductively organized into overarching themes.

Results: Participants were mainly White (17/29, 59%), heterosexual (23/29, 79%), currently working (17/29, 59%), and held at least a bachelor's degree (17/29, 59%). Participants had been living with MBC for a median of 2.5 years (range 0.2 - 16.6 years). Most were diagnosed with hormone receptor-positive (23/29, 79%) and HER2-negative (21/29, 72%) breast cancer. Qualitative analysis identified 5 major themes. The participants reflected on their reasons for enrolling in the study, including interest in study activities and giving back to the MBC community (theme 1: participant engagement). Most participants found the EMA format easy to complete, and they provided recommendations for improving the design of future studies, including altering the timing and frequency of questionnaires (theme 2: feedback on study design). While less common, some participants discussed how engaging with the EMA protocol altered their experiences or behaviors (theme 3: impact of self-monitoring). Participants also discussed their reactions to seeing their EMA data, including their mixed thoughts on real-time data sharing (theme 4: responses to data). Finally, participants suggested programs and resources to improve their overall quality of life, reflecting on their interest in real-time interventions and peer-to-peer support (theme 5: recommendations for future interventions).

Conclusions: Patients with MBC are willing to use EMA methodologies for data collection on quality of life and are open to EMIs with varying content and formats. Given variations in daily functioning between and within patients, a just-in-time adaptive intervention framework may be well suited for this population.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e80467) doi:10.2196/80467

KEYWORDS

quality of life; metastatic cancer; breast cancer; ecological momentary assessment; ecological momentary intervention; feasibility; acceptability; qualitative data

Introduction

Metastatic breast cancer (MBC) is breast cancer that has spread to other parts of the body; it is typically considered to be treatable but incurable [1]. Approximately 5% - 10% of breast cancers are metastatic at diagnosis, and 20% - 30% of early-stage breast cancers will progress to MBC [2-4]. In the past decade, novel treatment approaches have resulted in significantly increased survival times following MBC diagnosis [5-10]. For this reason, the MBC population is anticipated to increase in the coming years; by 2030, an estimated 246,000 people in the United States will be living with MBC (a 59% increase from 2017) [11,12].

MBC symptoms, treatment side effects, and their impact on daily functioning influence quality of life (QOL) in this population [13,14]. Given the incurable nature of the disease, QOL is a priority for patients with MBC [15-17]. Interventions that improve QOL for people with MBC are vital; not only is QOL an important outcome in and of itself, but improvements in QOL may also enable patients to adhere to life-sustaining therapies [18].

However, there are few supportive care interventions that significantly improve QOL for people with MBC. A 2023 systematic review identified 11 studies testing supportive care interventions specifically designed to improve QOL among individuals living with MBC [19]—only 3 demonstrated QOL improvements. More recent trials have found that health-related QOL in patients with MBC significantly increased 6 months after the onset of a supervised exercise program [20]. Two small pilot trials, one examining the impact of brief psychotherapy [21] and another evaluating a whole-food, plant-based diet [22], demonstrated effects in some QOL domains, but not others. Thus, further research is needed to develop, test, and optimize interventions targeting QOL for patients with MBC.

Ecological momentary interventions (EMIs) may be particularly beneficial to help ameliorate QOL in patients with MBC. EMIs deliver dynamic, individually tailored treatments to patients in real time and in natural settings [23,24], using ecological momentary assessment (EMA) methodology (ie, repeated, in situ reporting) to identify optimal times to provide support. This type of adaptive intervention approach has several potential benefits over traditional "one-size-fits-all" interventions: maintaining participant engagement, sustaining continued behavior change for longer durations, and ultimately achieving greater intervention effects [25]. Given the substantial variability in QOL trajectories among patients with metastatic cancer generally [26]—and the widely acknowledged heterogeneity of the disease and treatment response in MBC specifically [27-29]—EMIs could be ideally suited to meet the needs of patients with MBC with the most QOL decrements without overburdening those who are currently stable. However, EMIs have rarely been used in cancer survivorship research [30]. To our knowledge, there are no studies examining EMIs in the context of metastatic cancer, and further information is therefore needed to inform their use within this population.

The purpose of this study was to qualitatively assess preferences for intervention content and mode among patients living with MBC, with an emphasis on EMIs. Participants were recruited from a longitudinal, observational study of QOL using an EMA design [31,32]. In-depth qualitative interviews assessed participants' perspectives on the EMA design, including its feasibility, acceptability, and relevance to participants' MBC experiences. As symptom monitoring can improve symptom control for patients with cancer [33-35], the interviews also examined the impact of self-monitoring. Finally, guided by participants' EMA data, interviews solicited suggestions for possible interventions when patients reported high symptom burden and/or low QOL.

Methods

Procedures and Participants

This study used qualitative exit interview data from a longitudinal study of QOL among patients with MBC. Data collection and analyses were guided by a pragmatist philosophical orientation and align with COREQ (Consolidated Criteria for Reporting Qualitative Research) [36].

Between January 2023 and April 2024, participants were recruited from 5 sites within the Georgetown Lombardi Comprehensive Cancer Center Network (in Washington DC, Maryland, and New Jersey). Eligible participants were assigned female at birth, aged 18 years and older, diagnosed with MBC, able to speak English or Spanish, and had a working telephone capable of receiving SMS text messages and accessing the internet. Study team members screened the electronic medical record to identify potentially eligible patients, obtained approval to approach participants from their treating oncologist, approached potentially eligible patients via telephone, and confirmed patient eligibility. Eligible, interested patients provided written informed consent via electronic signature on web-based consent forms. Following informed consent, participants completed a brief baseline survey and a 4-week EMA protocol. Participants completed 3 EMA prompts per day, 1 day per week (randomly assigned each week), for 4 consecutive weeks; thus, participants completed up to 12 EMA surveys. This "measurement burst" design incorporates short periods of intensive, recurrent assessment that are repeated longitudinally and spaced out over longer intervals. This design reduces participant burden while maintaining density of assessments [37,38]. EMA prompts consisted of a text message with a link to a REDCap (Research Electronic Data Capture) survey (Multimedia Appendix 1). Each survey assessed global QOL, symptom severity (depression, anxiety, pain, fatigue,

XSL•FO
RenderX

slowed cognitive functioning, appetite loss, nausea, gastrointestinal [GI] distress, and decreased libido), and positive experiences (social connectedness, peace, and joy). Participants had a 3-hour window to complete each EMA. Thus, EMA surveys enabled assessment of change in each domain over the course of the day and over the course of 1 month (the approximate length of a treatment cycle for a patient with MBC).

Upon completion of EMA data collection, participants were queried regarding their interest in and willingness to participate in an exit interview. Initially, we planned to offer the exit interview to all study participants. However, in the first 3 months of data collection, a relatively large number of participants were interested in participating in the exit interview (20/24, 83% of participants enrolled in the first 3 months). As it would not be feasible to conduct interviews and analyze qualitative data from more than 80% of study participants (target n=120), we elected to change our approach and randomly sample 15% of participants who were interested in the exit interview. Sampling for the interview was stratified by recruitment site, such that 15% of participants from each site who were interested in the exit interview were invited to participate. We engaged in constant comparative analysis [39] and conducted interviews until theoretical saturation was reached [40].

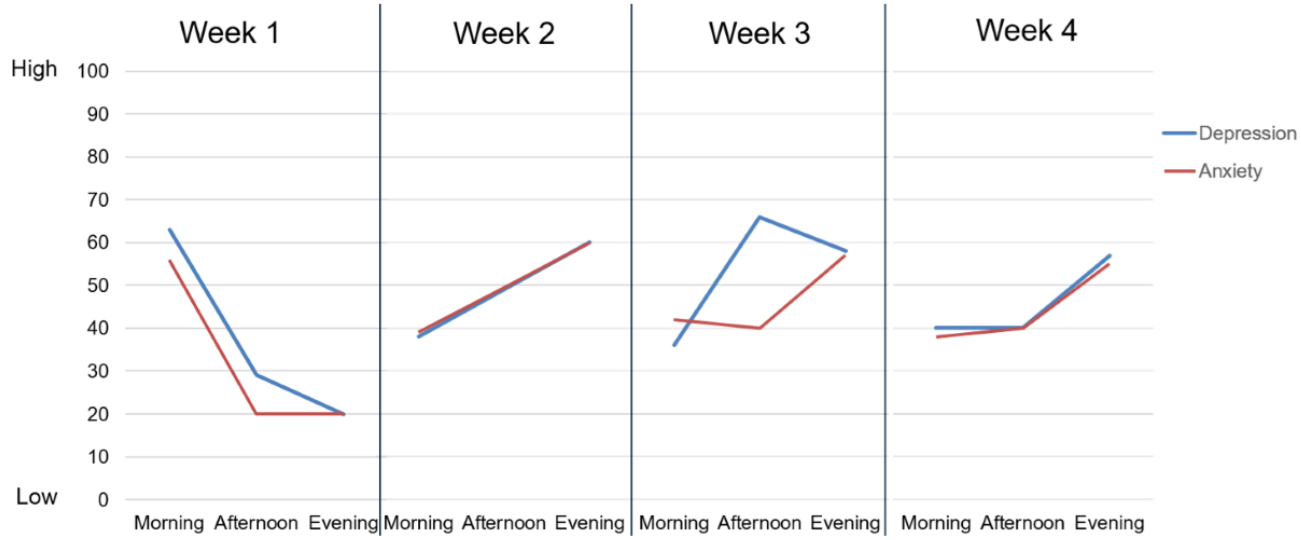
Three female members of the study team with prior training and experience in qualitative interviewing (CCC, MAVS, and JDR) conducted semistructured qualitative interviews in English or Spanish with selected participants. The interviewers did not have a relationship with the interviewees prior to their participation in the study but had sometimes interacted with the participants during the EMA phase. Interviewers reintroduced themselves and reminded participants of the rationale for the study at the start of the interview. The interviewers had foundational knowledge of MBC but not firsthand experience. They also tended to be younger than the interview participants. These characteristics may introduce bias by affecting the interviewer's application of probing questions. Potential bias was managed through postinterview debriefing with colleagues.

The interview guide (Multimedia Appendix 2) was developed by the study team in collaboration with patient advocates with MBC. We first identified broad topic areas of interest that included both feedback on their experiences in this study and the perceived acceptability of future research projects, including EMIs. Once the topic areas for the interview were identified, the group brainstormed a wide variety of probing questions that could be used to elicit additional input from participants. Thus, the resulting interviews began with open-ended questions about participants' experiences with the EMA protocol and symptom self-monitoring. After the open-ended questions, participants were shown their EMA data in graphical format (see Figure 1 for an example). Specifically, participants were shown a series of line graphs representing their responses to the EMA surveys over the 4 weeks in the study. Participants viewed graphs representing overall QOL, mental health (anxiety and depression), physical health (pain, fatigue, nausea, and GI distress), thinking and feeling (slowed cognitive functioning, appetite loss, and decreased libido), and positive experiences (social connectedness, peace, and joy). The EMA data display figures were developed by the study team and iteratively revised based on feedback from patient advocates with MBC to ensure clarity prior to their use in the interviews. For each graph, the interviewer oriented the participant to the horizontal axis, vertical axis, and legend. The interviewer also answered any questions about the data visualization. Participants were then asked about their reactions to their data, how they might change their behavior in response to these data, and their preferences for data-guided symptom management interventions.

Interviews took place via telephone or web-based video platform, based on participant preference. When interviews took place via telephone, participants were mailed or emailed the EMA data display figures and were instructed to refer to them during the telephone call. Interviews lasted 44 minutes on average (range 26‐76 minutes), were audio-recorded, transcribed verbatim, and translated to English (if necessary). The interviewer took field notes during the interviews. As participants did not consent to be recontacted, we did not conduct repeat interviews, return transcripts to participants for review, or have participants provide feedback on the findings.

XSL•FO
RenderX

Figure 1. Example ecological momentary assessment (EMA) data display figure, showing one participant's responses to EMA questions on anxiety and depression across 12 time points (morning, afternoon, and evening; 1 day per week for 4 consecutive weeks). The vertical axis represents the visual analog scale ranging from 0 to 100 and the horizontal axis represents the EMA time point. The blue lines indicate depression scores and the red lines indicate anxiety scores.



Data Analysis

We used a combined deductive and inductive qualitative data analysis approach [41,42]. First, 3 members of the study team (CCC, MAVS, and EPT) reviewed a randomly selected subset of transcripts (n=5) and collaboratively developed a detailed codebook consisting of data-driven codes from the raw data [43]. These codes were then organized based around core concepts of feasibility and acceptability as reflected by the topic areas in the interview guide. Thus, the data-driven codes represented inductive subthemes that were deductively organized into themes based on our belief that certain core concepts exist in the data [44,45]. Following the development of the codebook, all transcripts were independently coded by 2 of 6 raters (MAVS, ZR, NK, EPT, MS, and HK). Coding disagreements were resolved through discussion with a third rater (CCC) until consensus was reached. Coding was conducted using Dedoose.

Ethical Considerations

All procedures were reviewed and approved by the Institutional Review Board at Georgetown University (institutional review board number 00005081). All participants provided written informed consent via electronic signature on web-based consent forms and confirmed ongoing consent verbally at the time of the qualitative interview. Informed consent included consent for the use of quotes in research publications. Qualitative data are deidentified to protect participant confidentiality. Participants received US $10 for completing the baseline survey, up to US $25 for completing EMA surveys, and US $20 upon interview completion. Thus, participants could earn up to US $55 for completing all study activities. Participant incentives were provided as a gift card.

Results

Participant Demographics

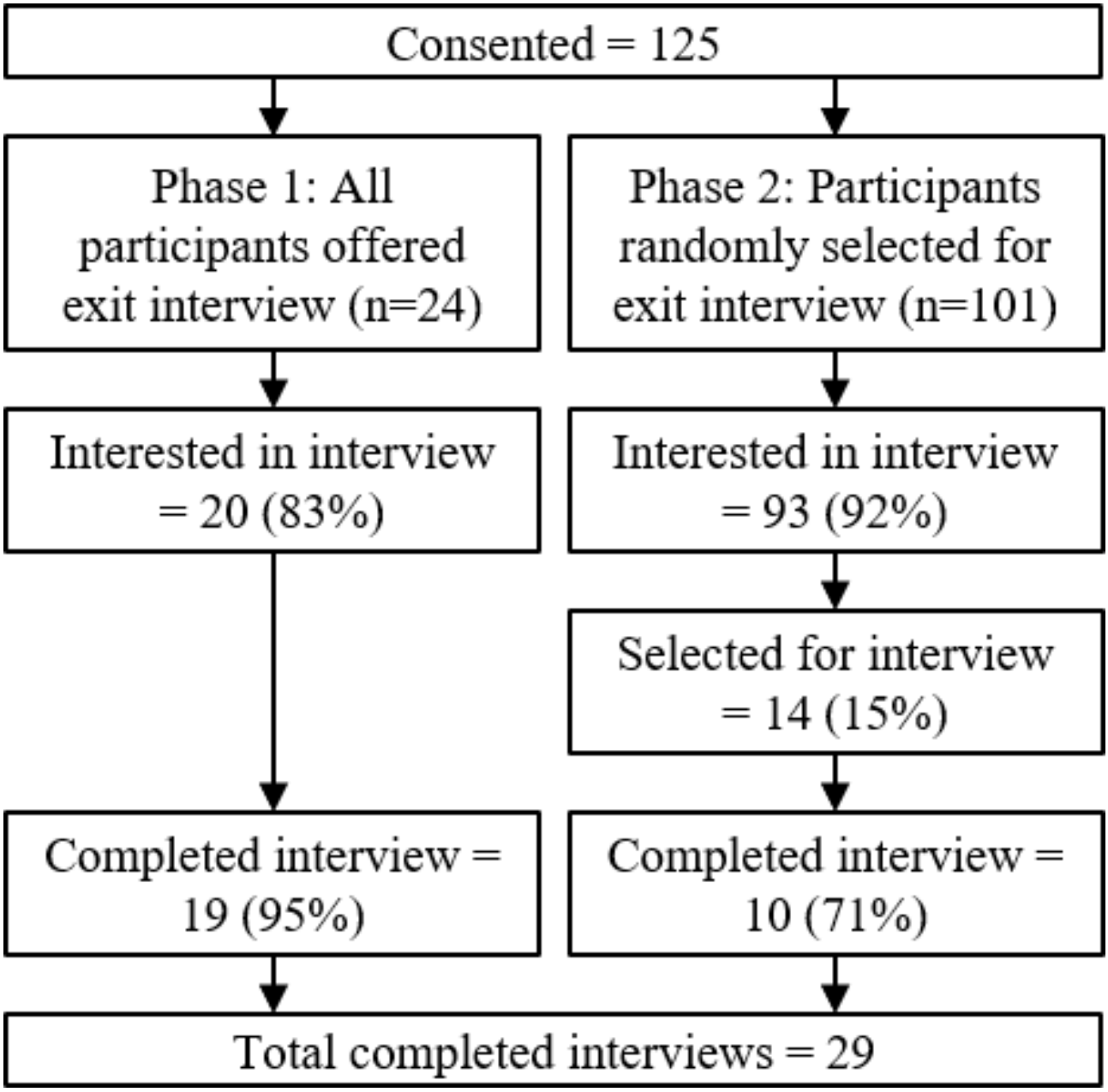
Of the 125 eligible patients who consented, 113 out of 125 (90%) expressed interest in the exit interview, 34 out of 113 (30%) were invited to participate in the interview, and 29 out of 34 (85%) completed the interview (Figure 2).

See Table 1 for a full description of the interview participants. The median participant age was 57 years. The majority were non-Hispanic White (17/29, 59%), currently working (17/29, 59%), and held a bachelor's degree or greater (17/29, 59%).

Participants had been living with MBC for a median of 2.5 years (range 0.2 - 16.6). Most participants were diagnosed with hormone receptor positive (23/29, 79%) and HER2 negative (21/29, 72%) breast cancer. The most common treatments received were targeted therapy (20/29, 69%) and hormone therapy (16/29, 55%); chemotherapy (7/29, 24%), immunotherapy (5/29, 17%), and radiation therapy (1/29, 3%) were less common. Most participants had an Eastern Cooperative Oncology Group performance status rating of 0 (13/29, 45%) or 1 (11/29, 38%), indicating no performance restrictions and restrictions in strenuous physical activity, respectively.

XSL•FO
RenderX

Figure 2. Study flow diagram.



XSL•FO
RenderX

Table . Participant demographic and clinical characteristics (N=29).

Characteristics	Values
Demographic characteristics (self-reported)	
Age (years)	
Mean (SD)	54.8 (11.5)
Median (range)	57 (33-74)
Race and ethnicity[a], n (%)	
Asian	1 (3)
Black/African American	8 (28)
Hispanic/Latina	3 (10)
Native Hawaiian/Pacific Islander	1 (3)
White	17 (59)
Preferred language, n (%)	
English	27 (93)
Spanish	2 (7)
Sexual orientation, n (%)	
Heterosexual or straight	23 (79)
Lesbian	3 (10)
Bisexual	1 (3)
Missing/prefer not to answer	2 (7)
Partner status, n (%)	
Partnered	18 (62)
Not partnered	10 (35)
Missing/prefer not to answer	1 (3)
Education, n (%)	
Less than bachelor’s degree	12 (41)
Bachelor’s degree or greater	17 (59)
Employment status, n (%)	
Currently working	17 (59)
Not working	10 (35)
Missing/prefer not to answer	2 (7)
Annual household income, n (%)	
<US $100,000	18 (62)
≥US $100,000	7 (24)
Missing/prefer not to answer	4 (14)
Health insurance type[a], n (%)	
Private	23 (79)
Public (eg, Medicare and Medicaid)	9 (31)
Missing/prefer not to answer	1 (3)
Clinical characteristics (abstracted from EMR)[b]	
Years since MBC[c] diagnosis	
Mean (SD)	4.3 (4.1)
Median (range)	2.5 (0.2-16.6)

XSL•FO
RenderX

Characteristics	Values
Hormone receptor status, n (%)	
Positive	23 (79)
Negative	6 (21)
HER2 status, n (%)	
Positive	8 (28)
Negative	21 (72)
Treatment(s) during study[a], n (%)	
Chemotherapy	7 (24)
Radiation therapy	1 (3)
Immunotherapy	5 (17)
Targeted therapy	20 (69)
Hormone therapy	16 (55)
ECOG[d] performance status, n (%)	
0	13 (45)
1	11 (38)
2	1 (3)
Missing	4 (14)
Study characteristics	
Study site, n (%)	
MedStar Georgetown University Hospital	10 (34)
MedStar Washington Hospital Center	9 (31)
MedStar Good Samaritan Hospital	2 (7)
MedStar Franklin Square Medical Center	4 (14)
Hackensack Meridian Health/John Theurer Cancer Center	4 (14)
Number of EMA[e] surveys completed	
Mean (SD)	10.7 (1.6)
Median (range)	11 (6-12)

[a]As participants could select more than 1 response for race and ethnicity and health insurance type, and could be receiving multiple treatments simultaneously, counts for these categories sum to more than the sample total.

[b]EMR: electronic medical record.

[c]MBC: metastatic breast cancer.

[d]ECOG: Eastern Cooperative Oncology Group.

[e]EMA: ecological momentary assessment.

Themes

Qualitative analysis identified 5 major themes. Table 2 provides theoretically driven themes and the component subthemes derived from the raw qualitative data. Textbox 1 provides exemplar quotes from participants for themes 1 - 4. Exemplar quotes for theme 5 are included in Textbox 2.

XSL•FO
RenderX

Table . Theoretically driven themes and data-driven subthemes.

Themes	Subthemes
Participant engagement	• Motivation for participation or altruism • Feeling seen • Interest in study activities

Feedback on study design	• Demystifying or clarity about study
	• Flexibility and accessibility
	  • Frequency of assessments
	  • Timing of assessments
	  • Response options
	• Technology
	• Relevance of the questions

Impact of self-monitoring	• Emotional reactions to study activities or procedures • Heightened awareness • Behavior changes
Responses to data	• Reflecting on the data • Positive experiences • Reaction to the idea of real-time data sharing
Recommendations for future interventions	• Content • Timing • Format • Audience • Peer support

XSL•FO
RenderX

Textbox 1. Representative quotes illustrating themes 1-4.

Participant engagement

- "I like being part of a study in that I think it's good to get a wide range of how people experience cancer and everything like that, so." [Participant 16, 9 years post-MBC diagnosis, completed 12 EMA surveys]
- "These things are good experiences for people. I mean, you guys get insight and I get to think about how I actually am feeling at those times." [Participant 19, 1 year post-MBC diagnosis, completed 8 EMA surveys]
- "Well, I really respect my oncologist. So the fact that she had passed me along for it made me want to do it. And then I think looking at quality of life issues for people with metastatic breast cancer, I felt a connection to the study." [Participant 20, 9 years post-MBC diagnosis, completed 12 EMA surveys]

Feedback on study design

- "It was very easy to participate. And I haven't given it much thought because I would just do it. And it wasn't difficult in any way." [Participant 12, 5 years post-MBC diagnosis, completed 10 EMA surveys]
- "Instead of looking at overall quality of life, we are being asked to look at it in snippets of time. And I think you can get a skewed answer that way. It really depends on how the participant is looking at the question." [Participant 27, 3 years post-MBC diagnosis, completed 7 EMA surveys]
- "I think the three times in one day, I don't know if that was really needed. I would rather have done more [days] because it would key in more on if you're having a good day or a bad day. Because I can remember some of my bad days. I thought, 'Oh, I wish I had the questions.' [My answers] would be so different." [Participant 17, 2 years post-MBC diagnosis, completed 12 EMA surveys]
- "I appreciated the fact that I could take the survey on my phone, so no matter where I was, I was able to access it. That was really good." [Participant 4, 10 years post-MBC diagnosis, completed 10 EMA surveys]
- "I wish that I could have done it over email because sometimes it's hard for me to get to my phone. But I liked it when you guys texted me the night before so I knew that I would have to check my phone a few times a day." [Participant 14, 5 months post-MBC diagnosis, completed 12 EMA surveys]
- "I thought the questions were very good. They touched on a lot of different parts of your life that you really wouldn't normally think about." [Participant 17, 2 years post-MBC diagnosis, completed 12 EMA surveys]
- "Some of [the questions] didn't apply to me. I'm not sexually active, so I didn't know how to gauge that." [Participant 4, 10 years post-MBC diagnosis, completed 10 EMA surveys]

Impact of self-monitoring

- "It did make me more cognizant and mindful. It made me think, 'Is there something I can do for myself?' Having those little pauses was actually nice. And at times it made me kind of grateful in that moment that not everything was challenging." [Participant 24, 3 years post-MBC diagnosis, completed 12 EMA surveys]
- "It just made me think about what I'm going through because sometimes you just go through the day, and you don't pay much attention to what you're feeling. And this really made me think more about what's happening, and if it wasn't good, how can I change it or how could I feel different...It makes you ask yourself, 'How am I taking care of myself? Am I doing a good job or am I not?' Because that's important." [Participant 17, 2 years post-MBC diagnosis, completed 12 EMA surveys]
- "I don't think I changed [my behavior]. I think I became more mindful of things...I can't think of an instance where I was like, 'Oh, let me go do more self-care.' I think it just made me mindful of what I was doing." [Participant 6, 2 years post-MBC diagnosis, completed 11 EMA surveys]
- "I was more aware of my behavior, aware of myself like I was more conscious of how I was feeling on what I was doing. For instance, when I tell myself, 'I don't feel like exercising today,' or, 'I'm too tired to go downstairs on the treadmill,' but then I have to tell myself, 'You'll feel better. You need it.' So just things like that. It just made me more aware...It's like you're holding yourself accountable." [Participant 18, 3 years post-MBC diagnosis, completed 12 EMA surveys]

Responses to data

- "I was a little hesitant [to look at my data] because I kind of was expecting this...Unfortunately, I think this is probably pretty accurate and describes my highs and lows and the stress and overwhelm that I had been feeling...But I guess the anxiety led to more of the overwhelm, feeling alone, challenging, more depressive thoughts and feelings. So unfortunately, this is probably really, really accurate." [Participant 24, 3 years post-MBC diagnosis, completed 12 EMA surveys]
- "Overall, I try to stay very positive, but when you get granular with specific feelings and experiences and whatnot, yeah, it probably doesn't equate to the quality of life being so high...As I look at [my data], I'm like, 'Well, how the heck do you get that [high] quality of life then?' Because it just doesn't seem like anything should—none of these average to the quality of life that I thought I had. But at the end of the day—so food didn't taste good to me today and whatever, but is it really affecting my quality of life? I'm just happy to be here. Just very lucky. Very, very thankful." [Participant 25, 1 year post-MBC diagnosis, completed 11 EMA surveys]

XSL•FO
RenderX

Textbox 2. Representative quotes illustrating theme 5 and related subthemes.

Content

- "[I would like] suggestions that I could implement myself without needing assistance from someone else. Definitely things that would have the quickest effect." [Participant 3, 2 years post-MBC diagnosis, completed 11 EMA surveys]
- "You can send more tips of that meditation, of just that deep breathing...I think deep breathing makes a big difference. Even if you don't know how to meditate, to teach someone how to do the deep breathing, to help them through those anxiety attacks will help." [Participant 18, 3 years post-MBC diagnosis, completed 12 EMA surveys]
- "Maybe you can send ways to calm yourself down in the moment. Like take a step back, breathe in, breathe out, put your arms up, whatever. Things like that. Or if you need to punch a pillow or something. So it might be more intense, but it has ideas that are safe for you to get it out. Go in the bathroom and let it out. Find yourself a place or just take a pause in whatever you're doing, just different things like that." [Participant 7, 2 months post-MBC diagnosis, completed 10 EMA surveys]
- "Some kind of question to prompt thought. Like, 'When have you last eaten, or have you had some water recently? Can you get a few minutes of fresh air or talk to somebody or think of an affirmation or write something down?' Any of those kinds of things, I guess. Just something to be more mindful and practical." [Participant 24, 3 years post-MBC diagnosis, completed 12 EMA surveys]
- "I guess for nausea, to remind you of all the interventions and to make sure that you're on top of it. It's so important with nausea to stay ahead of it. Same with pain management or GI distress, really, too." [Participant 20, 9 years post-MBC diagnosis, completed 12 EMA surveys]
- "I know that there are studies about how getting out in nature lifts your spirits and exercise and that sort of thing. So maybe encouragement to do that...Or kind of preemptive solutions that might hit a larger number of people. Like, 'Oh, look, they reminded me this week that it's spring so your allergies might be worse because your immune system is compromised so do this. Exercise inside instead of jogging outside.'" [Participant 15, 8 years post-MBC diagnosis, completed 12 EMA surveys]
- "My doctors tell me that they have resources out there that will help a person. So if I was feeling depressed, I would probably want to be guided somewhere where I can talk to a mental health person, or talk to someone about a type of medication, or support groups or stuff like that. Maybe you can tell me, 'Okay. If you use this link, that can take you to where you need to get the help that you think you might need.' Because there's so much information out there. Sometimes you can't concentrate." [Participant 1, 1 year post-MBC diagnosis, completed 10 EMA surveys]

Timing

- "If you rank between this and this [on the survey], maybe they could send them to you just as a, 'Hey, you're not alone. We appreciate your feedback. And while it may not totally resolve your matter, here are some suggestions.' Right after you finish that survey for that time period or something because some people may not know what to do, just suggestions." [Participant 7, 2 months post-MBC diagnosis, completed 10 EMA surveys]
- "I feel like [it would suggest] things that I already do, which are listen to music and meditate and pray. But just, I guess, suggestions to help in the moment." [Participant 13, 1 year post-MBC diagnosis, completed 12 EMA surveys]
- "I think a prompt would be nice. [Something] that's nudging you to go and look at it, but if you don't want to, you can opt out. I think the idea of having a nudge to go look at something would be nice based off reporting something as being lower and then that kind of correlating with the app prompting." [Participant 23, 2 years post-MBC diagnosis, completed 12 EMA surveys]
- "I think it depends. That week when I was having a really tough time, I might not have been able to take another piece of advice that day. So maybe it's a follow-up on the next day...Just based on that one particular day, I might have just been like, 'No, you have no idea. Nothing is going to help me today.' But maybe the next day? Like, 'Okay. The next time you find yourself having that kind of thinking or this, why don't you maybe try one of these suggestions? Keep this in your pocket.'" [Participant 25, 1 year post-MBC diagnosis, completed 11 EMA surveys]

Format

- "If there's a website you can go to and it has different topics—'I'm feeling a lot of fatigue today. What are some of the things I can do?' You click on a button and see, 'Go outside and take a brisk walk.' Yes, practical application is what I'm looking for." [Participant 3, 2 years post-MBC diagnosis, completed 11 EMA surveys]
- "Nothing too lengthy. For me, my attention span isn't all that great...I don't mind reading but nothing that's going to take more than like 15 minutes of my time. I just get distracted and impatient." [Participant 13, 1 year post-MBC diagnosis, completed 12 EMA surveys]
- "Well, I mean, people could choose. There's websites you can go on to get free [meditation videos] and there's also websites that have some wonderful healing music on too." [Participant 9, 4 years post-MBC diagnosis, completed 9 EMA surveys]

Audience

- "It's so complicated because there are people who suffer from very difficult symptoms, who sometimes don't feel like going for a walk. But if the family supports them, encouraging them to take a break, 'Let's just walk for 15, 20, 30 minutes' that can help a lot. Because there are times when those chemotherapies are so strong, it's really difficult to tell a person, 'Go for a walk.' So, if the family supports them, or friends accompany them, maybe the person will be a little more encouraged." [Participant 29, 10 years post-MBC diagnosis, completed 12 EMA surveys]
- "My husband is not one of those people that likes to talk or even believes in therapy, but he's a people person. So he's related to some of his coworkers and things like that that have had wives or spouses that have had cancer. I know that that's helped him. I just wish that there was somebody for my kids and my husband to talk to." [Participant 7, 2 months post-MBC diagnosis, completed 10 EMA surveys]

XSL•FO
RenderX

Peer support

- "What would help? Sometimes talking to somebody...If I'm really down in the dumps, I would not call my friends, though, because they don't understand. But I would call another cancer patient. That's how I would handle it for me. I think that would help me, because I know who I can call." [Participant 8, 8 years post-MBC diagnosis, completed 11 EMA surveys]
- "I think reaching out just by phone to talk to somebody, just having a phone buddy or something like that. I think having someone to talk to would be important. Not so much to complain, but to let them know what's going on. And then maybe have the person talk about good self-care." [Participant 9, 4 years post-MBC diagnosis, completed 9 EMA surveys]

Theme 1: Participant Engagement

Participants primarily enrolled in the study because they wanted to give back to the MBC community. Participants cited an altruistic drive as motivation for joining the study, stating "it's my turn to give back a little bit" [Participant 26, 9 years post-MBC diagnosis, completed 10 EMA surveys]. Other participants expressed a desire to join the study to provide representation for specific subpopulations within the MBC community. For example, one participant stated that they enrolled in the study for "the opportunity to give some data points for women of color" [Participant 22, 2 years post-MBC diagnosis, completed 11 EMA surveys]. Furthermore, participants appreciated the study's focus on their QOL, which they felt was an understudied dimension of MBC: "This is an area where there could be more support and research and more information out there for people who are just so busy and dealing with [MBC] and living with [MBC]" [Participant 24, 3 years post-MBC diagnosis, completed 12 EMA surveys].

Participants also appreciated the low-burden nature of the study procedures. The 1-month assessment time frame meant that they "didn't have to make a big commitment" [Participant 26, 9 years post-MBC diagnosis, completed 10 EMA surveys], and the lack of financial restrictions to enroll encouraged participation because "it doesn't cost anything" [Participant 7, 2 months post-MBC diagnosis, completed 10 EMA surveys].

Theme 2: Feedback on Study Design

Most participants felt the EMA design was straightforward. Almost all participants noted that the daily assessments were simple to complete, especially given the reminders provided the day prior. One participant reflected that "It was quick and straightforward, very, very user-friendly, and very easy for somebody who could be forgetful like me and just be dealing with a lot all at once. It didn't feel like it was one more thing that I had to do, which was really nice" [Participant 24, 3 years post-MBC diagnosis, completed 12 EMA surveys].

However, some participants reflected on a lack of clarity in study activities. Several participants noted that they did not fully understand the study procedures until they were actively completing them. For example, one participant had thought that they would receive surveys daily, rather than 1 day each week. Other participants did not grasp the overarching goal of the study and/or how the design would answer the research questions. As one participant shared, "I guess I just didn't really understand what [the researchers] were looking for" [Participant 21, 17 years post-MBC diagnosis, completed 11 EMA surveys].

When asked about alternative study designs, participants suggested fewer assessments per day, with a larger response window, for multiple days in a row. Some participants had trouble responding to surveys within 3 hours. One participant noted, "Sometimes, I found it difficult to keep up with actually answering the questions in a timely manner. Just regular everyday time management sometimes didn't work out the way I needed it to" [Participant 3, 2 years post-MBC diagnosis, completed 11 EMA surveys]. Thus, allowing a longer response window could improve response rates. Generally, participants felt that the 3 surveys administered per day were more than sufficient to adequately capture their experiences that day; some noted that 1 or 2 surveys would have sufficed. However, participants frequently stated that completing surveys several days in a row would be a better reflection of their experiences. As one participant put it, "There were days that I wasn't doing well, but those weren't the days that the questions were. So in a way, I felt like it wasn't accurate to how I was feeling all the time" [Participant 17, 2 years post-MBC diagnosis, completed 12 EMA surveys].

Participants had mixed experiences with the technology used to deliver the study surveys. Some participants enjoyed completing surveys on their cell phone, as this allowed them to do the survey in any location and did not interrupt their daily plans. Other participants would have preferred email surveys because it was sometimes difficult for them to get to the phone.

Another commonly cited technology issue was the use of sliders to capture symptom data, rather than a Likert-style response scale. Participants felt that the slider did not always reflect their experiences and that inadvertently moving the slider resulted in inaccuracies in their data. Numerical answers were preferred over sliders "because, with numbers, I could have done what I wanted to tell you with more precision" [Participant 10, 1 year post-MBC diagnosis, completed 11 EMA surveys].

Relatedly, many participants wished the study had involved more open-ended questions or room for comments. One participant noted:

> *I thought it was a little too restrictive regarding the type of questions they were asking and my limited ability to be able to express what I was actually trying to think about, being forced to gauge high or low, where it would have helped me [is] if I could put in a little comment section. This is what I'm feeling. This is why I'm feeling it.* [Participant 4, 10 years post-MBC diagnosis, completed 10 EMA surveys]

Participants felt that most questions resonated with their MBC experiences, with the exception of one: libido. Several participants noted that the questions resonated with their lived experiences with MBC, with the surveys serving as both a clarifying and reflective experience. For example, one

XSL•FO
RenderX

participant reflected, "It was easy [to answer the questions] because I'm living it. So it's not really something I have to think about" [Participant 19, 1 year post-MBC diagnosis, completed 8 EMA surveys].

The one exception to this pattern was the item assessing changes in libido. This item prompted many negative reactions; participants felt "uncomfortable" [Participant 7, 2 months post-MBC diagnosis, completed 10 EMA surveys] or "sad" [Participant 16, 9 years post-MBC diagnosis, completed 12 EMA surveys] and were "sick of being asked about [it]" [Participant 15, 8 years post-MBC diagnosis, completed 12 EMA surveys]. Confusion was another common reaction. Many participants did not realize that changes in libido can be a side effect of cancer treatment and were confused about why this would be included in a study of QOL in MBC. Others understood libido "as sex" [Participant 8, 8 years post-MBC diagnosis, completed 11 EMA surveys], which led to some confusion on the premise of the question and what was being asked, with the sense that it was not relevant for their personal situation (particularly for participants who were not partnered). Participants suggested that future studies assessing libido should include a "not applicable" response option.

Theme 3: Impact of Self-Monitoring

Some participants described negative emotional responses to the study activities or procedures. For example, one participant stated they felt "upset" answering questions about their QOL on "one of the days [they were not] really feeling good" [Participant 28, 3 years post-MBC diagnosis, completed 10 EMA surveys]. Another participant described their experience in detail:

> *I mean, I'm not sure if upsetting is the right word, but it was kind of like, "I don't want to go to that space right now." Sometimes I was like, "I'm focused on something completely different, and that's going to shift my perspective." The survey itself doesn't take very long to fill out, but yeah, your mind does kind of go in that space, and you start to think about that, which is distracting, potentially, to what your current day's achievement outcomes and focus that you want it to be...when I finished the surveys or if it was a survey day, I don't think I left feeling upset. But I think, yeah, a lighter version of distressing...* [Participant 23, 2 years post-MBC diagnosis, completed 12 EMA surveys]

In other words, the survey prompt tended to interrupt participants' usual routines and, in doing so, sometimes caused distraction or distress as they reflected on their symptoms and experiences.

While participants generally expressed heightened awareness in response to the study assessments, the extent to which this resulted in behavior change varied. Several participants expressed that completing the study assessments made them feel more self-aware, and they were appreciative of the opportunity to reflect on their feelings and experiences:

> *It was amazing to actually capture my daily activities so I would be prepared to answer the questions. So I really was very focused on me and my feelings, which sometimes I will try to avoid and just stay busy. But this actually made me kind of reassess myself and see that I want to keep focusing on certain things, or just go on and enjoy life and do some of the things I wanted to do.* [Participant 5, 2 years post-MBC diagnosis, completed 6 EMA surveys]

While this reflective experience prompted some to adopt certain behavior changes—such as exercise, yoga, or cooking a meal—many participants reported no explicit behavior change as a result of study participation. As one participant stated, "I thought the survey was very helpful, but it didn't change me" [Participant 4, 10 years post-MBC diagnosis, completed 10 EMA surveys].

Theme 4: Responses to Data

Participants appreciated the opportunity to see their data at the end of the study. In reviewing their data, participants felt that they could more clearly see how it was going to be used to accomplish the goals of the study. As one participant noted:

> *I really appreciate the fact that you are sharing all the information with me, and the opportunity to know what happened at the end. So that's interesting for me because I never had that before, and I've been in a couple other studies.* [Participant 28, 3 years post-MBC diagnosis, completed 10 EMA surveys]

Many participants found their data interesting or surprising but individual reactions varied. Some participants had positive responses when presented with data that they characterized as "good," while others expressed sadness upon "seeing that [quality of life] was trending down" [Participant 16, 9 years post-MBC diagnosis, completed 12 EMA surveys]. While most participants felt that the graphs accurately reflected their weekly experiences, others commented on the inaccuracy of some data due to "user error" [Participant 1, 1 year post-MBC diagnosis, completed 10 EMA surveys]. This "user error" most frequently referred to inadvertently choosing the wrong response value on the slider (see theme 2).

Of particular note were participant reactions to their data on positive experiences, including joy, peace, and social connection. Some attributed low scores directly to their physical or mental health, or even felt that positive experiences were not possible for someone with cancer: "As my other friend with cancer and I have discussed, joy is not really much of a part of our experience. So I'm not surprised it never got over 50%" [Participant 15, 8 years post-MBC diagnosis, completed 12 EMA surveys]. However, many of the participants reported positive experiences *despite* difficulties with their physical or mental health:

> *[My] physical health has declined. I have neuropathy so bad. It's hard to hold things. My feet just dangle. I have no control over my feet. I don't even get out of my pajamas, or some days I don't even brush my teeth. But positive experiences, I do have some still. I still have my friends that call me and visit me sometimes.* [Participant 8, 8 years post-MBC diagnosis, completed 11 EMA surveys]

XSL•FO
RenderX

> *I have chronic neck pain. I have [cancerous] lesions top to bottom on every vertebra of my spine, and the worst being my cervical spine. And that's where I had the surgery last year. So it's healed. It's strong. I mean, I was almost paralyzed last year, and I'm nowhere near that. I have a normal life, except I do have chronic pain in my neck sometimes...[but] my family and my kids bring me so much peace and joy.* [Participant 13, 1 year post-MBC diagnosis, completed 12 EMA surveys]

Reactions to the idea of real-time data sharing were mixed. Interviewers asked participants how they would feel if their data were available in real time. Some participants felt that this would be beneficial:

> *I think it would have been helpful. If I could see them in real time, then I would be aware of the things that I could change and make an effort to change them. Obviously, some stuff you can't really change with pain and anxiety and stuff. That stuff is hard to change. But in terms of joy and peace, there are things you can do to try to change those.* [Participant 3, 2 years post-MBC diagnosis, completed 11 EMA surveys]

Others felt that real-time data sharing, especially in moments of high depression or anxiety, would not be helpful. Many participants shared that they "don't think [they] would have responded any differently" or changed their behaviors if they had received data in real time and preferred to "look at it all together in a month" to "get a better picture" [Participant 18, 3 years post-MBC diagnosis, completed 12 EMA surveys].

Theme 5: Recommendations for Future Interventions

Real-time interventions were of interest to most, but not all, patients with MBC with various suggestions for content and format. Participants were asked about their openness to real-time interventions in moments of high symptom burden. Many participants were open to receiving resources in real time. Suggested content for EMIs varied widely and included mindfulness, meditation, and breathing exercises; tips for emotion regulation; reminders to take medications for side effects such as pain, nausea, and GI distress; suggestions to engage in physical activity or spend time outside or in nature; affirmations and journaling; and prompts to connect with other people. Participants overall emphasized desire to see "practical application(s)" in real-time interventions to alleviate symptom burden [Participant 3, 2 years post-MBC diagnosis, completed 11 EMA surveys]. Some participants wanted real-time interventions that could "[be implemented] without the need for someone else" [Participant 3, 2 years post-MBC diagnosis, completed 11 EMA surveys]. These wishes for quick, pragmatic, and implementable interventions were consistent across participants who expressed interest in real-time interventions. In terms of the format of the intervention, participants overall expressed a preference for receiving resources via email or text message, preferably with a link to a website where they could access further information as needed. As one participant put it:

> *If you rank between this and this [on the survey], maybe they could send them to you just as a, "Hey, you're not alone. We appreciate your feedback. And while it may not totally resolve your matter, here are some suggestions." Right after you finish that survey for that time period or something because some people may not know what to do, just suggestions.* [Participant 7, 2 months post-MBC diagnosis, completed 10 EMA surveys]

Relatedly, one participant described how the frequency of real-time interventions could affect their adherence, sharing, "I like messages sporadically, something that draws my attention to it. It can't be too habitual" [Participant 24, 3 years post-MBC diagnosis, completed 12 EMA surveys]. Less commonly, some participants suggested that a responsive, but delayed, intervention would be preferable. As one participant shared:

> *That week when I was having a really tough time, I might not have been able to take another piece of advice that day...maybe the next day?* [Participant 25, 1 year post-MBC diagnosis, completed 11 EMA surveys]

Some participants saw potential benefits of EMIs for their friends and family. While most participants felt that resources should be delivered directly to the person living with MBC, a few noted that their friends and family might also benefit from resources. For example, one participant expressed a desire for "somebody for my kids and my husband to talk to" [Participant 7, 2 months post-MBC diagnosis, completed 10 EMA surveys]. Another noted that engaging friends and family members might help bolster the overall effect of the intervention, stating "if the family supports them, or friends accompany them, maybe the person will be a little more encouraged" [Participant 29, 10 years post-MBC diagnosis, completed 12 EMA surveys].

Many participants were interested in interventions that included engagement with other people with similar experiences. Participants expressed a desire to connect with other patients with MBC, particularly in moments of high distress: "If I'm really down in the dumps, I would not call my friends, though, because they don't understand. But I would call another cancer patient" [Participant 8, 8 years post-MBC diagnosis, completed 11 EMA surveys]. Some participants had been able to establish their own peer support networks but felt that providing support resources could help others. As one participant noted:

> *I think that there are people who don't know how to access those supports. So maybe [you could provide] the information for therapists or support groups. Having access to those resources, maybe places where you could access them for free.* [Participant 12, 5 years post-MBC diagnosis, completed 10 EMA surveys]

Thus, an EMI that connected patients with peers or provided resources for seeking additional support was widely seen as potentially beneficial.

Although less common, some participants were not interested in additional resources. A smaller group of participants, typically those further from MBC diagnosis, was not interested in receiving resources for managing symptoms. As one participant described, "There's nothing right now that I need

XSL•FO
RenderX

addressed. I'm already addressing it. Like I said, I've already been around the block a few times, so I've already come up with my own way if there is a way" [Participant 27, 5 years post-MBC diagnosis, completed 7 EMA surveys].

Discussion

Principal Findings

This study qualitatively explored patient experiences with EMA and their preferences for an EMI to ameliorate QOL. We identified 5 major themes from qualitative interviews: participant engagement, feedback on study design, impact of self-monitoring, responses to data, and recommendations for future interventions. These results highlight that patients with MBC are open to EMA methodologies for data collection and that EMIs can be valuable to this population.

Patients with MBC were engaged and interested in participating in this EMA study. Many patients were motivated by increasing research inclusivity for this population, underscoring the importance to participants of representativeness in research samples. These findings are consistent with other studies, which identified that altruism highly impacts patients' decisions to enroll in research studies [46,47]. Designing recruitment strategies and study methodology to emphasize inclusivity and the altruistic benefits of the research may facilitate recruitment and retention in research studies [48]. However, we also identified that several patients did not fully understand the study design and purpose of the research. This has important implications for the design and implementation of the informed consent process for future studies to ensure that all participants fully understand study procedures before enrolling. For example, future studies with complex EMA designs might incorporate evidence-based tools for improving informed consent, such as the "teach-back" method [49,50].

Furthermore, participant feedback indicated that a preferred study procedure would have included fewer surveys per day for multiple days in a row, with additional time to complete each survey. Some participants noted that their experiences did not change dramatically over the course of the day, which aligns with and builds upon our previously published quantitative analyses [31,32], demonstrating that time of day of the EMA was not significantly associated with QOL or symptom burden. Taken together, these results can help guide the design of future EMA studies that align with the preferences of the MBC community. Despite their suggestions for improving the EMA schedule, participants generally found the EMA design to be straightforward and easy to understand, similar to other studies using EMA in MBC populations. One such study looked specifically at the feasibility of using EMA for symptom burden and management and demonstrated high adherence and retention [51]. This study extends these findings to demonstrate the feasibility of using EMA to assess physical symptoms, psychological symptoms, positive experiences, and overall QOL among patients with MBC.

One unique result from this study was participant discomfort and confusion discussing libido as relevant to their MBC experience. This finding reflects a broader pattern reported in clinical settings, where the majority of women with breast cancer do not discuss the sexual health effects of cancer with their providers [52]. When conversations about sexual function do occur, they are more frequently initiated by the clinical provider [53]. Participants in this study were also confused about the definition of libido and its relevance as a side effect of many MBC treatments. More effective communication between patients and providers about sexual health is needed to increase patient knowledge [54]. Therefore, future studies should explain the rationale for assessing libido, define the term, and offer an opt-out response option to both increase participant knowledge and limit any potential participant discomfort.

Finally, these results have important implications for future interventions. While patients with MBC are open to EMIs with varying content and formats, more research is needed on the best time to intervene for each person. Intervention timing is an important avenue for future research to explore, as participants in this study varied on when they would have preferred help (ie, in moments of high distress vs the next day). Future research should also further examine patient preferences toward receiving help during moments of high distress to determine feasibility and acceptability of various options (eg, connecting with a peer vs receiving resources). We also observed variability in participant openness toward receiving their data in real time, with some participants feeling positively toward the idea and others negatively. Given this variability in preferences, a just-in-time adaptive intervention (JITAI) may be well suited for this population; JITAIs deliver personalized support in real time, typically using mobile technology, with the aim of providing the right type and amount of support at the right time and place to maximize effectiveness [55]. A JITAI for people with MBC could personalize intervention delivery based on participant and contextual factors, including participant preferences. In terms of the intervention components to be included in a JITAI framework, participants suggested a wide variety of resources, primarily targeted toward patients living with MBC but potentially also engaging family members and support people. These findings build upon previous research, as other studies also document the importance of more MBC-specific support groups to better meet this population's needs, as well as the lack of formal resources for family members and support people [56]. This study builds on this prior literature and further illuminates current gaps in interventions to improve QOL for patients with MBC.

Strengths and Limitations

This study is among the first to explore the acceptability of EMA study designs among patients with MBC, a group with significant QOL decrements and few options for tailored supportive care interventions. Our sample included participants who had been living with MBC for varying lengths of time and could therefore contribute unique perspectives on their disease. The qualitative interview design incorporated patients' EMA data, adding depth and clarity to study findings. We used a robust coding approach involving multiple coders for each transcript and discussion of discrepancies until consensus was reached, which allowed us to mitigate potential individual biases in the coding process.

XSL•FO
RenderX

However, the interpretation of these results must be informed by the study's limitations. First, interview participants represent a subset of individuals who were eligible for and agreed to participate in an EMA-based observational study of QOL. Thus, results may be subject to selection bias; participants who elected to participate in such a study may be more open to EMA and EMI or may have higher QOL than individuals who declined to participate. Second, responses may be subject to social desirability bias due to lack of anonymity during study interviews. Third, most participants were non-Hispanic White and highly educated. Although representative of patients at National Cancer Institute-designated Comprehensive Cancer Centers [57], it is possible that theoretical saturation was achieved only due to the homogenous nature of the cohort [58]. Relatedly, all participants in this study were patients at academic medical centers; their experience may not reflect the majority of patients with MBC who are seen in community oncology practices [59]. Additional research, including heterogeneous patient populations receiving care in varied settings, would increase the fidelity and usefulness of the findings presented here. Finally, this is a small, qualitative study. Future studies should quantitatively examine the feasibility and acceptability of EMI in a large, representative sample of patients with MBC.

Conclusions

This qualitative study identified that patients with MBC are willing to use EMA methodologies for data collection on QOL and that EMI could be beneficial for this population. Participants engaged well with the study design, found the EMA framework easy to understand, and would be open to EMIs. A JITAI framework may complement EMA well for this population, with the most frequently suggested interventions including MBC-specific support groups, peer navigation, and resources for family members and support people. Future research will take into consideration this qualitative participant feedback to formulate an updated EMA approach paired with a useful EMI.

Funding

This work was supported by the National Cancer Institute (R03 CA273489, PI: Conley; T32CA261787, PI: O'Neill, Adams-Campbell, Schwartz; P30 CA051008, PI: Weiner). The content presented here is solely the responsibility of the authors and does not represent the official views of the National Cancer Institute.

Data Availability

The datasets analyzed during this study are not publicly available, as qualitative interview data cannot be completely anonymized, even when aliases are used to replace names. Data are available from the corresponding author (CCC) on reasonable request.

Authors' Contributions

All authors contributed to the study conception and design. MAVS, JDR, and CCC performed material preparation and data collection. EPT, BM, MAVS, ZR, NK, MS, HK, and CCC performed data analysis. EPT, BM, MAVS, ZR, NK, MS, HK, and CCC wrote the first draft of the manuscript. All authors commented on subsequent versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

CCC has received research support (to institution) from Pfizer. LG has participated in consulting activities for which she has received honoraria from Pfizer and Stemline/Menarini. RG has participated in consulting activities for which she has received honoraria from AstraZeneca. CI has participated in consulting activities for which she has received honoraria from AstraZeneca, Genentech, Seattle Genetics, Pfizer, Gilead, and Merck; received royalties from Wolters Kluwer (UptoDate) and McGraw Hill (Goodman and Gillman); and received research support (to institution) from Tesaro/GSK, Seattle Genetics, Pfizer, AstraZeneca, Bristol Myers Squibb, Genentech, and Novartis. SCO has received research support (to institution) from Pfizer and Gilead Sciences. No other authors have competing interests to disclose.

Multimedia Appendix 1
Ecological momentary assessment surveys.
[PDF File, 117 KB - cancer_v12i1e80467_app1.pdf]

Multimedia Appendix 2
Qualitative interview guide.
[PDF File, 156 KB - cancer_v12i1e80467_app2.pdf]

Checklist 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File, 137 KB - cancer_v12i1e80467_app3.pdf]

References

XSL•FO
RenderX

1. Liang Y, Zhang H, Song X, Yang Q. Metastatic heterogeneity of breast cancer: molecular mechanism and potential therapeutic targets. Semin Cancer Biol 2020 Feb;60:14-27. [doi: 10.1016/j.semcancer.2019.08.012] [Medline: 31421262]
2. NIH Surveillance Research Program. SEER*Explorer: an interactive website for SEER cancer statistics. National Cancer Institute. 2023. URL: https://seer.cancer.gov/explorer [accessed 2025-01-03]
3. Malmgren JA, Mayer M, Atwood MK, Kaplan HG. Differential presentation and survival of de novo and recurrent metastatic breast cancer over time: 1990-2010. Breast Cancer Res Treat 2018 Jan;167(2):579-590. [doi: 10.1007/s10549-017-4529-5] [Medline: 29039120]
4. Courtney D, Davey MG, Moloney BM, et al. Breast cancer recurrence: factors impacting occurrence and survival. Ir J Med Sci 2022 Dec;191(6):2501-2510. [doi: 10.1007/s11845-022-02926-x] [Medline: 35076871]
5. Finn RS, Martin M, Rugo HS, et al. Palbociclib and letrozole in advanced breast cancer. N Engl J Med 2016 Nov 17;375(20):1925-1936. [doi: 10.1056/NEJMoa1607303] [Medline: 27959613]
6. Arnedos M, Vicier C, Loi S, et al. Precision medicine for metastatic breast cancer--limitations and solutions. Nat Rev Clin Oncol 2015 Dec;12(12):693-704. [doi: 10.1038/nrclinonc.2015.123] [Medline: 26196250]
7. Caswell-Jin JL, Plevritis SK, Tian L, et al. Change in survival in metastatic breast cancer with treatment advances: meta-analysis and systematic review. JNCI Cancer Spectr 2018 Nov;2(4):ky062. [doi: 10.1093/jncics/pky062] [Medline: 30627694]
8. Santa-Maria CA, Gradishar WJ. Changing treatment paradigms in metastatic breast cancer: lessons learned. JAMA Oncol 2015 Jul;1(4):528-534. [doi: 10.1001/jamaoncol.2015.1198] [Medline: 26181262]
9. Sundquist M, Brudin L, Tejler G. Improved survival in metastatic breast cancer 1985-2016. Breast 2017 Feb;31(46-50):46-50. [doi: 10.1016/j.breast.2016.10.005] [Medline: 27810699]
10. Breast cancer. National Comprehensive Cancer Network (NCCN). 2025. URL: https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf [accessed 2025-04-09]
11. Gogate A, Wheeler SB, Reeder-Hayes KE, et al. Projecting the prevalence and costs of metastatic breast cancer from 2015 through 2030. JNCI Cancer Spectr 2021 Aug;5(4):pkab063. [doi: 10.1093/jncics/pkab063] [Medline: 34409255]
12. Mariotto AB, Etzioni R, Hurlbert M, Penberthy L, Mayer M. Estimation of the number of women living with metastatic breast cancer in the United States. Cancer Epidemiol Biomarkers Prev 2017 Jun;26(6):809-815. [doi: 10.1158/1055-9965.EPI-16-0889] [Medline: 28522448]
13. Lin C, Clark R, Tu P, Bosworth HB, Zullig LL. Breast cancer oral anti-cancer medication adherence: a systematic review of psychosocial motivators and barriers. Breast Cancer Res Treat 2017 Sep;165(2):247-260. [doi: 10.1007/s10549-017-4317-2] [Medline: 28573448]
14. Butler LD, Koopman C, Classen C, Spiegel D. Traumatic stress, life events, and emotional support in women with metastatic breast cancer: cancer-related traumatic stress symptoms associated with past and current stressors. Health Psychol 1999 Nov;18(6):555-560. [doi: 10.1037//0278-6133.18.6.555] [Medline: 10619528]
15. Hart NH, Nekhlyudov L, Smith TJ, et al. Survivorship care for people affected by advanced or metastatic cancer: MASCC-ASCO standards and practice recommendations. JCO Oncol Pract 2024 Sep;20(9):1160-1172. [doi: 10.1200/OP.23.00716] [Medline: 38684036]
16. Ginter AC. “The day you lose your hope is the day you start to die”: quality of life measured by young women with metastatic breast cancer. J Psychosoc Oncol 2020;38(4):418-434. [doi: 10.1080/07347332.2020.1715523] [Medline: 32067600]
17. Conley CC, Cumbo S, Chavez Ochoa J, et al. Iterative development of an interactive website to support shared decision-making in metastatic breast cancer. J Cancer Educ 2024 Dec;39(6):681-690. [doi: 10.1007/s13187-024-02451-8] [Medline: 38730103]
18. Kawahara T, Iwamoto T, Takashima I, et al. Association of change in health-related quality of life and treatment discontinuation in metastatic breast cancer: a post hoc, exploratory analysis of two randomized clinical trials. Support Care Cancer 2022 Oct;30(10):8367-8375. [doi: 10.1007/s00520-022-07283-0] [Medline: 35857127]
19. Keane D, Phillips G, Mitchell N, Connolly RM, Hegarty J. Improving quality of life and symptom experience in patients with metastatic breast cancer: a systematic review of supportive care interventions. Psychooncology 2023 Aug;32(8):1192-1207. [doi: 10.1002/pon.6183] [Medline: 37434307]
20. Hiensch AE, Depenbusch J, Schmidt ME, et al. Supervised, structured and individualized exercise in metastatic breast cancer: a randomized controlled trial. Nat Med 2024 Oct;30(10):2957-2966. [doi: 10.1038/s41591-024-03143-y] [Medline: 39054374]
21. Zhang Y, Pang Y, He Y, You M, Tang L. Feasibility of online managing cancer and living meaningfully (CALM) in Chinese patients with metastatic breast cancer: a pilot randomized control trial. Sci Rep 2024 Feb 28;14(1):4892. [doi: 10.1038/s41598-024-52574-7] [Medline: 38418478]
22. Campbell TM, Campbell EK, Culakova E, et al. A whole-food, plant-based randomized controlled trial in metastatic breast cancer: weight, cardiometabolic, and hormonal outcomes. Breast Cancer Res Treat 2024 Jun;205(2):257-266. [doi: 10.1007/s10549-024-07266-1] [Medline: 38446316]
23. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]

XSL•FO
RenderX

24. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39. [doi: 10.1348/135910709X466063] [Medline: 19646331]
25. Dao KP, De Cocker K, Tong HL, Kocaballi AB, Chow C, Laranjo L. Smartphone-delivered ecological momentary interventions based on ecological momentary assessments to promote health behaviors: systematic review and adapted checklist for reporting ecological momentary assessment and intervention studies. JMIR Mhealth Uhealth 2021 Nov 19;9(11):e22890. [doi: 10.2196/22890] [Medline: 34806995]
26. O'Connor BP, Pesut B. A longitudinal study of within-person trajectories in quality of life in patients receiving early palliative care. Qual Life Res 2024 Oct;33(10):2733-2742. [doi: 10.1007/s11136-024-03722-z] [Medline: 38907833]
27. Guo L, Kong D, Liu J, et al. Breast cancer heterogeneity and its implication in personalized precision therapy. Exp Hematol Oncol 2023 Jan 9;12(1):3. [doi: 10.1186/s40164-022-00363-1] [Medline: 36624542]
28. Pasha N, Turner NC. Understanding and overcoming tumor heterogeneity in metastatic breast cancer treatment. Nat Cancer 2021 Jul;2(7):680-692. [doi: 10.1038/s43018-021-00229-1] [Medline: 35121946]
29. Kida K, Olver I, Yennu S, Tripathy D, Ueno NT. Optimal supportive care for patients with metastatic breast cancer according to their disease progression phase. JCO Oncol Pract 2021 Apr;17(4):177-183. [doi: 10.1200/OP.20.00622] [Medline: 33492987]
30. Thong MSY, Chan RJ, van den Hurk C, et al. Going beyond (electronic) patient-reported outcomes: harnessing the benefits of smart technology and ecological momentary assessment in cancer survivorship research. Support Care Cancer 2021 Jan;29(1):7-10. [doi: 10.1007/s00520-020-05648-x] [Medline: 32844316]
31. Mossman B, Small BJ, Velazquez-Sosa MA, Rodriguez JD, Bobbit A, Mathurin A, et al. Ecological momentary assessment of quality of life in women with metastatic breast cancer. 2025 Presented at: American Preventive Oncology Society Annual Meeting URL: https://aspo.org/2025abstracts/ [accessed 2026-02-20]
32. Mossman B, Small BJ, Addington EL, et al. Longitudinal quality of life in adults with metastatic breast cancer: the impact of between- and within- person changes in symptoms and positive affect. Res Sq. Preprint posted online on Nov 19, 2025. [doi: 10.21203/rs.3.rs-8138106/v1] [Medline: 41333396]
33. Lizán L, Pérez-Carbonell L, Comellas M. Additional value of patient-reported symptom monitoring in cancer care: a systematic review of the literature. Cancers (Basel) 2021 Sep 15;13(18):4615. [doi: 10.3390/cancers13184615] [Medline: 34572842]
34. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565. [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
35. Warrington L, Absolom K, Conner M, et al. Electronic systems for patients to report and manage side effects of cancer treatment: systematic review. J Med Internet Res 2019 Jan 24;21(1):e10875. [doi: 10.2196/10875] [Medline: 30679145]
36. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
37. Heron KE, Everhart RS, McHale SM, Smyth JM. Using mobile-technology-based ecological momentary assessment (EMA) methods with youth: a systematic review and recommendations. J Pediatr Psychol 2017 Nov 1;42(10):1087-1107. [doi: 10.1093/jpepsy/jsx078] [Medline: 28475765]
38. Doherty K, Balaskas A, Doherty G. The design of ecological momentary assessment technologies. Interact Comput 2020 May 17;32(3):257-278. [doi: 10.1093/iwcomp/iwaa019]
39. Fram SM. The constant comparative analysis method outside of grounded theory. TQR 2013;18:1. [doi: 10.46743/2160-3715/2013.1569]
40. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
41. Azungah T. Qualitative research: deductive and inductive approaches to data analysis. QRJ 2018 Nov 15;18(4):383-400. [doi: 10.1108/QRJ-D-18-00035]
42. Lindlof T, Taylor B. Qualitative Communication Research Methods, 4th edition: Thousand Oaks; 2017.
43. DeCuir-Gunby JT, Marshall PL, McCulloch AW. Developing and using a codebook for the analysis of interview data: an example from a professional development research Project. Field Methods 2011 May;23(2):136-155. [doi: 10.1177/1525822X10388468]
44. Miles MB, Huberman AM, Saldaña J. Qualitative Data Analysis: A Methods Sourcebook: Thousand Oaks; 2019.
45. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug;42(4):1758-1772. [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]
46. Sheridan R, Martin-Kerry J, Hudson J, Parker A, Bower P, Knapp P. Why do patients take part in research? An overview of systematic reviews of psychosocial barriers and facilitators. Trials 2020 Mar 12;21(1):259. [doi: 10.1186/s13063-020-4197-3] [Medline: 32164790]
47. Truong TH, Weeks JC, Cook EF, Joffe S. Altruism among participants in cancer clinical trials. Clin Trials 2011 Oct;8(5):616-623. [doi: 10.1177/1740774511414444] [Medline: 21813584]

XSL•FO
RenderX

48. Williams B, Entwistle V, Haddow G, Wells M. Promoting research participation: why not advertise altruism? Soc Sci Med 2008 Apr;66(7):1451-1456. [doi: 10.1016/j.socscimed.2007.12.013] [Medline: 18222579]
49. Seely KD, Higgs JA, Nigh A. Utilizing the "teach-back" method to improve surgical informed consent and shared decision-making: a review. Patient Saf Surg 2022 Mar 5;16(1):12. [doi: 10.1186/s13037-022-00322-z] [Medline: 35248126]
50. Seely KD, Higgs JA, Butts L, et al. The "teach-back" method improves surgical informed consent and shared decision-making: a proof of concept study. Patient Saf Surg 2022 Oct 28;16(1):33. [doi: 10.1186/s13037-022-00342-9] [Medline: 36307856]
51. Brasky TM, Newton AM, Stephens JA, et al. Testing the feasibility of mobile ecological momentary assessment for symptom burden and management among metastatic cancer patients. Digit Health 2024;10:20552076241261843. [doi: 10.1177/20552076241261843] [Medline: 38854924]
52. Reese JB, Sorice K, Lepore SJ, Daly MB, Tulsky JA, Beach MC. Patient-clinician communication about sexual health in breast cancer: a mixed-methods analysis of clinic dialogue. Patient Educ Couns 2019 Mar;102(3):436-442. [doi: 10.1016/j.pec.2018.10.003] [Medline: 30314828]
53. Reese JB, Sorice KA, Zimmaro LA, Lepore SJ, Beach MC. Communication about sexual health in breast cancer: what can we learn from patients' self-report and clinic dialogue? Patient Educ Couns 2020 Sep;103(9):1821-1829. [doi: 10.1016/j.pec.2020.03.029] [Medline: 32349896]
54. Reese JB, Beach MC, Smith KC, et al. Effective patient-provider communication about sexual concerns in breast cancer: a qualitative study. Support Care Cancer 2017 Oct;25(10):3199-3207. [doi: 10.1007/s00520-017-3729-1] [Medline: 28451911]
55. Goldstein SP, Evans BC, Flack D, et al. Return of the JITAI: applying a just-in-time adaptive intervention framework to the development of m-Health solutions for addictive behaviors. Int J Behav Med 2017 Oct;24(5):673-682. [doi: 10.1007/s12529-016-9627-y] [Medline: 28083725]
56. Lyons-Rahilly T, Meskell P, Carey E, Meade E, O' Sullivan D, Coffey A. Exploring the experiences of women living with metastatic breast cancer [MBC]: a systematic review of qualitative evidence. PLoS One 2024;19(1):e0296384. [doi: 10.1371/journal.pone.0296384] [Medline: 38181009]
57. Wolfson JA, Sun CL, Wyatt LP, Hurria A, Bhatia S. Impact of care at comprehensive cancer centers on outcome: results from a population-based study. Cancer 2015 Nov 1;121(21):3885-3893. [doi: 10.1002/cncr.29576] [Medline: 26218755]
58. Levitt HM. Qualitative generalization, not to the population but to the phenomenon: reconceptualizing variation in qualitative research. Qualitative Psychology 2021;8(1):95-110. [doi: 10.1037/qup0000184]
59. Kirkwood MK, Hanley A, Bruinooge SS, et al. The state of oncology practice in America, 2018: results of the ASCO Practice Census Survey. J Oncol Pract 2018 Jul;14(7):e412-e420. [doi: 10.1200/JOP.18.00149] [Medline: 29906211]

Abbreviations

EMA: ecological momentary assessment
EMI: ecological momentary intervention
GI: gastrointestinal
JITAI: just-in-time adaptive intervention
MBC: metastatic breast cancer
QOL: quality of life
REDCap: Research Electronic Data Capture

Edited by M Balcarras, N Cahill; submitted 10.Jul.2025; peer-reviewed by E Kemp, I Mindlis; accepted 26.Jan.2026; published 03.Mar.2026.

Please cite as:
Taylor EP, Mossman B, Velazquez-Sosa MA, Rashid Z, Kloster N, Schwarz M, Kang H, Rodriguez JD, Addington EL, Moskowitz JT, Glenn L, Drakeford S, Guerra R, Isaacs C, Chitalia A, Gallagher C, Graham D, O'Neill SC, Conley CC
Evaluation of Experiences With Ecological Momentary Assessment Among Women With Metastatic Breast Cancer: Qualitative Study
JMIR Cancer 2026;12:e80467
URL: https://cancer.jmir.org/2026/1/e80467
doi:10.2196/80467

© Emily P Taylor, Brenna Mossman, Mikaela A Velazquez-Sosa, Zaynab Rashid, Natalie Kloster, Maureen Schwarz, Hannah Kang, Jennifer D Rodriguez, Elizabeth L Addington, Judith Tedlie Moskowitz, Lesley Glenn, Shontè Drakeford, Roxana Guerra, Claudine Isaacs, Ami Chitalia, Christopher Gallagher, Deena Graham, Suzanne C O'Neill, Claire C Conley. Originally published in JMIR Cancer (https://cancer.jmir.org), 3.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Metastasis Extraction From Non–Small Cell Lung Cancer Clinical Notes Using Large Language Model–Based Classification: Retrospective Comparative Study

Sweta Balaji[1], BS; Kiersten Campbell[1], MS; Rou-Zhen Chen[2], MPH; Daniel Smith[3], MS; Matthew A Reyna[1], PhD; Abeed Sarker[1], PhD; Joshua D Wallach[2], PhD; Ravi B Parikh[4], MD; Selen Bozkurt[1], PhD

[1]Department of Biomedical Informatics, School of Medicine, Emory University, Atlanta, GA, United States

[2]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[3]Department of IT and Research Informatics, Winship Cancer Institute, Atlanta, GA, United States

[4]Department of Hematology and Medical Oncology, School of Medicine, Emory University, Atlanta, GA, United States

Corresponding Author:
Selen Bozkurt, PhD
Department of Biomedical Informatics
School of Medicine
Emory University
100 Woodruff Circle
Atlanta, GA, 30322
United States
Phone: 1 (404) 727 4631
Email: selen.bozkurt@emory.edu

Abstract

Background: Identification of metastasis status in non–small cell lung cancer (NSCLC) is a critical part of understanding disease prognosis, treatment courses, trial eligibility, and population-level cancer surveillance. However, metastasis status is inconsistently recorded in structured cancer registry fields, since manual abstraction of clinical notes is often an error-prone process. This challenge highlights an opportunity for leveraging large language models (LLMs) to conduct high-scale metastasis extraction from real-world clinical documentation.

Objective: We conducted a retrospective, multicohort comparative evaluation of 3 distinct LLMs for 2 independent classification tasks: overall metastasis presence at any site and brain/central nervous system (CNS) metastasis presence. We evaluated model performance on 2 independent NSCLC cohorts: one with registry labels and another with manual annotations. We further explored whether our methods could analyze clinical documentation and recover missing information in structured registry fields.

Methods: Patient cohorts were derived from the Winship Cancer Institute. Cohort 1 (n=579 patients; 24,887 notes across 69 note types; 2023-2025) used registry-linked metastasis fields as the reference standard. Cohort 2 (n=22 patients; 644 radiology notes; 2010-2021) was drawn from 2 completed randomized trials and used dual-annotated labels (Cohen κ: 0.93 overall metastasis; 0.88 CNS metastasis) as the reference standard. We fine-tuned the GatorTron-base encoder model for each independent binary classification task. We evaluated MedGemma-27B-text and Llama 3.1-70B using zero-shot prompting. A separate cohort of 675 patients with missing registry labels was used for an exploratory missingness-recovery analysis, validated against annotations of a random subsample. Our best-performing model was used to conduct further site-specific metastasis classification, based on available registry data.

Results: More than half (675/1260, 54%) of initially identified cohort 1 patients had missing or unknown registry metastasis labels. MedGemma demonstrated the best performance in overall metastasis classification (cohort 1: F_1-score=0.80; cohort 2 patient-level: F_1-score=1.0; cohort 2 note-level: F_1-score=0.93). For brain/CNS metastasis, Llama3 performed best in both cohorts (cohort 1: F_1-score=0.79; cohort 2 patient-level: F_1-score=0.93; cohort 2 note-level: F_1-score=0.86). The fine-tuned GatorTron model showed strong performance for classification of overall metastasis in cohort 1 (F_1-score=0.72). Error analysis indicated that several model errors reflected incomplete registry labels, ambiguous clinical language, or missing documentation. In the exploratory recovery analysis, model predictions agreed with manual annotations at an accuracy of 0.90 and F_1-score of 0.89.

XSL•FO
RenderX

Additional site-specific metastasis classification by MedGemma demonstrated high performance for liver metastasis (F_1-score=0.82) and bone metastasis (F_1-score=0.79).

Conclusions: All models demonstrated relatively high performance. The zero-shot generative models were more robust to nuanced documentation and context-dependent brain/CNS metastasis extraction. The fine-tuned encoder model demonstrated strong classification performance but may have been limited by potential inaccuracies in the registry reference standards during model training. This study further demonstrated the potential of LLMs in recovering plausible structured labels from narrative text, complementing cancer registries for metastasis ascertainment.

(JMIR Cancer 2026;12:e99573) doi:10.2196/99573

KEYWORDS

electronic health records; large language models; metastasis; natural language processing; non-small-cell lung cancer

Introduction

Non–small cell lung cancer (NSCLC) is the leading cause of cancer-related mortality worldwide, with prognosis and treatment shaped by the presence and extent of metastatic disease at diagnosis and throughout the clinical course [1-3]. Among the most consequential and clinically complex sites of spread is the central nervous system (CNS) [4]. Brain metastasis develops in up to 50% of patients with NSCLC over the course of their disease, carrying distinct prognostic implications, driving treatment decisions, and serving as a key eligibility criterion in oncology trials [5,6]. Accurate ascertainment of metastatic status is crucial for determining eligibility for systemic therapies, radiation planning, enrollment in clinical trials, and cancer surveillance [7,8]. These realities underscore that metastasis information, particularly CNS metastasis, must be captured accurately and comprehensively.

However, metastatic status is inconsistently recorded in structured electronic health record (EHR) fields and is often dispersed across radiology reports, oncology progress notes, discharge documentation, and other free-text clinical narratives [9]. Cancer registries, the primary infrastructure for population-level cancer surveillance, rely on manual review by trained cancer registrars, an approach that is resource-intensive, subject to human error, and limited in its ability to capture disease progression documented after the initial abstraction window [10]. These gaps in data capture can have meaningful consequences. In a large National Cancer Database analysis of more than 1.2 million patients with NSCLC, 71% had missing data in at least one variable of interest, including cancer stage and treatment [11]. Missing data were especially prevalent among patients with lower 2-year overall survival, indicating that incomplete documentation may have been more prevalent for patients with severe illnesses or complex, metastatic cases [11]. Thus, studies using cancer registry data may be biased if missingness is ignored [11].

Another limitation of cancer registries is that data often reflect disease status primarily at the time of initial diagnosis and do not capture metastatic progression documented later in the clinical course [12,13]. In contrast, clinical notes provide longitudinally detailed records of disease evolution, including imaging interpretation, treatment modifications, and clinician assessments across the patient journey. Manual abstraction of clinical notes is a lengthy, resource-intensive process that is prone to human error and may not comprehensively capture disease complexities documented throughout the course of a patient's medical history [14]. These challenges highlight an opportunity for informatics approaches, namely natural language processing (NLP) methods, to enable large-scale extraction of metastasis records from real-world clinical documentation, improving accuracy and comprehensiveness of existing structured data sources [15].

There has been rapid evolution of NLP methods in the field of clinical information extraction. Early rule-based methods leveraged medical ontologies to create structured rules for detection of relevant information in unstructured text [16]. This method continues to be used today, particularly for text filtering and standardization [17]. Rule-based methods have previously demonstrated high precision and interpretability but with limited ease of implementation and generalizability to notes beyond the study context [18,19]. Their limited generalizability is characterized by their dependence on specific linguistic patterns, along with requiring specialized human expertise to create custom rules [20]. These rules can be affected by the unique documentation style and phrasing of the data used [20]. Machine learning (ML) approaches demonstrated improvement over rule-based systems due to their ability to automatically learn linguistic patterns and complexities rather than solely capturing patterns based on predefined rules [21]. Transformer-based ML models, specifically, have demonstrated strong performance in real-world data and can be fine-tuned for use in various clinical contexts [21]. Transformers pretrained on clinical and medical data have further demonstrated immense promise in detecting semantic interactions and complex, varied phrasing in clinical text [22,23]. These models, further optimized on smaller clinical cohorts, have been shown to outperform traditional NLP and ML approaches, including regular expression-based text searches and classic supervised learning approaches [24]. However, lower performance has been observed when applying models, including transformer-based, rule-based, and classic ML methods, to heterogeneous note types or switching institutions from which the notes are sourced [21,23]. Furthermore, differences in vocabulary and documentation style can be seen to lower performance in these models [21,23]. More recently, generative large language models (LLMs), including both domain-adapted and general-purpose instruction-tuned variants, have demonstrated strong performance on a range of clinical extraction tasks [11,25-27]. However, because generative LLMs produce free, unconstrained text rather than directly predicting

XSL•FO
RenderX

predefined labels, additional prompting and/or postprocessing may be required to standardize their outputs [28]. These additional steps may introduce variability in output formatting and interpretation, whereas BERT-based classifiers map inputs directly to a fixed label set [28]. These models also require a substantial amount of memory, which can detract from widespread use in clinical settings [23].

Despite this progress, critical gaps remain, particularly in the types of clinical text analyzed for extraction tasks. First, a recent systematic review of LLM applications in clinical medicine found that over three-quarters of included studies did not conduct analyses on real EHR data, relying instead on simulated clinical scenarios or synthetic notes [29]. This limits the generalizability of findings to the nuanced documentation encountered in routine oncology practice. Second, few studies have directly compared fine-tuned domain-adapted encoder models, for supervised classification, against instruction-tuned generative models, particularly for information extraction from real-world oncology notes [23,28]. Third, and particularly for NSCLC, brain and CNS metastasis extraction is underexamined as a distinct NLP task, despite its clinical importance. Given the complex language and the plethora of terminologies used to document different sites of CNS metastasis, this extraction task further tests the ability of NLP methods or LLMs in detecting and distinguishing between nuanced phrases and site-specific verbiage.

To address these gaps, we examined 2 research questions. First, how accurately can LLM-based methods classify the presence of overall metastasis and site of metastasis from real-world clinical notes? Second, can these models recover clinically relevant metastasis information that is missing from structured registry data? Metastasis status recorded in structured registry fields is a crucial part of observational cancer research and trial recruitment, so systematic undercapture propagates into every downstream use. A pipeline addressing this gap would operate retrospectively on documentation already in the EHR. In cancer registry abstraction, where metastasis status is currently determined by manual chart review, it would prescreen notes and help identify candidate documentation for abstractors. In trial eligibility screening, it would flag patients whose metastasis status is absent or incomplete in structured fields and who may be misclassified during cohort identification. In both cases, the model narrows what a human reviews rather than deciding the classification.

Methods

Study Design and Setting

We conducted a retrospective comparative evaluation of 3 LLM-based methods for extracting (1) overall metastasis at any site and (2) brain/CNS metastasis from real-world clinical notes in patients with NSCLC. We leveraged 2 cohorts with distinct but complementary roles in model development and validation, in addition to a dataset with missing metastasis records to explore the LLM's capabilities for data recovery. All analyses were conducted under a waived-consent protocol approved by the Emory University Institutional Review Board.

Data Sources and Cohorts

Cohort 1: Primary Registry-Linked Development Cohort

Overview

Cohort 1 was used for model development and internal validation. It was derived from the Winship Cancer Registry, which is linked to the Georgia Comprehensive Cancer Registry. Inclusion was restricted to patients for whom Winship Cancer Institute was the institution of record for diagnosis and/or first-course treatment, with both registry data and linked clinical notes available from 2023 through 2025.

Cohort Construction and Exclusions

Initial data extraction yielded 1260 patients with NSCLC with 62,191 notes across 81 note types and 12,174 encounters. Patients were excluded from model development if at least one registry metastasis field was missing (n=663) or labeled unknown with no confirmed positive site (n=12), yielding 585 patients with 24,908 notes across 69 note types. Six additional patients were excluded from model development and validation because their available notes contained no cancer-relevant clinical content (eg, administrative telephone encounters only), yielding a final development cohort of 579 patients with 24,887 notes across 69 note types for overall metastasis analysis. Three additional patients with unconfirmed brain metastasis registry records were excluded from the brain/CNS metastasis analysis, yielding a cohort of 576 patients. The full cohort selection process is illustrated in Figure 1. Demographic characteristics, following initial data extraction and cohort selection, are described in Table 1.

XSL•FO
RenderX

Figure 1. Flowchart of selection of cohort 1 following initial data extraction.

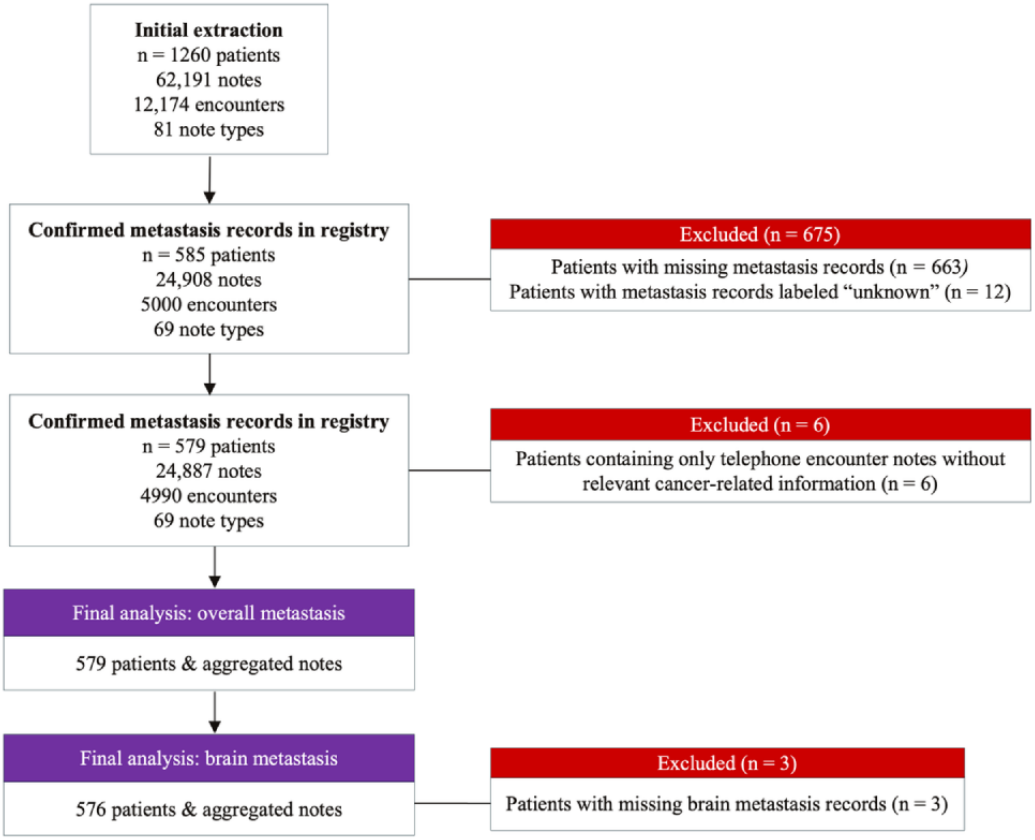


Table 1. Cohort 1 demographic characteristics.

	Total patients identified (n=1260)	Patients included in analysis (n=579)
Age at diagnosis (years), median (IQR)	72 (65-79)	73 (66-79)
Male, n (%)	670 (53.2)	273 (47.1)
Race, n (%)		
White	599 (47.5)	279 (48.2)
Black or African American	516 (40.9)	220 (38)
Asian	72 (5.7)	37 (6.4)
Other	20 (1.7)	13 (2.2)
Unknown	53 (4.2)	30 (5.2)
Ethnicity, n (%)		
Hispanic	27 (2.1)	14 (2.4)
Non-Hispanic	1147 (91)	521 (90)
Unknown	86 (6.8)	44 (7.6)

Reference Standard

We used the Winship Cancer Registry as our reference standard, which reports site-specific metastasis (liver, brain, bone, lung, distant lymph nodes, and other distant sites). Using these fields, we created our binary outcome label to identify overall metastasis. This label, encompassing all metastasis sites, was derived as follows: patients with at least one site-specific positive metastasis record were classified as metastasis-positive; patients with all-negative records were classified as metastasis-negative, and patients with at least one missing (n=663) or unknown (n=12) field, in the absence of a confirmed positive, were classified as metastasis unknown and excluded from model development. The registry brain metastasis field was used directly as the reference standard for brain metastasis analysis.

Note Inclusion Window

Because registry metastasis fields primarily capture disease status at initial diagnosis, clinical note inclusion was restricted to a window spanning 30 days before through 30 days after each patient's formal cancer diagnosis date. This window maximized

XSL•FO
RenderX

temporal correspondence between notes and registry labels. All note types within this window were eligible, including radiology reports, discharge summaries, progress notes, consultation notes, provider communication, nursing documentation, and other note types. The full distribution of notes across all note types is detailed in Table S1 in Multimedia Appendix 1.

Cohort 2: Independent Temporal Validation Cohort

To assess model generalizability across an earlier documentation period, we assembled an independent cohort of 22 patients with NSCLC whose records were extracted retrospectively from 2 completed randomized clinical trials conducted between 2010 and 2021. Trial enrollment required systematic imaging, resulting in dense radiology documentation in the historical records of these patients. A total of 644 radiology notes, including computed tomography, magnetic resonance imaging (MRI), X-ray, and other imaging reports, were extracted. Notes were not restricted to a diagnosis-adjacent window, and registry records were not used as the reference standard, since they did not capture documented metastatic progression. Manual note-level annotations served as the reference standard labels for this cohort (see ”Manual Annotation of Cohort 2” section). All models were applied to this cohort without retraining to serve as further external validation. Characteristics of both cohorts are provided in Table 2.

Table 2. Overall cohort characteristics.

	Cohort 1: primary registry-linked	Cohort 2: temporal validation
Reference standard	Cancer registry	Manual annotation
Documentation period	2023-2025	2010-2021
Total patients with NSCLC[a] identified, n	1260	22
Patients with missing/unknown registry metastasis records, n (%)	675 (54)	0 (0)
Patients included in analysis, n		
Overall	579	22
Brain/CNS[b]	576	22
Clinical notes included in analysis, n	24,887	644
Note types	Heterogeneous (69 types)	Radiology only
Patients with overall metastasis, n (%)	239 (41)	22 (100)
Patients with brain/CNS metastasis, n (%)	89 (15)	8 (36)

[a]NSCLC: non–small cell lung cancer.

[b]CNS: central nervous system.

Missingness-Recovery Dataset

The 675 patients excluded from primary model development, due to at least one missing or unknown registry metastasis record, represented 54% of the initial cohort 1 extraction pool. A total of 12 of these patients had unknown brain metastasis records. Notes from these patients were retained for a downstream exploratory analysis. This dataset was used solely to assess whether our methods could recover structured metastasis information absent from registry fields. No reference standard was available for this group, and model predictions were qualitatively reviewed by selecting a small random sample of patients and manually assessing their clinical notes, which were aggregated to the patient level. (see “Exploratory Downstream Missingness-Recovery Analysis” section). This analysis is reported to describe the feasibility of our methods and not as a primary performance evaluation.

Clinical Note Preprocessing

Prior to model input, all clinical notes were preprocessed through a keyword-based filtering pipeline to reduce noise while retaining clinically relevant content. Keywords were drawn from the publicly available Clinical Event Recognizer (CLEVER) dictionary, which contains clinical synonyms for oncology-relevant terms, supplemented by custom oncology-specific terms developed by the study team [30]. Representative CLEVER terms included “metastasis,” “widespread disease,” and “stage IV,” among other words and phrases. Custom additions included “high grade” and “malignant mass,” in addition to other words. All keywords used for preprocessing are available in Table S2 in Multimedia Appendix 1. For notes containing at least one keyword match, only sentences containing keywords were retained. For notes without any match, the full note text was preserved. This step was applied strictly as a noise-reduction measure to eliminate sentences irrelevant to the patient’s disease course. In cohort 1, all notes from each patient were aggregated at the patient level before preprocessing and model input, consistent with the patient-level registry reference standard. In cohort 2, notes were processed and input individually.

Manual Annotation of Cohort 2

Two trained abstractors, both of whom have a background in public health and biomedical informatics, independently annotated radiology notes from cohort 2 under the supervision of a board-certified oncologist. A standardized schema targeted 2 binary note-level labels: (1) overall metastasis present vs absent and (2) CNS metastasis present vs absent. Annotators

XSL•FO
RenderX

were instructed to label both explicit evidence (eg, "brain metastases confirmed on MRI") and contextually implied evidence strongly suggestive of current metastatic disease, such as imaging findings or treatment references indicating metastatic involvement.

Annotators were provided with two forms of each note: (1) the preprocessed text on which the label was assigned, and (2) the complete original note available for additional context. A predefined annotation schema was used to assign two binary note-level labels: (1) overall metastasis present vs absent, and (2) CNS metastasis present vs absent (Supplemental Material 1 in Multimedia Appendix 1). Disagreements were resolved by discussion or oncologist adjudication. Annotation proceeded in 4 iterative rounds of increasing size (n=10, 100, 200, and 334 notes), with interannotator agreement assessed and discrepancies reconciled after each round. Final agreement, quantified by Cohen κ values across all rounds of annotation, was 0.93 for overall metastasis and 0.88 for CNS metastasis, indicating high agreement (Table 3). Final labels were assigned based on unanimous consensus or oncologist suggestion.

Annotators reviewed each note in full before assigning a label. Explicit evidence of metastasis presence was often contained within a single sentence such as "hepatic lesions consistent with metastatic spread," "brain metastases confirmed on MRI," or "malignant pleural effusion," and one such statement was sufficient for a positive label. A metastasis negative label could be determined either by explicit evidence of a metastasis-negative test result or by absence of metastasis-positive language in the note. Metastasis negative labels therefore required review of the complete note. Annotators were also instructed to label positive imaging findings or treatment references indicating metastatic involvement as metastasis positive. Findings in which the documenting clinician expressed uncertainty, such as a lesion described as suspicious or worrisome for metastasis without confirmation, were labeled negative. Because evidence of metastasis was distributed across multiple phrases within a note, and because the negative class required complete review of the note to establish absence of metastasis-positive language, annotators did not mark individual text spans and instead assigned note-level labels based on the full note.

Table 3. Interannotator agreement for cohort 2.

Annotation round	Overall metastasis, Cohen κ	CNS[a] metastasis, Cohen κ
Round 1 (n=10)	1	1
Round 2 (n=100)	0.92	0.80
Round 3 (n=200)	0.92	0.85
Round 4 (n=334)	0.93	0.91

[a]CNS: central nervous system.

Model Development and Inference

We evaluated three pretrained LLMs representing distinct architectures: (1) GatorTron-base, (2) MedGemma-27B-text-it, and (3) Llama 3.1-70B [31-33]. Selection was restricted to open-weight models so that all inference could be performed within our institutional Health Insurance Portability and Accountability Act (HIPAA)–secure environment without transmitting protected health information externally, allowing for reproducibility at other institutions. Within this constraint, models were chosen to represent the approaches currently used for clinical extraction: a clinically pretrained encoder applied with supervised fine-tuning, a general-domain instruction-tuned generative model applied zero-shot, and a medically pretrained model of the same generative type.

GatorTron is a transformer encoder pretrained on more than 90 billion words of deidentified clinical text, biomedical literature, and other English text [31]. Encoder architectures are well suited to sequence classification using real-world EHR data, and our registry-derived metastasis fields provided labels sufficient for supervised fine-tuning [34]. Additionally, a comparative analysis of pretrained biomedical models reported that GatorTron's ability to learn nuanced clinical information within unstructured clinical text makes it suitable for clinical extraction tasks, compared to other BERT-based models [35]. Prior published comparisons have also demonstrated comparable performance between the GatorTron-base and GatorTron-large models. He et al [36] evaluated both GatorTron-base and GatorTron-large on clinical temporal relation extraction, finding that GatorTron-base with hard prompting achieved the highest micro F_1-score (89.54%), exceeding GatorTron-large (88.06%). Standard fine-tuning of GatorTron-base achieved a micro F_1-score of 85.95% [36]. Peng et al [37] similarly reported strong performance of GatorTron-base across tasks, including clinical concept extraction and medical relation extraction, achieving F_1-scores of 0.89 and 0.96, respectively, on the 2018 n2c2 dataset [38]. Given this comparable performance and the lower computational cost required for deployment, GatorTron-base was selected for this study.

We began by benchmarking the non–fine-tuned GatorTron-base model (345M parameters) for binary classification of overall metastasis on a subset of cohort 1. This initial benchmarking yielded random chance predictions (F_1-score=0.46), indicating a need for further fine-tuning on our institutional cohort [31]. We fine-tuned 2 separate GatorTron-base models for binary classification of overall metastasis (train: n=463, 80%; test: n=116, 20%) and brain/CNS metastasis (train: n=461, 80%; test: n=115, 20%). To mitigate label noise arising from ambiguous phrasing in clinical note text, training incorporated extraction of key linguistic features in the clinical text. First, a list of metastasis-relevant terms, from the CLEVER dictionary, was used to capture mentions of metastasis presence. For fine-tuning of brain/CNS metastasis classification, these terms

XSL•FO
RenderX

were expanded to include brain/CNS site-specific verbiage. Next, NegEx-based negation detection was implemented to handle negated metastasis mentions [38]. Finally, a curated list of uncertainty expressions (eg, “suspicious for metastasis” and “may represent metastasis”) was implemented to handle unconfirmed metastasis mentions. These features were encoded as binary indicators.

Notes in the 80% train set were further partitioned into an 80% train set (overall metastasis: n=370; brain/CNS metastasis: n=369) and 20% validation set (overall metastasis: n=93; brain/CNS metastasis: n=92), stratified based on the binary outcome label. Since clinical notes were aggregated at the patient level prior to preprocessing, the unit of analysis was the aggregated patient document. Following preprocessing, cohort 1 documents averaged 7082.3 tokens in length (median 3382, IQR 1179–8795; range 12-87,293). Summary statistics of note lengths per cohort and average number of notes per patient are provided in Tables S3 and S4 in Multimedia Appendix 1. Patients were first divided into a model training set (80%) and a held-out test set (20%: overall metastasis n=116, brain/CNS metastasis n=115). The training set was then partitioned into training (80%: overall metastasis n=370; brain/CNS metastasis n=369) and validation (20%: overall metastasis n=93; brain/CNS metastasis n=92) subsets, with all splits stratified on the binary outcome label. GatorTron accepts a maximum input of 512 tokens, so documents were tokenized and segmented into overlapping 512-token chunks with a stride of 128 tokens, yielding a mean of approximately 19 chunks per patient document.

We trained a custom classifier head that concatenated the 3 linguistic features, described above, into a single feature vector. This feature vector, along with the text representation generated by GatorTron, was input into a feedforward neural network, which incorporated contextual and rule-based information to generate a binary prediction of metastasis presence. The 20% validation set was used to identify the optimal probability threshold for metastasis classification by evaluating thresholds between 0 and 1 and selecting the probability value that optimized the F_1-score per patient. Final inference was conducted on the 20% hold-out test set (overall metastasis: n=116; brain/CNS metastasis: n=115). Final patient-level predictions were assigned based on the maximum probability predicted across chunks corresponding to each patient’s note. Predictions on cohort 2 were generated using the final fine-tuned model without further retraining. Probability thresholds were optimized on the cohort 1 validation set and applied directly to cohort 2. Optimizing thresholds on cohort 2 would have required its outcome labels and would therefore no longer constitute independent validation. Applying the cohort 1 threshold maintained a consistent decision rule across cohorts and allowed evaluation of whether the approach transfers to an independent cohort with different note characteristics without cohort-specific fine-tuning.

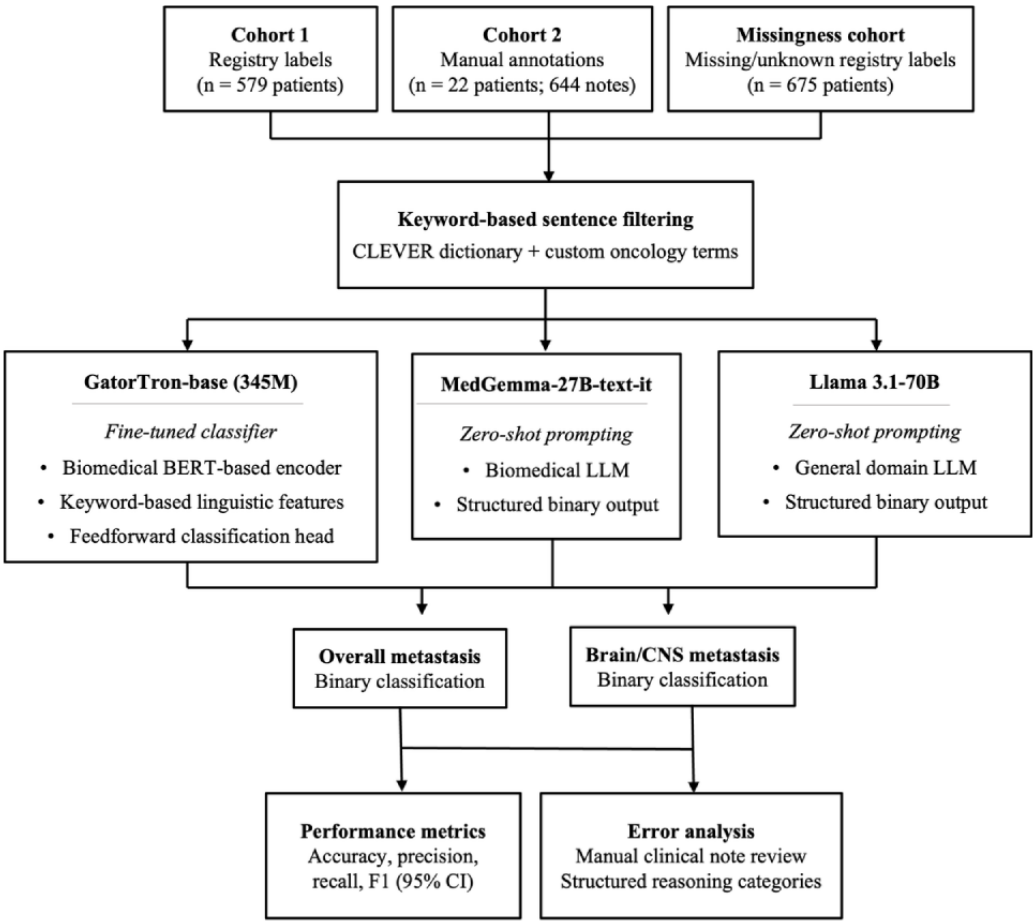
MedGemma-27B-text-it is a biomedical variant of Google’s Gemma model, optimized for medical text and image understanding [32]. Prior literature has demonstrated MedGemma’s ability to perform clinical information extraction from unstructured medical notes using prompt-based strategies, supporting its application in clinical information extraction tasks [39]. However, given MedGemma’s recent release at the time of this study, there was limited evidence of its use in oncology notes for cancer-related information extraction, reflecting a knowledge gap that this study sought to address. We evaluated the 27 billion parameter instruction-tuned text model in a zero-shot setting using structured prompts that instructed the model to classify whether the provided note text documented the target condition and to return a constrained binary response (“Metastasis: YES” or “Metastasis: NO”). Greedy decoding was used, in which stochastic sampling was disabled (do_sample=False), such that the highest probability token was selected at each generation step. A fixed maximum output length was imposed to constrain the models’ responses. Prompt templates are provided in Supplemental Materials 2 and 3 in Multimedia Appendix 1.

Llama 3.1-70B-Instruct is a general-domain open-source LLM developed by Meta, trained on a large-scale corpus of general text and optimized for instruction-following [40]. Given that the prior 2 models were pretrained on biomedical data, evaluation of the Llama model allows for an understanding of the capability of general-domain models to comprehend dense clinical text. Furthermore, the Llama model has previously been shown to outperform BERT-based models and successfully conduct structured extraction tasks from clinical notes [23]. Therefore, its evaluation in this study, alongside GatorTron and MedGemma, may challenge current notions that encoder biomedical models are superior for clinical extraction tasks [23]. Llama3 was evaluated using an identical zero-shot prompt structure to MedGemma, decoding configuration, and constrained binary output format as MedGemma for both extraction tasks. Because the two are similar in architecture and prompting strategy and differ primarily in their pretraining, their comparison is informative about whether medical-domain pretraining is associated with better performance on this task.

Since the encoder model was fine-tuned while the generative models were prompted zero-shot, this study is an evaluation of distinct NLP approaches for metastasis classification rather than a direct comparison of model architectures. A pipeline depicting an overview of note preprocessing and inference is provided in Figure 2. All model weights were obtained from Hugging Face and deployed locally on an institutional, HIPAA-secure computing infrastructure [41-43]. All analyses were implemented in Python 3.10. All code is available on GitHub [44]. Models were run on a high-performance cluster with an AMD EPYC 9254 Processor (24 CPU cores/48 CPU threads), 348 GB RAM, and 2 NVIDIA RTX 6000 Ada Generation GPUs (48 GB VRAM per GPU), and 60 TB of NVMe-backed local scratch space in /tmp. The computational job used to run these models was allocated 2 CPU cores, 128 GB RAM, and 2 GPUs. Training time for the GatorTron model took 32 minutes on the training cohort (n=463). Inference time for all 3 models was approximately 4.5 minutes per model on the test sets (n=116), averaging approximately 2.3 seconds per note.

XSL•FO
RenderX

Figure 2. Note preprocessing and model inference pipeline. CLEVER: Clinical Event Recognizer; CNS: central nervous system; LLM: large language model.



Data Partitioning and Model Evaluation

GatorTron fine-tuning and testing were conducted on cohort 1. MedGemma and Llama3 did not undergo fine-tuning and were instead evaluated under a zero-shot prompt setting. Cohort 1 evaluations for all models were conducted on the 20% test set used in the evaluation of the fine-tuned GatorTron model (overall metastasis: n=116; brain metastasis: n=115).

Cohort 2 served as an independent, out-of-sample test set for all 3 models. Since note-level manual annotations served as the reference standard for cohort 2, model evaluation was conducted at both the note-level (n=644) and the patient-level (n=22). Patient-level reference standards were derived by aggregating note-level annotations, such that if any note of a patient is labeled metastasis positive, the patient-level label is metastasis positive. For each model, cohort, and prediction level, we computed accuracy, precision, recall, and F_1-score. The 95% CIs were estimated using bootstrapped resampling with 10,000 iterations. All metrics were reported separately for the overall metastasis and brain/CNS metastasis tasks.

Furthermore, to assess whether our methods extend beyond aggregated metastasis status, we additionally evaluated our best-performing model on 3 site-specific classification tasks based on available metastasis fields in the cancer registry: bone, liver, and distant lymph node metastasis. The evaluation used the same held-out test set (n=116).

Error Analysis

For the best and poorest performing models on each extraction task in the development cohort, all misclassified cases were examined qualitatively. Clinical notes from each error were manually reviewed and assigned to one of four prespecified error categories: (1) incorrect model prediction despite obvious note content; (2) ambiguous clinical text, in which metastatic status was uncertain or unconfirmed by imaging, laboratory testing, or biopsy; (3) correct model prediction with an erroneous registry label; or (4) missing clinical text, in which relevant documentation was likely absent from the available note. Category definitions and representative examples are provided in Table 4.

XSL•FO
RenderX

Table 4. Structured reasoning categories used in qualitative analysis of prediction errors.

Error category	Description	Example
Incorrect prediction	The model assigned an incorrect prediction despite clear indication of positive or negative metastasis presence in the clinical note text.	• "Suspected metastasis confirmed to be benign meningioma." • In this case, the note explicitly confirms negative metastasis, but the model predicts a positive case.
Ambiguous text	The clinical note text has unclear confirmation of metastasis presence. The note expresses uncertainty of metastasis presence without documented confirmation via imaging, laboratory test, or biopsy. The clinical note text may also indicate suspicion of an additional primary cancer, which the model predicts as metastasis.	• "Brain metastasis cannot be ruled out." • "There are several malignant nodular lesions." • If no imaging, laboratory test, or biopsy confirmation is documented, the model may predict the incorrect label.
Correct prediction and incorrect registry label	In some cases of false positive predictions, the available clinical note demonstrates clear mentions of metastasis presence, either indicating an error in the registry label or an update in the patient's cancer diagnosis not reflected in the registry.	• "Biopsy results confirm osseous mets." • If this confirmation is documented around the time of initial cancer diagnosis but is not reflected in the cancer registry, this may reflect an error in the registry label.
Missing clinical text	In some cases of false negative predictions, the available clinical note does not clearly state any mention of metastasis. Therefore, rather than a failure of the model, these errors indicate the possible presence of additional notes, unavailable to us, which may have included metastasis information.	• If the registry indicates metastasis to a specific site, but there is no language in the clinical note relevant to metastasis, metastasis testing, or even mention of the site, the model prediction will not align with the registry label.

Exploratory Downstream Missingness-Recovery Analysis

The best-performing model for each extraction task was applied to the 675 patients in the missingness-recovery dataset. This analysis was conducted to evaluate the feasibility of leveraging our models to recover metastasis information from narrative text. To assess predictive performance, random samples of 10 predicted-positive and 10 predicted-negative cases were drawn. Their notes were manually reviewed and labeled prior to viewing model predictions to ensure unbiased assessment. Concordance between manual labels and model predictions was then evaluated. Given the small validation sample, findings are exploratory rather than definitive of our model performance.

Ethics Approval

This study was approved by the Emory University Institutional Review Board (protocol 2025P010922) under the expedited review process, with a complete waiver of informed consent and HIPAA authorization.

Results

Cohort Characteristics

Among 1260 patients with NSCLC identified from the Winship Cancer Registry and linked to clinical notes, restricted to analytic cases diagnosed and/or treated at Emory, metastasis status was missing or recorded as unknown in 54% (675/1260) of patients. Cohort 1 included 579 patients with complete registry metastasis data and 24,898 clinical notes; 41% (239/579) had metastasis at any site, and 15% (89/579) had brain metastasis at diagnosis. Cohort 2 included 22 patients and 644 radiology notes. All patients had metastasis at any site, and 36% (8/22) had CNS metastasis. At the note level, 51% (328/644) of notes were positive for metastasis and 11% (71/644) for CNS metastasis. Keyword-based preprocessing reduced mean document length by approximately 71% in both cohorts, from 24,193 to 7082 tokens per aggregated patient note in cohort 1 and from 364 to 107 tokens per note in cohort 2 (Table S3 in Multimedia Appendix 1).

Model Performance: Overall Metastasis Extraction

All models demonstrated strong performance (Table 5). In cohort 1, MedGemma achieved the highest F_1-score (0.80), followed by Llama3 (F_1-score=0.76) and fine-tuned GatorTron (F_1-score=0.72). In cohort 2, all models achieved high patient-level scores, which was expected given the dense radiological documentation and the presence of metastasis in all patients. At the note level, MedGemma achieved the highest F_1-score (0.93), followed by Llama3 (F_1-score=0.88) and GatorTron (F_1-score=0.76). The optimal probability threshold for overall metastasis classification was 0.58.

XSL•FO
RenderX

Table 5. Performance metrics and 95% CIs for overall metastasis prediction.

Model and cohort	Prediction, n	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
GatorTron-base (fine-tuned)					
Cohort 1 (primary development)	116[a]	0.74 (0.66-0.83)	0.71 (0.65-0.81)	0.74 (0.65-0.84)	0.72 (0.63-0.80)
Cohort 2	22[a]	0.95 (0.86-1.0)	1.0 (1.0-1.0)	0.95 (0.86-1.0)	0.98 (0.93-1.0)
Cohort 2	644[b]	0.71 (0.67-0.74)	0.65 (0.61-0.70)	0.91 (0.88-0.94)	0.76 (0.73-0.79)
Llama 3.1-70B-Instruct					
Cohort 1	116[a]	0.76 (0.68-0.84)	0.69 (0.58-0.80)	0.85 (0.74-0.94)	0.76 (0.67-0.84)
Cohort 2	22[a]	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (1.0-1.0)
Cohort 2	644[b]	0.87 (0.85-0.90)	0.84 (0.80-0.88)	0.92 (0.89-0.95)	0.88 (0.85-0.91)
MedGemma-27B-text-it					
Cohort 1	116[a]	0.80 (0.72-0.87)	0.75 (0.64-0.86)	0.85 (0.76-0.94)	0.80 (0.72-0.87)
Cohort 2	22[a]	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (1.0-1.0)
Cohort 2	644[b]	0.93 (0.91-0.95)	0.97 (0.95-0.99)	0.88 (0.85-0.92)	0.93 (0.90-0.95)

[a]Patient-level.

[b]Note-level.

Model Performance: Brain/CNS Metastasis Extraction

In cohort 1, Llama3 achieved the highest F_1-score (0.79), followed by MedGemma (F_1-score=0.76) and GatorTron (F_1-score=0.50). In cohort 2, Llama3 achieved the highest F_1-scores (patient-level F_1-score=0.93; note-level F_1-score=0.86). MedGemma achieved comparable note-level performance (F_1-score=0.86) with a lower patient-level score (F_1-score=0.70). GatorTron achieved a patient-level F_1-score of 0.72 and a note-level F_1-score of 0.50. The optimal probability threshold for brain/CNS metastasis classification was 0.40. Performance metrics for brain/CNS metastasis are provided in Table 6.

Table 6. Performance metrics and 95% CIs for brain/central nervous system metastasis prediction.

Model and cohort	Prediction, n	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
GatorTron-base (fine-tuned)					
Cohort 1 (primary development)	115[a]	0.88 (0.80-0.96)	0.50 (0.38-0.64)	0.50 (0.26-0.71)	0.50 (0.34-0.66)
Cohort 2	22[a]	0.73 (0.55-0.91)	0.57 (0.30-0.82)	1.0 (1.0-1.0)	0.72 (0.46-0.90)
Cohort 2	644[b]	0.81 (0.78-0.84)	0.35 (0.28-0.42)	0.83 (0.75-0.92)	0.50 (0.42-0.67)
Llama 3.1-70B-Instruct					
Cohort 1	115[a]	0.96 (0.91-0.99)	0.91 (0.69-1.0)	0.71 (0.46-0.94)	0.79 (0.59-0.96)
Cohort 2	22[a]	0.95 (0.86-1.0)	0.88 (0.62-1.0)	1.0 (1.0-1.0)	0.93 (0.77-1.0)
Cohort 2	644[b]	0.96 (0.95-0.98)	0.82 (0.73-0.90)	0.90 (0.82-0.97)	0.86 (0.79-0.91)
MedGemma-27B-text-it					
Cohort 1	115[a]	0.95 (0.90-0.98)	0.83 (0.58-1.0)	0.72 (0.45-0.94)	0.76 (0.54-0.92)
Cohort 2	22[a]	0.82 (0.74-0.95)	0.83 (0.43-1.0)	0.63 (0.25-1.0)	0.70 (0.33-0.93)
Cohort 2	644[b]	0.97 (0.95-0.98)	0.83 (0.75-0.92)	0.89 (0.81-0.96)	0.86 (0.79-0.91)

[a]Patient-level.

[b]Note-level.

XSL•FO
RenderX

Error Analysis

Prediction errors from the best and poorest performing models for each classification task in the held-out test set of cohort 1 (overall metastasis n=116; brain/CNS metastasis n=115) were qualitatively analyzed and assigned to a category of reasoning. Corresponding predictions from the other 2 models were assessed in parallel.

Overall Metastasis Prediction

GatorTron

GatorTron generated 14 false negative predictions and 16 false positive predictions. Less than half of the errors generated (n=13) were due to incorrect model errors, all of which were predicted correctly by MedGemma and Llama3. The primary reason for other errors was ambiguous phrasing in the clinical text (n=10), in which clinical documentation indicated "suspicious," "concerning," or "probable" metastasis without clear confirmation through laboratory test results. Another reason was that the model correctly generated positive predictions (n=4), despite a metastasis-negative registry label. In these cases, patient notes explicitly mentioned a diagnosis of metastatic cancer, which was not reflected in the registry but identified by our model. Finally, multiple false negative cases (n=3) were due to missing clinical text. A detailed error analysis is provided in Table S6 in Multimedia Appendix 1.

MedGemma

For prediction of overall metastasis, MedGemma generated 8 false negative predictions and 15 false positive predictions. In 9 instances, the model generated correct positive predictions despite conflicting registry records. Errors also occurred due to missing clinical text (n=6) and ambiguous clinical text (n=6; 4 false positive and 2 false negative predictions). Finally, the model produced one false positive prediction and one false negative prediction, both categorized as incorrect and predicted correctly by Llama3 and GatorTron. A detailed error analysis is provided in Table S7 in Multimedia Appendix 1.

Brain/CNS Metastasis Prediction

GatorTron

For brain/CNS metastasis, GatorTron generated 7 false positive predictions and 7 false negative predictions. Most errors were caused by incorrect predictions (n=10), all of which were predicted correctly by MedGemma and Llama3. Other errors were due to missing text (n=3) and ambiguous text (n=1). A detailed error analysis is provided in Table S8 in Multimedia Appendix 1.

Llama3

Llama3 generated 4 false negative predictions due to missing clinical text and one false positive prediction, which was determined to be incorrect. The incorrect false positive prediction was predicted correctly by MedGemma and GatorTron. A detailed error analysis is provided in Table S9 in Multimedia Appendix 1.

Site-Specific Metastasis Extraction

Since MedGemma was the best-performing model for overall metastasis, we further examined its site-specific metastasis classification performance (Table 7). Performance for bone metastasis (F_1-score=0.79) and liver metastasis (F_1-score=0.82) was comparable to that observed for overall metastasis classification. Performance for distant lymph node metastasis was lower (F_1-score=0.29), reflecting high recall (0.87) with low precision (0.18) at a prevalence of 1.7%.

Table 7. Performance metrics and 95% CIs for site-specific metastasis prediction.

Model and metastasis site (n=116)	Prevalence[a]	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
MedGemma-27B-text-it					
Bone	0.16	0.93 (0.88-0.97)	0.79 (0.58-0.95)	0.79 (0.59-0.95)	0.79 (0.62-0.91)
Liver	0.12	0.96 (0.91-0.99)	0.80 (0.57-1.0)	0.86 (0.64-1.0)	0.82 (0.64-0.96)
Distant lymph nodes	0.02	0.92 (0.87-0.96)	0.18 (0.00-0.45)	1.0 (0.00-1.0)	0.31 (0.00-0.62)

[a]Proportion of test-set patients with registry-recorded metastasis at the specified site (n=116).

Exploratory Missingness-Recovery Analysis

We evaluated whether models could recover metastasis information among 675 patients with at least one missing or unknown registry label (including 12 with unknown brain metastasis status). Using the best-performing model for overall metastasis, MedGemma classified 302 of 675 (44.7%) patients as metastasis-positive, closely resembling the prevalence of metastasis in cohort 1 (239/579, 41%). To assess validity, a random sample of 20 cases (10 predicted positive and 10 predicted negative) was manually annotated prior to viewing model results. Agreement with model predictions yielded accuracy=0.90 and F_1-score=0.89. Two false positive errors were identified in the prediction of overall metastasis due to several ambiguous phrases regarding suspicion of metastasis without clear confirmation. For brain/CNS metastasis, all 12 patients with unknown registry status were manually annotated as negative and correctly classified by Llama3.

Discussion

Principal Findings

LLM-based methods demonstrated strong performance for extracting overall and CNS metastasis from real-world NSCLC clinical notes, with performance varying by task and evaluation setting. Additional exploratory analyses suggest that clinically meaningful information about metastasis can be recovered from narrative documentation when structured registry data are missing. These findings highlight a core limitation of current oncology data systems. Cancer registries, while essential, are

XSL•FO
RenderX

temporally limited and incomplete, whereas clinical notes can provide longitudinal, detailed documentation of disease evolution. LLMs offer a practical mechanism to extract clinically relevant information within routine documentation, complementing structured data.

Earlier studies have largely focused on single cohorts or limited note types, particularly radiology reports and discharge summaries [24,25]. By leveraging a registry-linked cohort with heterogeneous oncology documentation, we provide a more realistic evaluation of real-world performance. Furthermore, few studies have evaluated and compared distinct NLP strategies, including clinically pretrained encoder models and large instruction-tuned generative models, on the same task [8,11,12]. In our study, baseline performance metrics demonstrated that the generative models (Llama3 and MedGemma) achieved higher performance in classifying both overall and brain/CNS metastasis presence, respectively, compared to the fine-tuned GatorTron model. Upon further inspection of model errors, generative models and fine-tuned encoder models appeared to serve complementary strengths, where certain prediction errors made by one model were correctly predicted by another. Additionally, while GatorTron was fine-tuned on clinical notes using linguistic features, Llama 3.1 and MedGemma were evaluated under zero-shot prompting strategies. Therefore, rather than a direct comparison of model architectures, this study showcases distinct LLM-based strategies for metastasis classification of clinical notes. CIs for several estimates overlap substantially, particularly for brain/CNS metastasis in cohort 1 and across tasks in cohort 2, and differences in reported scores should not be interpreted as establishing that one approach outperforms another.

Fine-tuned encoder models show promise for classifying metastasis presence and identifying potentially inaccurate registry records based on clinical note text; however, they may be sensitive to discrepancies present in their input data. Instruction-tuned generative models, even in a zero-shot prompt setting, demonstrate holistic understanding of clinical note text without this limitation, which may explain the differences in model performance in this study.

In this multicohort evaluation, we make 3 principal observations. First, metastasis information was substantially incomplete even within a cancer registry-linked dataset, with more than half of patients lacking definitive structured metastasis labels. Second, model performance varied, with the instruction-tuned generative models demonstrating holistic understanding of complex clinical language and the fine-tuned encoder model demonstrating strong classification performance but likely hindered by potentially inaccurate registry labels during training. Third, LLMs appear adept at correcting inaccurate registry records and recovering missing data.

Our first observation in this study is the high prevalence of missing data in structured metastasis labels, representing 54% (675/1260) of the patients identified in our initial extraction of cohort 1. This study extends prior clinical NLP work in important ways. While cancer registries serve as the gold standard for cancer surveillance, future studies may seek to use more curated real-world datasets for more robust model fine-tuning. Though inaccuracies and missingness in the cancer registry may have inhibited the fine-tuned encoder model from more accurately distinguishing between positive and negative metastasis cases, our model still demonstrated the ability to correct inaccurate data based on clinical text documentation. The use of metastasis labels from the cancer registry presents both an advantage and a challenge to fine-tuning our GatorTron model. Our cohort size (n=579 patients for overall metastasis; n=576 patients for brain/CNS metastasis) was comparable to prior work that fine-tuned the same model for metastasis extraction, which leveraged manually annotated computed tomography reports from 550 patients [45]. That work, however, manually annotated labels at the report level. Because registry metastasis fields are recorded at the patient level, each of our labels applied instead to an aggregated document (mean 7082, SD 9841 tokens) drawn from 69 distinct note types in which metastatic disease may be expressed in different forms. This heterogeneity better reflects real-world documentation but provides weaker supervision for a fine-tuned model, and it is intrinsic to registry-linked data. We did not evaluate performance as a function of training set size and therefore cannot determine whether the fine-tuned model was data-limited. The zero-shot models required no labeled data and were not subject to these constraints; their comparable or higher performance suggests that zero-shot approaches may be better suited to settings where only patient-level reference labels are available.

Our second observation in this study is that model performance patterns underscore that no single approach is uniformly optimal. Encoder models align well with structured labels; however, discrepancies in registry records may influence model outputs. In contrast, generative models appear more robust to heterogeneous, implicit, and context-dependent clinical language, which is particularly relevant for CNS metastasis, where documentation is often indirect and radiology-driven. This difference in performance may be attributed to the ability of generative models to better identify nuanced and complex site-specific language used in the documentation of brain/CNS without the influence of structured registry labels. Furthermore, GatorTron's prediction of brain/CNS metastasis may have been inhibited by the class imbalance in brain/CNS metastasis cases during model fine-tuning, whereas this class imbalance was not impactful in a zero-shot prompting setting. These results support a use-case-specific deployment strategy rather than a single-model solution.

Our third observation in this study is that LLM-based extraction appears to be adept at correcting and recovering incomplete metastasis data in structured registry fields. Our error analysis indicated that many discrepancies were due to data limitations in reference labels or available clinical documentation rather than explicit model failure. Differences in registry labels compared to clinical documentation, missing clinical text, and ambiguity in clinical language accounted for most errors. In several cases, model predictions were clinically plausible despite conflicting with registry labels, underscoring the imperfect nature of structured ground truth in oncology datasets. Both the error analysis and exploratory missingness recovery analysis provide an initial signal that NLP may augment registry accuracy

XSL•FO
RenderX

and completeness. The error analysis identified several positive metastasis cases that were not reflected in the registry labels, despite clinical documentation near the time of initial diagnosis. The exploratory missingness recovery analysis demonstrated high accuracy in recovering missing metastasis labels, based on manual review of a subset of notes. While preliminary, this suggests that clinically relevant signals embedded in narrative documentation can be systematically recovered. These findings have direct clinical and operational implications. NLP-based extraction can support cancer registry workflows by prioritizing high-yield cases and improving completeness. It may also enable more accurate cohort construction and clinical trial screening, where metastasis status, particularly CNS involvement, is a key eligibility criterion.

Our study contributes to the current field of clinical NLP in a few key ways. First, previous metastasis extraction studies have predominantly focused on single document types, namely radiology reports and discharge summaries, while our evaluation included 69 distinct note types [24,25,45]. The incorporation of heterogeneous clinical documentation more closely reflects the setting in which LLM-based cancer abstraction would be deployed. It also provides more robust evidence of each model's ability to comprehend and extract structured data from complex, varied clinical language. Second, we evaluated not just distinct LLMs but distinct modeling approaches for metastasis extraction. Recent literature has begun to compare different LLM-based approaches for clinical note extraction, including encoder-only, decoder-only, and prompt-based instruction-tuned LLMs in zero-shot and few-shot settings [37,46]. Such comparisons have not previously been reported for metastasis extraction. Because models were evaluated using different adaptation strategies, this study does not permit a direct comparison of model architectures, but it does provide an evaluation of distinct LLM-based methodologies for clinical note extraction. Third, we used both registry data and human annotations as reference standards, in cohort 1 and cohort 2 respectively. This allowed model performance to be assessed against existing structured registry fields while independently evaluating it against human annotation. Metastasis status was missing or recorded as unknown for 54% (675/1260) of registry-linked patients in the source population, consistent with reports of incomplete and discordant registry documentation [47].

Our error analysis further indicated that many apparent model errors reflected gaps in the registry rather than misclassification, illustrating the limitations of registry-derived labels as a reference standard for model development. Finally, despite these inconsistencies, LLM-based approaches may support the recovery of missing or incorrect registry labels: in the exploratory recovery analysis, predictions for patients with missing registry values agreed with manual annotation at accuracy 0.90 and F_1-score 0.89. Few studies have used LLM-based methods to assess the recovery of information missing from structured registry fields.

Our study has several limitations. First, cohort 1 (the primary development cohort) relied on cancer registry labels that contain limitations, including accuracy of label and inability to capture metastatic progression. Thus, discordance between registry labels and model predictions does not necessarily indicate model error, since the reference standard itself may be imperfect. Furthermore, there is a possibility that, though documentation spanned one month before and one month after initial diagnosis, metastasis cases diagnosed within one month after initial cancer diagnosis were not updated in the registry.

Second, cohort 2 was small and derived from clinical trial populations with dense radiologic follow-up, which may limit generalizability to routine care settings. All models demonstrated higher performance in cohort 2, likely due to the minimal heterogeneity of note types. Radiology notes are likely to contain dense documentation of metastatic disease presence, which may have contributed to improved model performance in cohort 2. Moreover, all patients in cohort 2 had overall metastasis, so patient-level metrics for that task do not reflect the models' ability to distinguish metastasis-positive from metastasis-negative patients. The note-level evaluation is therefore the more informative measure of discriminative performance in this cohort, and patient-level results should be interpreted as complementary rather than as an indication of the model's discriminative ability. Note-level metrics may, themselves, be optimistic, since each patient contributed multiple notes that often contained similar or repeated phrasing and are therefore not independent observations.

Third, the performance metrics of cohort 2 require further contextualization, given its small cohort size and class imbalance in the patient-level evaluation. Since all patients were found to have metastasis, the patient-level performance metrics do not demonstrate the models' ability to distinguish between metastasis-positive and metastasis-negative patients. Additionally, the note-level performance metrics may reflect a biased overestimation of the models' ability to classify metastasis presence. Each patient contributed multiple notes, which may have contained similar or repeated phrasing, likely leading the model to assign identical metastasis predictions for multiple notes based on similar language seen across notes.

Fourth, the use of different reference standards across cohorts (registry-based at the patient level in cohort 1 vs manual note-level annotation in cohort 2) complicates direct comparison, particularly given differences in temporal scope and definition of CNS metastasis. Certain prediction errors made by the models on cohort 1 may have been a result of incorrect registry labels. Since the ground truth label was more reliable in cohort 2 notes, model performance may have appeared to be better in this cohort.

Fifth, models of differing underlying architectures were evaluated in distinct ways, wherein GatorTron was fine-tuned, while generative models were evaluated in a zero-shot setting. Additionally, encoder models with extended context windows, compared to GatorTron, such as those built on the ModernBERT architecture, have since become available and can process inputs of up to 8192 tokens. Such models would accommodate most documents in this cohort without the chunking and max-pooling strategies used in our study. Evaluating whether long-context modeling improves performance of metastasis classification would be a useful extension of this work. The difference in

XSL•FO
RenderX

computational requirements across the 3 models is a crucial consideration for deployment of these models in real-world clinical settings. The Llama 3.1 70-billion-parameter model has larger memory requirements compared to the GatorTron 345-million-parameter model, and this difference should be considered when inference is run on constrained hardware. Model selection in real-world clinical settings should consider this trade-off between performance and computational requirements.

Sixth, the keyword-based preprocessing strategy may have excluded relevant contextual information, potentially limiting sensitivity to implicitly expressed metastasis. Seventh, the study was conducted at a single institution, and broader external validation is needed. Finally, the missing recovery analysis was exploratory and based on a small sample. A larger-scale validation should be conducted to further evaluate the potential of recovering missing metastasis records. Similarly, site-specific evaluation was informative for bone and liver metastasis but not for distant lymph node metastasis, where prevalence in the test set was 1.7%, and the resulting estimates were unstable. Our estimates demonstrated a high presence of false positive predictions, so future evaluations of site-specific metastasis at infrequently documented sites would require a much larger test set.

Several factors affect the generalizability of our findings. Both patient cohorts were specific to one health care institution, and all patients were diagnosed with NSCLC. These constraints pose a challenge in assessing the models' generalizability to cohorts at other institutions and with other cancer types, and clinical terminologies and documentation practices may differ elsewhere. Our study design does, however, offer 2 features relevant to interpretation. First, the variation in clinical note types allows the model evaluations to be more representative of real-world EHR documentation. Second, the use of 2 cohorts from different documentation periods, with different reference standards, provided a second evaluation setting and a preliminary view of model performance across distinct clinical language and documentation styles. Cohort 2 was small (n=22), drawn from a clinical trial population with dense radiologic follow-up, and comprised entirely of metastasis-positive patients, so it does not by itself establish generalizability of our methods. It provides a second evaluation setting with an independent reference standard and a different documentation period, but its size and composition prevent its results from generalizing to routine care. Additionally, the patient-level estimates in this cohort carry wide CIs. Similarly, the missingness-recovery analysis was validated against manual annotation of 20 cases, demonstrating the feasibility of our methods but not sufficient to estimate recovery performance with precision. Therefore, the results of the error analysis should be regarded as preliminary. Both findings require confirmation in larger and more diverse cohorts in order to demonstrate the reliability of these methods in other settings.

Conclusion

In this study, the instruction-tuned generative models, evaluated in a zero-shot setting, performed best at classifying metastasis. The high degree of missing metastasis data, even within a cancer registry-linked cohort, underscores a key limitation of structured oncology data and highlights the value of note-based ascertainment. Exploratory analyses further suggest that metastasis information documented in clinical narratives can be partially recovered when structured fields are incomplete. These findings support the use of NLP to complement cancer registry data for metastasis ascertainment, cohort identification, and trial screening, while requiring further validation before broader use.

Acknowledgments

Patient cohort identification was conducted by the Winship Cancer Institute's Data and Technology Applications Shared Resource (WinDATA). Clinical note extraction for select patients was conducted by Emory University's Medical Informatics and AI Core team.

Claude Sonnet 4.6 (Anthropic) was used between April 7-13, 2026 to suggest fixes for Python code errors or improvements to code. Any changes made to our code based on Claude's suggestions were manually verified. Claude was also used occasionally to rephrase sentences.

Funding

This project is funded by The Fund for Innovation in Cancer Informatics (ICI). S Bozkurt is also supported by the National Library of Medicine (R01LM015395). The funders had no role in study design, data collection, analysis, interpretation, or preparation of the manuscript.

Data Availability

The datasets analyzed in this study are not publicly available due to institutional policies regarding Health Insurance Portability and Accountability Act (HIPAA)–protected medical files.

Authors' Contributions

Conceptualization: S Bozkurt, RBP
Data curation: S Balaji, RZC, S Bozkurt, DS

XSL•FO
RenderX

Formal analysis: S Balaji, KC
Funding acquisition: S Bozkurt
Investigation: S Balaji, S Bozkurt
Methodology: S Balaji, S Bozkurt, KC
Project administration: S Balaji, S Bozkurt
Supervision: S Bozkurt, RBP, JDW, AS
Validation: S Bozkurt, KC
Visualization: S Balaji
Wiring – original draft: S Balaji, S Bozkurt
Writing – review & editing: S Balaji, S Bozkurt, KC, RZC, RBP, JDW, MAR, DS

Conflicts of Interest

JDW receives support from Johnson & Johnson, Food and Drug Administration (FDA), National Institute on Alcohol Abuse and Alcoholism/National Institutes of Health (NIAAA/NIH), and Arnold Ventures for research outside the submitted work.

Multimedia Appendix 1
Note distributions, preprocessing and feature extraction keywords, metastasis annotation guidelines, large language model prompt templates, and model error analyses.
[DOCX File , 65 KB - cancer_v12i1e99573_app1.docx]

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263 [FREE Full text] [doi: 10.3322/caac.21834] [Medline: 38572751]
2. What is lung cancer? Types of lung cancer. American Cancer Society. URL: https://www.cancer.org/cancer/types/lung-cancer/about/what-is.html [accessed 2026-04-08]
3. Sleeman J, Steeg PS. Cancer metastasis as a therapeutic target. Eur J Cancer 2010;46(7):1177-1180 [FREE Full text] [doi: 10.1016/j.ejca.2010.02.039] [Medline: 20307970]
4. Donnelly C. Newer-generation TKIs show promise for CNS metastasis in NSCLC. Oncol Times 2025;47(11):1. [doi: 10.1097/01.cot.0000000000000225]
5. Amsbaugh MJ, Kim CS. Brain Metastasis. Treasure Island, FL: StatPearls Publishing; 2026.
6. Ostrom Q, Huang WC, Barnholtz-Sloan JS. Brain metastases: epidemiology. In: Handbook of Clinical Neurology. Amsterdam: Elsevier; 2018:27-42.
7. Kirshner J, Cohn K, Dunder S, Donahue K, Richey M, Larson P, et al. Automated electronic health record-based tool for identification of patients with metastatic disease to facilitate clinical trial patient ascertainment. JCO Clin Cancer Inform 2021;5:719-727. [doi: 10.1200/CCI.20.00180] [Medline: 34197178]
8. Menezes ME, Das SK, Minn I, Emdad L, Wang XY, Sarkar D, et al. Detecting tumor metastases: the road to therapy starts here. Adv Cancer Res 2016;132:1-44 [FREE Full text] [doi: 10.1016/bs.acr.2016.07.001] [Medline: 27613128]
9. Seneviratne MG, Banda JM, Brooks JD, Shah NH, Hernandez-Boussard TM. Identifying cases of metastatic prostate cancer using machine learning on electronic health records. AMIA Annu Symp Proc 2018;2018:1498-1504 [FREE Full text] [Medline: 30815195]
10. Petkov VI, Penberthy LT, Dahman BA, Poklepovic A, Gillam CW, McDermott JH. Automated determination of metastases in unstructured radiology reports for eligibility screening in oncology clinical trials. Exp Biol Med (Maywood) 2013;238(12):1370-1378 [FREE Full text] [doi: 10.1177/1535370213508172] [Medline: 24108448]
11. Yang DX, Khera R, Miccio JA, Jairam V, Chang E, Yu JB, et al. Prevalence of missing data in the national cancer database and association with overall survival. JAMA Netw Open 2021;4(3):e211793 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.1793] [Medline: 33755165]
12. Cancer surveillance programs and data tracking. American Cancer Society. URL: https://www.cancer.org/cancer/preparing-for-treatment/cancer-surveillance-programs-and-data-tracking.html [accessed 2026-04-20]
13. Alba PR, Gao A, Lee KM, Anglin-Foote T, Robison B, Katsoulakis E, et al. Ascertainment of veterans with metastatic prostate cancer in electronic health records: demonstrating the case for natural language processing. JCO Clin Cancer Inform 2021;5:1005-1014. [doi: 10.1200/CCI.21.00030] [Medline: 34570630]
14. Alzu'bi AA, Watzlaf VJM, Sheridan P. Electronic health record (EHR) abstraction. Perspect Health Inf Manag 2021;18(Spring):1g [FREE Full text] [Medline: 34035788]
15. Bozkurt S, Magnani CJ, Seneviratne MG, Brooks JD, Hernandez-Boussard T. Expanding the secondary use of prostate cancer real world data: automated classifiers for clinical and pathological stage. Front Digit Health 2022;4:793316 [FREE Full text] [doi: 10.3389/fdgth.2022.793316] [Medline: 35721793]

XSL•FO
RenderX

16. Mykowiecka A, Marciniak M, Kupść A. Rule-based information extraction from patients' clinical data. J Biomed Inform 2009;42(5):923-936 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.007] [Medline: 19646551]
17. Liu S, McCoy AB, Chen Q, Wright A. Integrating rule-based NLP and large language models for statin information extraction from clinical notes. Int J Med Inform 2026;205:106104 [FREE Full text] [doi: 10.1016/j.ijmedinf.2025.106104] [Medline: 40925145]
18. Goryachev SD, Wu JT, Lin E, Friedman DR, Zwolinski R, Dhond R, et al. Rule-based natural language processing to extract clinical trial and research study enrollment history from unstructured notes. Health Informatics J 2026;32(1):14604582261430614 [FREE Full text] [doi: 10.1177/14604582261430614] [Medline: 41912375]
19. Kaster L, Hillis E, Oh IY, Cordell EC, Foraker RE, Lai AM, et al. Comparison of rule- and large language model-based phenotype extraction from clinical notes for neurofibromatosis type 1. J Am Med Inform Assoc 2025;32(11):1663-1673 [FREE Full text] [doi: 10.1093/jamia/ocaf155] [Medline: 40966762]
20. Thakkar V, Silverman GM, Kc A, Ingraham NE, Jones EK, King S, et al. A comparative analysis of large language models versus traditional information extraction methods for real-world evidence of patient symptomatology in acute and post-acute sequelae of SARS-CoV-2. PLoS One 2025;20(5):e0323535. [doi: 10.1371/journal.pone.0323535] [Medline: 40373001]
21. Rijcken E, Zervanou K, Mosteiro P, Scheepers F, Spruit M, Kaymak U. Machine learning vs. rule-based methods for document classification of electronic health records within mental health care—a systematic literature review. Nat Lang Process J 2025;10:100129. [doi: 10.1016/j.nlp.2025.100129]
22. Arango-Argoty G, Kipkogei E, Stewart R, Sun GJ, Patra A, Kagiampakis I, et al. Pretrained transformers applied to clinical studies improve predictions of treatment efficacy and associated biomarkers. Nat Commun 2025;16(1):2101 [FREE Full text] [doi: 10.1038/s41467-025-57181-2] [Medline: 40025003]
23. Hu Y, Zuo X, Zhou Y, Peng X, Huang J, Keloth VK, et al. Information extraction from clinical notes: are we ready to switch to large language models? J Am Med Inform Assoc 2026;33(3):553-562 [FREE Full text] [doi: 10.1093/jamia/ocaf213] [Medline: 41533750]
24. Yuan K, Yoon CH, Gu Q, Munby H, Walker AS, Zhu T, et al. Transformers and large language models are efficient feature extractors for electronic health record studies. Commun Med (Lond) 2025;5(1):83 [FREE Full text] [doi: 10.1038/s43856-025-00790-1] [Medline: 40119150]
25. Raghavan P, Chen JL, Fosler-Lussier E, Lai AM. How essential are unstructured clinical narratives and information fusion to clinical trial recruitment? AMIA Jt Summits Transl Sci Proc 2014;2014:218-223 [FREE Full text] [Medline: 25717416]
26. Kang YJ, Lee H, Yi JP, Kim H, Yoon CI, Baek JM, et al. Large language model versus manual review for clinical data curation in breast cancer: retrospective comparative study. JMIR Med Inform 2025;13:e73605 [FREE Full text] [doi: 10.2196/73605] [Medline: 41197113]
27. Peng C, Yang X, Chen A, Smith KE, PourNejatian N, Costa AB, et al. A study of generative large language model for medical research and healthcare. NPJ Digit Med 2023;6(1):210 [FREE Full text] [doi: 10.1038/s41746-023-00958-w] [Medline: 37973919]
28. Arzideh K, Schäfer H, Allende-Cid H, Baldini G, Hilser T, Idrissi-Yaghir A, et al. From BERT to generative AI - comparing encoder-only vs. large language models in a cohort of lung cancer patients for named entity recognition in unstructured medical reports. Comput Biol Med 2025;195:110665 [FREE Full text] [doi: 10.1016/j.compbiomed.2025.110665] [Medline: 40554973]
29. Chen SF, Alyakin A, Seas A, Yang E, Choi JJ, Lee JV, et al. LLM-assisted systematic review of large language models in clinical medicine. Nat Med 2026;32(3):1152-1159. [doi: 10.1038/s41591-026-04229-5] [Medline: 41776077]
30. stamang/CLEVER. GitHub. URL: https://github.com/stamang/CLEVER/blob/master/res/dicts/base/clever_base_terminology.txt [accessed 2026-04-06]
31. Yang X, Chen A, PourNejatian N, Shin HC, Smith KE, Parisien C, et al. A large language model for electronic health records. NPJ Digit Med 2022;5(1):194 [FREE Full text] [doi: 10.1038/s41746-022-00742-2] [Medline: 36572766]
32. MedGemma. Google DeepMind. URL: https://deepmind.google/models/gemma/medgemma/ [accessed 2026-03-17]
33. Introducing Meta Llama 3: the most capable openly available LLM to date. Meta AI. URL: https://ai.meta.com/blog/meta-llama-3/ [accessed 2026-04-17]
34. Hasheminasab SA, Jamil F, Afzal MU, Khan AH, Ilyas S, Noor A, et al. Evaluating large language models for clinical note processing: local fine-tuning and internal-external validation using electronic health records from South Asia. BMC Med Inform Decis Mak 2026;26(1):75 [FREE Full text] [doi: 10.1186/s12911-026-03366-8] [Medline: 41742185]
35. Wang B, Xie Q, Pei J, Chen Z. Pre-trained language models in biomedical domain: a systematic survey. ACM Comput Surv 2023;56(3):1-52 [FREE Full text] [doi: 10.1145/3611651]
36. He J, Rasmy L, Li H, Li J. Prompting large language models for clinical temporal relation extraction. arXiv Preprint posted online on December 4, 2024. [doi: 10.48550/arXiv.2412.04512]
37. Peng C, Dong X, Lyu M, Paredes D, Zhang Y, Wu Y. A study of large language models for patient information extraction: model architecture, fine-tuning strategy, and multi-task instruction tuning. J Biomed Inform 2026;178:105034. [doi: 10.1016/j.jbi.2026.105034] [Medline: 41905520]
38. Pizarro J. jenojp/negspacy. URL: https://github.com/jenojp/negspacy [accessed 2026-04-23]

XSL•FO
RenderX

39. Balachandran A, Durgapraveen B, Sudhagar G, Rajkumar S. EXL health AI lab at MEDIQA-OE 2025: evaluating prompting strategies with MedGemma for medical order extraction. 2025 Presented at: Proceedings of the 7th Clinical Natural Language Processing Workshop; 2025 October 27-30; Virtual p. 68-74 URL: https://aclanthology.org/2025.clinicalnlp-1.8/
40. Introducing Llama 3.1: our most capable models to date. Meta AI. URL: https://ai.meta.com/blog/meta-llama-3-1/ [accessed 2026-04-21]
41. UFNLP/gatortron-base. Hugging Face. URL: https://huggingface.co/UFNLP/gatortron-base [accessed 2026-03-17]
42. google/medgemma-27b-text-it. Hugging Face. 2026. URL: https://huggingface.co/google/medgemma-27b-text-it [accessed 2026-03-17]
43. meta-llama/Llama-3.1-70B. Hugging Face. URL: https://huggingface.co/meta-llama/Llama-3.1-70B [accessed 2026-03-17]
44. bozlab/MetsExtrac. GitHub. URL: https://github.com/bozlab/MetsExtract [accessed 2026-04-23]
45. Tay SB, Low GH, Wong GJE, Tey HJ, Leong FL, Li C, et al. Use of natural language processing to infer sites of metastatic disease from radiology reports at scale. JCO Clin Cancer Inform 2024;8:e2300122. [doi: 10.1200/CCI.23.00122] [Medline: 38788166]
46. Spiero I, Rijk MH, Scheeres MA, Rutten FH, Geersing G, Platteel TN, et al. Comparison of local large language models for extraction of signs and symptoms data from electronic health records. PLoS One 2026;21(6):e0350625 [FREE Full text] [doi: 10.1371/journal.pone.0350625] [Medline: 42268852]
47. Brauer DG, Sekiya A, Vobugari N, Prakash A, Ozed-Williams B, Ecker BL, et al. Assessing concordance of cancer registry data and electronic health record data in metastatic cancer: a cohort study of patients with colorectal liver metastasis in a U.S. integrated health system. Cancer Epidemiol 2026;102:103028. [doi: 10.1016/j.canep.2026.103028] [Medline: 41763113]

Abbreviations

CLEVER: Clinical Event Recognizer
CNS: central nervous system
EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
LLM: large language model
ML: machine learning
MRI: magnetic resonance imaging
NLP: natural language processing
NSCLC: non–small cell lung cancer

Edited by M Balcarras; submitted 27.Apr.2026; peer-reviewed by G Panda, M Torii, J Zheng, P Bolourchi; comments to author 17.Jul.2026; revised version received 09.Sep.2026; accepted 10.Sep.2026; published 16.Oct.2026.

Please cite as:
Balaji S, Campbell K, Chen RZ, Smith D, Reyna MA, Sarker A, Wallach JD, Parikh RB, Bozkurt S
Metastasis Extraction From Non–Small Cell Lung Cancer Clinical Notes Using Large Language Model–Based Classification: Retrospective Comparative Study
JMIR Cancer 2026;12:e99573
URL: https://cancer.jmir.org/2026/1/e99573
doi:10.2196/99573
PMID:

©Sweta Balaji, Kiersten Campbell, Rou-Zhen Chen, Daniel Smith, Matthew A Reyna, Abeed Sarker, Joshua D Wallach, Ravi B Parikh, Selen Bozkurt. Originally published in JMIR Cancer (https://cancer.jmir.org), 16.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Information-Seeking Patterns and Communication Preferences Among Japanese Survivors With Cancer: Cross-Sectional Analysis

Manae Hirayama*, MTech; Daisuke Kasai, PhD; Ken-ichi Tabei*, PhD

Graduate School of Industrial Technology, Advanced Institute of Industrial Technology, Tokyo Metropolitan Public University Corporation, 1-10-40, Higashiooi, Shinagawa, Tokyo, Japan

*these authors contributed equally

Corresponding Author:
Ken-ichi Tabei, PhD
Graduate School of Industrial Technology, Advanced Institute of Industrial Technology, Tokyo Metropolitan Public University Corporation, 1-10-40, Higashiooi, Shinagawa, Tokyo, Japan

Abstract

Background: With increasing numbers of survivors with cancer, the importance of patient-centered information provision and communication to alleviate psychological burdens, such as anxiety and depression, is growing. However, substantial individual differences exist in patients with cancer information–seeking behaviors and use of support services, and few studies have comprehensively examined cognitive and psychological factors such as treatment status, sex, trust in information sources, and patient-provider relationships.

Objective: This study aimed to integrate the theory of planned behavior and the patient-provider relationship model to identify latent subgroups among Japanese survivors with cancer using information-seeking behaviors, difficulties in information seeking, trust in information sources, and intentions to use psychosocial support services recommended by medical institutions.

Methods: A CHERRIES (Checklist for Reporting Results of Internet E-Surveys)–compliant cross-sectional web survey was conducted in December 2024 with 350 Japanese survivors with cancer (at least 1 year post diagnosis, either undergoing treatment, or within 5 years after completing treatment). Exploratory factor analysis examined items such as difficulties in information seeking, trust in information sources, and assessment of relationships with physicians. Using the resulting factor structure and sociodemographic and clinical characteristics, latent class analysis was conducted. Differences between classes were examined using the chi-square test, Kruskal-Wallis test, and post hoc analyses.

Results: Latent class analysis classified participants into 3 groups: women under observation, men under observation, and patients under treatment. In the women under observation group, evaluation of the reliability of information from nonmedical institutions was significantly higher than in the male group (χ^2_2 =12.30; *P*=.002). Although information-seeking behavior among men under observation was relatively limited, their evaluation of relationships with physicians was significantly higher than that of the treatment group (χ^2_2 =12.20; *P*=.002). The proportion of men who regarded promoting communication with doctors and health care professionals as a benefit of using support services was also significantly higher than that of women (χ^2_2 =12.57; *P*=.002 and Class 2 > Class 1; *P*=.001). In the treatment group, searches for information on life during treatment (χ^2_2 =7.22; *P*=.03), use of the Cancer Consultation Support Center (χ^2_2 =17.31; *P*<.001), and use of the National Cancer Center website (χ^2_2 =7.59; *P*=.02) were significantly higher than among men under observation. The treatment group also reported greater difficulty in seeking information (χ^2_2 =11.90; *P*=.003).

Conclusions: Information-seeking behaviors, trusted sources, and perceived difficulties differed by sex and treatment stage among Japanese survivors with cancer. Patients undergoing treatment showed high information needs but greater difficulty in seeking information, suggesting reduced perceived behavioral control. Men under follow-up emphasized relationships with physicians, whereas women relied more on nonmedical information sources. These findings indicate that psychosocial support and information provision should be optimized according to patient-provider communication patterns.

(JMIR Cancer 2026;12:e79065) doi:10.2196/79065

KEYWORDS

patient preference; latent class analysis; survivor with cancer; mental health; patient-centered communication; theory of planned behavior

XSL•FO
RenderX

Introduction

Currently, approximately 20 million people worldwide are newly diagnosed with cancer each year [1], and this number is expected to reach 35 million by 2050 [2]. In Japan, the number of new patients with cancer is projected to reach approximately 4 million between 2035 and 2039 [3]. Owing to advances in treatment methods and early detection, many individuals diagnosed with cancer are now long-term survivors, and the number of people classified as survivors with cancer who live with and beyond cancer is increasing [4]. Regardless of treatment status, survivors with cancer tend to face various difficulties, including physical, mental, economic, social, and emotional challenges [5]. In particular, the prevalence of psychological stress, such as anxiety and depression, is high, making it extremely important to address their emotional needs [6,7]. Against this background, health care professionals are expected to provide appropriate information and high-quality communication so that patients can confidently make decisions about their diagnosis, prognosis, and treatment options [8].

The most frequently reported unmet nonmedical need among patients with cancer is information [9]. Accurate knowledge of the disease and its treatment affects psychological well-being, quality of life, and satisfaction with treatment [10]. However, the quality of information, access to it, and the individual's level of understanding pose challenges. Improving the quality of the information provided and accurately identifying patient needs also leads to better communication between patients and physicians [11,12]. Nonetheless, many survivors with cancer report that the medical terms and explanations used by health care professionals are difficult to understand, resulting in insufficient information and communication [13,14]. Therefore, while patient-centered communication [15,16]—which enhances the reliability of information for survivors with cancer—is important, dissatisfaction may not stem solely from a simple lack of communication but may instead be intricately related to individual cognitive and psychological factors, which need to be elucidated.

Given this background, we conducted a pilot study using content analysis of free-text responses regarding the use of psychosocial support services among survivors with cancer aged 40 years and older [17]. The results suggested that factors such as trust in support providers, expectations for emotional regulation, difficulty in verbalizing emotions, a desire for independence, the personality of support providers, clarity of consultation procedures, and awareness and dissemination of services are related to the use of support services [17]. These findings suggest that the usage of psychosocial support services may be influenced not only by access and information deficiencies but also by individual cognitive and behavioral characteristics, including personal values. Thus, it is necessary to introduce a theoretical framework to comprehensively and quantitatively capture these complex factors.

The theory of planned behavior (TPB) has been widely used to explain behavior change among patients with chronic illnesses [18,19]. According to the TPB, behavioral intention is formed by 3 factors: "attitude," "subjective norms," and "perceived behavioral control" [20]. However, because the behavior of patients with chronic illnesses is influenced not only by internal cognitive factors but also by external factors such as the living environment, support networks, and health care systems, in this study we integrated external and relational factors, such as the living environment, support networks, and health care systems of survivors with cancer, into the TPB framework to deepen our understanding of the formation of behavioral intentions.

When addressing communication between patients and health care professionals, which is a central focus of this study, the patient-provider relationship model [21] is important. This model explains the impact of the relationship between health care professionals and patients on health outcomes, such as survival, treatment and recovery, relief of symptoms and distress, emotional well-being, and pain control. Communication functions are classified into relationship-oriented functions, which emphasize strengthening the therapeutic relationship and responding to emotions, and task-oriented functions, which involve exchanging information, decision-making, supporting self-management, and managing uncertainty [21].

Previous research has often examined information-seeking behaviors and the use of support services among survivors with cancer on the basis of single attributes, such as treatment phase or sex [22-24]. There have been limited attempts to comprehensively typify diverse individual behavioral characteristics and psychological factors and theoretically clarify their backgrounds. Therefore, in this study, we aimed to fill this gap by using latent class analysis (LCA) to identify behavioral patterns and by examining the characteristics of these patterns from the perspectives of the TPB and the relationship-oriented and task-oriented functions of the patient-provider relationship model.

On the basis of the above theoretical framework, we focused on the following two research objectives in this study:

1. To identify latent groups that can be typified from the viewpoint of the TPB regarding information-seeking behavior of survivors with cancer, trust in information sources, difficulties faced while seeking information, and use of psychosocial support services.
2. To examine how the identified latent groups are characterized from the perspectives of the relationship-oriented and task-oriented communication functions of the patient-provider relationship model and to clarify the contextual factors behind information-seeking behaviors and the use of support services.

Methods

Research Design

This cross-sectional survey was conducted on the basis of the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) [25] (Checklist 1).

The questionnaire, created in Japanese, consisted of 21 screens and 52 items on the website and was designed to be answerable within 25 minutes (Multimedia Appendix 1). To prevent missing data, all questions were set as mandatory in the web-based

XSL•FO
RenderX

questionnaire. To eliminate duplicate responses, any responses that matched IP address, sex, age, and place of residence were considered to be from the same respondent, and only the first response was included in the analysis.

Questions regarding patients with cancer information-seeking behavior were developed with reference to previous studies [26] that aimed to clarify the information needs and difficulties encountered when searching for information and using psychosocial support services among patients with ovarian cancer, as well as the patient experience survey [27] conducted by the National Cancer Center. During development, 3 researchers with experience in medical research and survey studies involving individuals with diseases examined whether each question appropriately reflected the research objectives and theoretical framework from the perspective of content validity and revised the questions for clarity of wording and to minimize respondent burden.

First, 12 items on sociodemographic information (including sex, age, place of residence, and occupation) and clinical information (including type of cancer, cancer stage, treatment status, and recurrence or metastasis) were collected using a single-choice format during the screening phase. This allowed for the collection of basic attribute information to uniquely classify each respondent. In addition, as a basic indicator related to information seeking, "device used to gather information" (1 item) was assessed using a single-choice format.

For information-seeking behavior regarding cancer, we started with the question "Have you ever searched for information about cancer?" Four multiple-choice checkbox items were used: "sources of cancer-related information," "social networking service (SNS) or apps used for collecting cancer-related information," and "advantages of psychosocial support and information-gathering services recommended by medical institutions." For these questions, mutually exclusive options such as "did not search," "do not wish to obtain information," "do not use SNS or apps," and "not applicable" were provided to ensure consistency in responses.

Furthermore, skip logic was introduced: if a respondent answered "did not search" to the question "Have you ever searched for information about cancer?" they were not required to answer questions about "sources of cancer-related information," "devices used to collect information," or "SNS or apps used for collecting cancer-related information," and were instead directed straight to the question about "difficulties in information seeking." Additionally, even if respondents had searched for information, if they selected either "do not wish to obtain information" or "do not use SNS or apps," they were not asked to answer the question about the "device used to gather information."

For the intention to use psychosocial support and information provision services recommended by medical institutions (corresponding to "behavioral intention" in the TPB), a yes or no question—"use of hospital-recommended apps and services for counseling support and information gathering"—was used (1 item). Only participants who answered "Yes" were able to proceed to the subsequent question on "advantages of psychosocial support and information-gathering services recommended by medical institutions."

"Difficulties in information seeking" consisted of 7 items, "trusted information sources" consisted of 8 items, and "evaluation of interpersonal relationships between patients and health care providers" consisted of 9 items. To reduce the respondent burden, these were presented in a matrix format and measured using a 5-point Likert scale.

To assess personality traits, the Japanese version of the 10-Item Personality Inventory (TIPI-J) [28], which uses 10 items and a 7-point Likert scale, was used. The TIPI-J measures the Big Five personality dimensions—extraversion, agreeableness, conscientiousness, neuroticism, and openness—with 2 items each and has shown high internal consistency: extraversion (α=.92), agreeableness (α=.85), conscientiousness (α=.82), neuroticism (α=.91), and openness (α=.86) [28]. The TIPI-J has also been shown to be associated with the emotional and cognitive health of patients with brain tumors [29], making it a valid indicator for this study.

To ensure accessibility and usability, a test page simulating the survey environment was created before the actual survey to confirm that the question display and skip logic functioned accurately.

Targeting and Survey Distribution

The survey was conducted privately by Asmarc Inc (a Japanese marketing research company) from December 10 to December 12, 2024. The company operates an online research panel with over 16 million registered members. After agreeing to the membership terms, registered members completed preliminary and final registration, during which they provided basic demographic information, such as age, sex, place of residence, employment status, and medical history.

The survey was conducted using the D-style web and Monitas platforms (Asmarc Inc) and was conducted using a list of available questionnaires. The eligibility criteria for this study were based on patient-reported outcome criteria for survivors with cancer 1 - 5 years after diagnosis [30]. Participants were required to be between 20 and 89 years of age at the time of the survey, to have been diagnosed with cancer at least 1 year prior, and to be either currently undergoing treatment or within 5 years of completing treatment. The following exclusion criteria were applied: participants who reported their treatment status as "before cancer treatment" or "other," those who had been diagnosed with cancer less than 1 year prior, and those for whom more than 10 years had passed since the end of treatment.

The sample size was calculated using cancer incidence statistics in Japan [31] as the reference population. Assuming a 90% CI and a 5% margin of error, the required sample size was estimated to be approximately 270 participants. This calculation was performed using the *pwr* package (version 1.3-0) in R. The *pwr* package was originally developed by Stéphane Champely and is currently maintained by Helios De Rosario. Considering a previous recommendation that at least 300 participants are desirable for conducting LCA [32], the minimum target sample size for this study was set at 300 participants.

XSL•FO
RenderX

Statistical Analysis

In this study, only data for participants with responses to all questions were included, and statistical analyses were performed using R (version 4.3.2; R Foundation for Statistical Computing). Sociodemographic information was stratified on the basis of the use of psychosocial support services, specifically by responses of "Yes" or "No" to "use of hospital-recommended apps and services for counseling support and information gathering." Descriptive statistics were calculated for each group. The sociodemographic information included sex, age, place of residence, occupation, and "device used to collect cancer information." Cancer-related information consisted of "cancer stage," "cancer type," "presence or absence of recurrence," "presence or absence of metastasis," "current treatment status," "time of cancer diagnosis," and "time of cancer treatment completion." Place of residence was recategorized into 6 categories: Hokkaido and Tohoku, Kanto, Chubu, Kinki, Chugoku and Shikoku, and Kyushu and Okinawa. To simplify the classification of cancer types, lung cancer and small cell lung cancer were grouped as lung cancer; esophageal, gastric, small intestine, and colorectal cancers were grouped as gastrointestinal cancer; liver, biliary tract, and pancreatic cancers were grouped as liver, biliary tract, and pancreatic cancer; kidney cancer and cancers of the renal pelvis and ureter and urethra, and bladder cancer were grouped as kidney, ureter, and bladder cancer; ovarian, endometrial, and cervical cancers were grouped as gynecologic cancer; leukemia, malignant lymphoma, and multiple myeloma were grouped as hematologic cancer; and brain tumors, head and neck cancer, thyroid cancer, cancer of unknown primary, skin cancer, and malignant melanoma were grouped as other. For information-seeking behaviors of patients with cancer, descriptive statistics were calculated regarding "having searched for information about cancer," "sources of cancer information," "SNS and apps used to gather cancer information," and "benefits of using psychosocial support and information-gathering services recommended by medical institutions." To understand the relational structure among items, analyses were conducted using tetrachoric correlations. Exclusive options such as "did not search" for "having searched for information about cancer," "did not want to obtain information" for "sources of cancer information," "do not use SNS or apps" for "SNS and apps used to gather cancer information," and "not applicable" for "benefits of using psychosocial support and information-gathering services recommended by medical institutions" were excluded from the statistical analyses. For the TIPI-J, descriptive statistics were calculated, and the relational structure among the items was examined using polychoric correlations.

Structural Validity of the Exploratory Factor Analysis

Exploratory factor analysis (EFA) was conducted according to the reporting standards proposed by Watkins [33]. EFA was chosen to identify the underlying factor structure. The measurement scales used were "behavioral barriers experienced when collecting information about cancer," "reliable information sources," and "relationships with doctors, medical professionals, and society." Each item was rated on a 5-point Likert scale ranging from 1 to 5. The appropriateness of conducting EFA was confirmed using the Bartlett test of sphericity [34]. The criteria for determining the number of factors included the Kaiser criterion of eigenvalues greater than 1 [35] and the minimum average partial (MAP) method [36]. Eigenvalues exceeded 1 up to the fifth factor (fifth factor=1.26), and the MAP exhibited its lowest value (0.03) at the fifth factor; therefore, a 5-factor solution was adopted. To choose the factor extraction method, multivariate normality was examined using the Mardia test [37]. Because skewness (skewness=62.14; P<.01) and kurtosis (kurtosis=613.96; P<.01) were significant, multivariate normality was rejected, and a method that did not assume normal distribution was required. Moreover, because the items used in this study were ordinal scales based on a 5-point Likert scale, factor analysis using the minimum residual method with a polychoric correlation matrix, which is suitable for ordinal data, was adopted [38,39]. Oblimin rotation [40] was used for simplification and theoretical convergence. Because oblique rotation was used, pattern coefficients of ≥0.37 were judged to be theoretically meaningful loadings, and coefficients of ≥0.40 were considered the primary criterion [41] for substantial interpretation. Cross-loadings of ≥0.3 were examined for interpretative caution. Internal consistency was determined using Cronbach α and McDonald ω, and factors with a reliability of ≥0.7 [42,43] and theoretical significance were judged as appropriate. The analysis was conducted using the *psych* package (version 2.6.2) in R. The *psych* package was developed and is maintained by William Revelle [44].

The LCA

In this study, we hypothesized that information-seeking behaviors and the use of psychosocial support services among survivors with cancer are not uniform and that multiple latent profiles exist. To identify these subgroups on the basis of statistical modeling, we used LCA [45]. Unlike traditional distance-based cluster analysis, LCA estimates latent structures using probabilistic models, allowing for objective model comparisons using information criteria, making it well suited for identifying latent profiles [46]. The analysis followed the reporting standards of Weller et al [47], and the optimal number of latent classes was determined on the basis of statistical fit indices and theoretical interpretability [48]. For the measurement variables, we recoded variables to ensure estimation stability, categorizing age, occupation, treatment status, and information-gathering devices. To examine the indicators included in the LCA, we conducted a polychoric correlation analysis and included 10 variables—sex, age, employment status, cancer type, stage, recurrence, metastasis, treatment status, time since completing treatment, and information-gathering devices—which were confirmed to have intervariable correlations of ±0.3 or higher. Model fit was assessed primarily using Bayesian information criterion (BIC), adjusted Bayesian information criterion (aBIC), and consistent Akaike information criterion (cAIC), and the model with the smallest value was selected as a candidate. Entropy was checked, with ≥0.8 considered acceptable and ≥0.9 considered ideal [48]. For the final model selection, we carefully evaluated whether each class had a sample size of n>50 or >5% of the total sample and whether each class profile was clinically meaningful. The R *poLCA* package (version 1.6.0.1), developed by Drew A Linzer and Jeffrey B Lewis, was used for the analysis [49].

XSL•FO
RenderX

Multiple Comparison Tests Between Identified Latent Classes

We examined the differences in each survey item among the latent classes. The variables compared included "counseling support and information gathering through apps and services recommended by the hospital," "survey items related to cancer," "sources of information about cancer," "SNS and apps used to gather cancer information," and "advantages of using psychosocial support and information-gathering services recommended by medical institutions," as well as the factors identified in the EFA and the TIPI-J. To assess statistical significance, we adopted nonparametric testing methods with a significance level of $P<.05$. For binary items, we used the chi-square test for independence, and for ordinal items, we used the Kruskal-Wallis test. For items where significant differences were observed, we conducted residual analysis for binary items and the Wilcoxon rank-sum test ($P=.02$) for ordinal items as post hoc analyses using the Bonferroni correction and evaluated the association between the latent classes and each survey item.

Ethical Considerations

This study was approved by the Research Safety and Ethics Committee of Tokyo Metropolitan University of Industrial Technology (approval no 23020). The survey was conducted in compliance with the Declaration of Helsinki and relevant ethical guidelines, and only those who provided informed consent participated in the study. The collected data were securely stored in a password-protected database that was accessible solely to the research team. Participants were compensated with reward points according to the regulations of each platform. The allocation and exchange of points were managed exclusively by the survey company, and the research team was not informed about individual point allocations. Participants who responded via the D-style web platform received 2 points for the preliminary survey and 15 points for the main survey, totaling 17 points upon completing both surveys. Participants who responded via the Monitas platform received 2 points for the preliminary survey and 9 points for the main survey, totaling 11 points upon completing both surveys.

Results

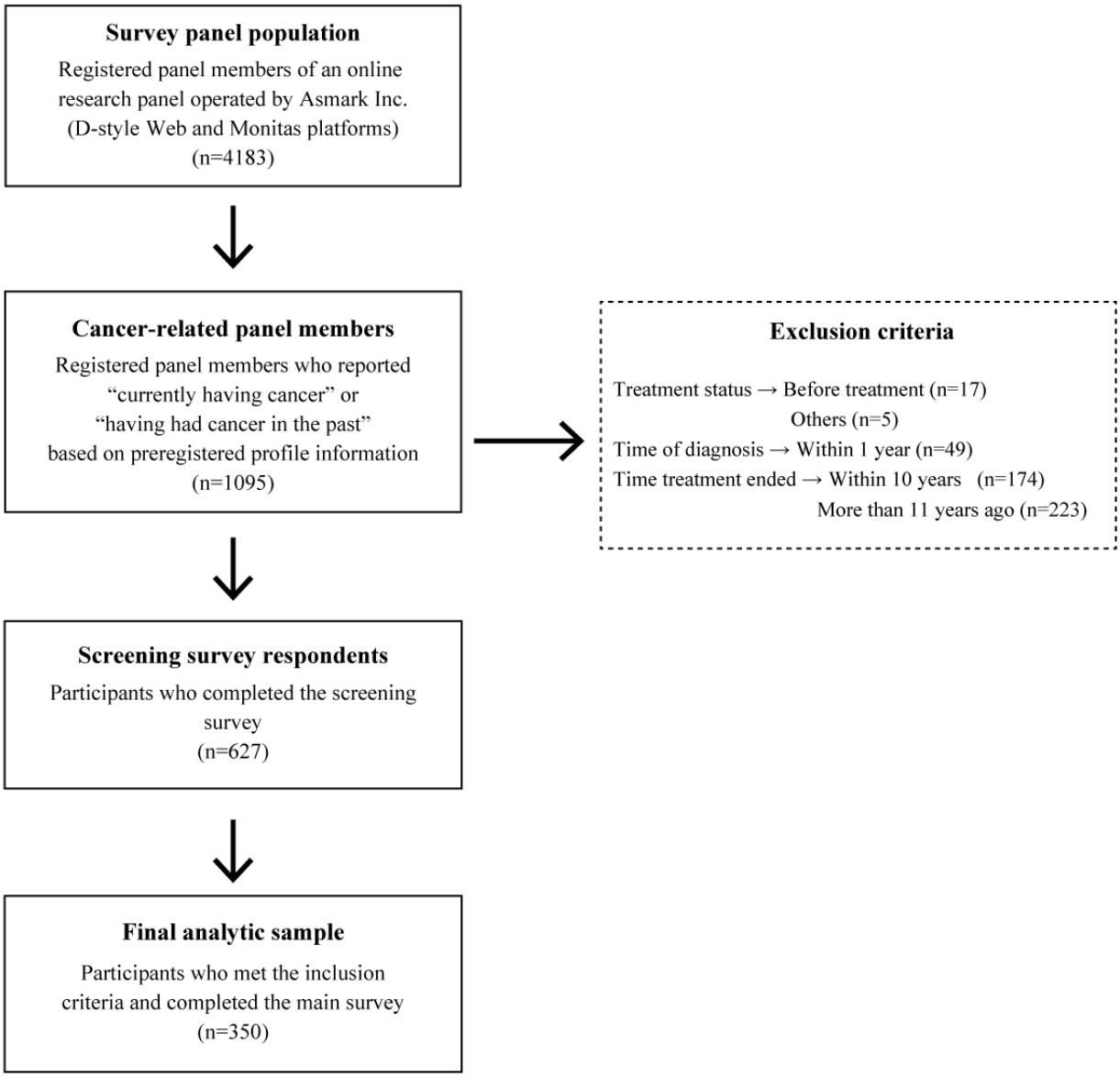
Participant Characteristics

Of the 4183 members registered on the D-style web and Monitas platforms, 1095 individuals who indicated in their preregistration information that they "currently have cancer or have had cancer in the past" were selected for the screening survey. Of these, 627 individuals responded (participation rate: 57.3%). Among the respondents, 350 met the eligibility criteria for the main survey and proceeded to participate (transition rate: 55.8%). Accordingly, the final analysis included 350 survivors with cancer (350/1095, 31.9%; Figure 1).

The survey participants comprised 175 men and 175 women (175/350, 50%). The average age was 57.9 (SD 10.8) years. The largest age group was 60 - 69 years (138/350, 39.4%), followed by 70 - 79 years (89/350, 25.4%). Regarding occupation, the most common response was unemployed (108/350, 30.9%), followed by company employees (84/350, 24%). Most participants resided in the Kanto region (156/350, 44.6%). Regarding clinical characteristics, the most common types of cancer were breast cancer (93/350, 26.6%) and gastrointestinal cancer (91/350, 26%). Many patients were at an early stage, with stage 1 accounting for 156 of 350 (44.6%) participants. Regarding treatment status, more than half had completed treatment and were undergoing regular follow-up visits at the hospital (202/350, 57.7%). Smartphones (193/350, 55.1%) were the most commonly used devices to gather information about cancer. In contrast, 105 of 350 (30%) participants reported that they did not search for information. Regarding the use of hospital-recommended apps and services for counseling support and information gathering, 282 of 350 (80.6%) participants answered "yes," and 68 of 350 (19.4%) participants answered "no." Among participants who reported that they did not search for information in response to the question about devices used to gather information about cancer (105/350, 30%), most also reported "yes" regarding the use of hospital-recommended apps and services for counseling support and information gathering (91/105, 86.7%), with only 14 of 105 (13.3%) reporting "no." Of those who answered "yes," 161 of 282 (57.1%) had completed treatment and were under regular follow-up, and 72 of 282 (25.5%) were receiving outpatient or home medical care. Smartphones were the most commonly used devices to gather information about cancer (161/282, 46%; Table 1).

XSL•FO
RenderX

Figure 1. Flowchart of participant selection process.



XSL•FO
RenderX

Table . Patient characteristics (N=350).

Category	Total, n (%)	Yes, n (%)	No, n (%)
Use of hospital-recommended apps and services for counseling support and information gathering			
Sex			
Male	175 (50)	141 (40.3)	34 (9.7)
Female	175 (50)	141 (40.3)	34 (9.7)
Age (years)			
Mean (SD)	57.9(10.8)		
20 - 29	3 (0.9)	3 (0.9)	0 (0)
30 - 39	7 (2)	6 (1.7)	1 (0.3)
40 - 49	32 (9.14)	28 (8)	4 (1.1)
50 - 59	76 (21.7)	56 (16)	20 (5.7)
60 - 69	138 (39.4)	114 (32.6)	24 (6.9)
70 - 79	89 (25.4)	71 (20.3)	18 (5.1)
80 - 89	5 (1.44)	4 (1.14)	1 (0.3)
Job			
Company employee	84 (24)	73 (20.9)	11 (3.1)
Civil servant	9 (2.6)	8 (2.3)	1 (0.3)
Self-employed	21 (6)	18 (5.1)	3 (0.9)
Company director	6 (1.7)	3 (0.9)	3 (0.9)
Freelance	6 (1.7)	4 (1.1)	2 (0.6)
Housewife and househusband	67 (19.1)	48 (13.7)	19 (5.4)
Part-time job	44 (12.6)	38 (10.9)	6 (1.7)
Unemployed	108 (30.9)	85 (24.3)	23 (6.6)
Other	5 (1.4)	5 (1.4)	0 (0)
Area			
Hokkaido and Tohoku	30 (8.6)	23 (6.6)	7 (2)
Kanto	156 (44.6)	129 (36.9)	27 (7.7)
Chubu	52 (14.9)	38 (10.9)	14 (4)
Kinki	68 (19.4)	53 (15.1)	15 (4.3)
Chugoku and Shikoku	22 (6.3)	19 (5.5)	2 (0.9)
Kyushu and Okinawa	22 (6.3)	20 (5.7)	2 (0.6)
Cancer type			
Lung cancer	29 (8.3)	24 (6.9)	5 (1.4)
Breast cancer	93 (26.6)	74 (21.1)	19 (5.5)
Gastrointestinal cancer	91 (26)	74 (21.1)	17 (4.9)
Liver, biliary tract, and pancreatic cancer	14 (4)	9 (2.6)	5 (1.4)
Kidney, ureter, and bladder cancer	17 (4.9)	14 (4)	3 (0.9)
Prostate cancer	39 (11.2)	31 (8.9)	8 (2.3)
Gynecological cancer	21 (6)	16 (4.6)	5 (1.4)
Blood cancer	30 (8.6)	27 (7.7)	3 (0.9)
Other	16 (4.6)	13 (3.7)	3 (0.9)
Stage			

XSL•FO
RenderX

Category	Total, n (%)	Yes, n (%)	No, n (%)
Use of hospital-recommended apps and services for counseling support and information gathering			
Stage 1	156 (44.6)	129 (36.9)	27 (7.7)
Stage 2	87 (24.9)	67 (19.1)	20 (5.7)
Stage 3	68 (19.4)	54 (15.4)	14 (4)
Stage 4	39 (11.1)	32 (9.1)	7 (2)
Cancer recurrence			
Yes	62 (17.7)	55 (15.7)	7 (2)
No	288 (82.3)	227 (64.9)	61 (17.4)
Cancer metastasis			
Yes	55 (15.7)	47 (13.4)	8 (2.3)
No	295 (84.3)	235 (67.1)	60 (17.1)
Current treatment status			
Inpatient treatment	5 (1.4)	5 (1.4)	0 (0)
Outpatient or home visit treatment	87 (24.9)	72 (20.6)	15 (4.3)
Completed treatment and under regular observation	202 (57.7)	161 (46)	41 (11.7)
Completed treatment, no hospital visits	56 (16)	44 (12.6)	12 (3.4)
When was the cancer diagnosed?			
Within 3 years	127 (36.3)	109 (31.1)	18 (5.1)
Within 5 years	85 (24.3)	61 (17.4)	24 (6.9)
Within 10 years	96 (27.4)	75 (21.4)	21 (6)
>11 years	42 (12)	37 (10.6)	5 (1.4)
Time when cancer treatment ended			
During treatment	58 (16.6)	49 (14.0)	9 (2.6)
Within 1 year	114 (32.6)	96 (27.4)	18 (5.1)
Within 3 years	86 (24.6)	60 (17.1)	26 (7.4)
Within 5 years	92 (26.3)	77 (22.0)	15 (4.3)
Devices used to gather information about cancer			
Smartphone	193 (55.1)	161 (46.0)	32 (9.1)
Personal computer	12 (3.4)	10 (2.9)	2 (0.6)
Tablet (eg, iPad)	8 (2.3)	6 (1.7)	2 (0.6)
Other	32 (9.1)	14 (4)	8 (5.1)
Not searching for information	105 (30)	91 (26)	14 (4)

Information-Seeking Behavior of Patients With Cancer

Topics Researched Regarding Cancer

The most frequently researched topic related to cancer was “my disease,” which accounted for 75.7% (265/350) of responses. This was followed by “about treatment” (68.6%, 240/350) and “progress and prognosis” (55.4%, 194/350). “Treatment costs” accounted for 38.9% (136/350) of responses, “life during treatment” for 36% (126/350), and “other patients’ treatment and daily life” for 27.1% (95/350). Regarding correlations between categories, “life during treatment” had a strong positive correlation with “about my own condition” (ρ=0.7) and “progress and prognosis” (ρ=0.64). “Other patients’ treatment and daily life” showed a moderate positive correlation with “about my own condition” (ρ=0.61).

XSL•FO
RenderX

Sources of Cancer Information and Use of SNS and Apps

The most common source of cancer-related information was "internet/SNS" (58.9%, 206/350), followed by "doctors, nurses, and other medical staff" (50.6%, 177/350). Other sources included "newspapers, magazines, and books" (17.7%, 62/350), "family, friends, and acquaintances" (15.7%, 55/350), "National Cancer Center website" (15.4%, 54/350), and "cancer counseling support center" (7.4%, 26/350). Regarding correlations between categories, a moderately strong correlation (ρ=0.69) was observed between "newspapers, magazines, and books" and "television and radio," Among the SNS and apps used, YouTube (Google LLC) was the most common (13.4%, 47/350), followed by blogs (10%, 35/350), LINE, provided by LY Corporation (7.4%, 26/350), and X (formerly Twitter, provided by X Corp; 6%, 21/350). These items exhibited positive correlations with each other (ρ=0.48-0.6).

Benefits of Using Psychosocial Support and Information-Gathering Services Recommended by Health Care Institutions

The most commonly identified benefits of psychosocial support services or apps were "being able to obtain reliable information" (59.7%, 209/350) and "deepening knowledge about illness and treatment" (58.6%, 205/350), with more than half of the respondents emphasizing the quality and understanding of information. "Receiving information and support tailored to one's own symptoms or situation" was cited by 45.4% (159/350) of patients, and "facilitating communication with doctors or medical staff" was cited by 29.4% (103/350) of patients. Regarding correlations between items, a strong positive correlation (ρ=0.71) was found between "deepening knowledge about illness and treatment" and "receiving information suited to one's own symptoms or situation." "Easy to operate and user-friendly" showed a moderate positive correlation with both "deepening knowledge about illness and treatment" (ρ=0.61) and "facilitating communication with doctors or medical staff" (ρ=0.6; Table 2 and Multimedia Appendix 2).

XSL•FO
RenderX

Table . Trends in information-seeking behavior among patients with cancer.

Item	Value, n (%)
Things I have researched about cancer	
About treatment	240 (68.6)
My disease	265 (75.7)
Course or prognosis	194 (55.4)
Information about other patients	95 (27.1)
Life during treatment	126 (36)
Cost of treatment	136 (38.9)
Other	20 (5.7)
Sources of cancer information	
Cancer Consultation Support Center	26 (7.4)
Doctors, nurses, and so on	177 (50.6)
Contact at public health center or health center	5 (1.4)
Newspapers, magazines, and books	62 (17.7)
Television and radio	40 (11.4)
National Cancer Center website	54 (15.4)
Internet and SNS[a]	206 (58.9)
Family, friends, and acquaintances	55 (15.7)
Other	6 (1.7)
SNS and apps used to gather cancer information	
Facebook (Meta Platforms, Inc)	14 (4)
Instagram (Meta Platforms, Inc)	15 (4.3)
LINE (LY Corporation)	26 (7.4)
X (formerly Twitter; X Corp)	21 (6)
YouTube (Google LLC)	47 (13.4)
Blog	35 (10)
Other	46 (13.1)
Benefits of using psychosocial support and information-gathering services recommended by health care institutions	
Access to reliable information	209 (59.7)
Easy to operate and use	76 (21.7)
Access to the latest medical information	152 (43.4)
Easy to understand correct usage and operation	78 (22.3)
Personal information is kept safe	58 (16.6)
Facilitate communication with doctors and medical staff	103 (29.4)
Gain knowledge about illnesses and treatments	205 (58.6)
Get information and support tailored to your symptoms and situation	159 (45.4)

[a]SNS: social networking services.

TIPI-J

The mean scores reported by respondents were highest for neuroticism at 8.62 (SD 2.47), followed by openness at 8.24 (SD 2.43), extraversion at 7.65 (SD 2.6), and conscientiousness at 7.15 (SD 2.29). Agreeableness exhibited the lowest value among the 5 traits, at 5.84 (SD 2.03). A Shapiro-Wilk test [50] was conducted to assess the distribution of each scale, and the results were significant for extraversion (W=0.98; P<.001), agreeableness (W=0.95; P<.001), conscientiousness (W=0.98; P<.001), neuroticism (W=0.98; P<.001), and openness (W=0.98; P<.001), indicating that the data were not normally distributed. The correlations among the subscales were uniformly low (r=−0.43 to 0.37; Multimedia Appendix 2).

XSL•FO
RenderX

EFA Results

The Bartlett test of sphericity was conducted to assess the applicability of factor analysis. On the basis of the results, the null hypothesis that the correlation matrix is an identity matrix was rejected (χ^2_{210}=5425.91; P<.001), indicating that there was a sufficiently robust correlation structure among the items. The Kaiser-Meyer-Olkin statistic was 0.81, confirming sampling adequacy and determining that factor analysis was appropriate. The number of factors was determined on the basis of eigenvalues, MAP, and theoretical consistency. Eigenvalues exceeded 1 up to the fifth factor (factor 1=6.33, factor 2=4.27, factor 3=2.56, factor 4=1.86, and factor 5=1.26), and the cumulative contribution of the 5-factor solution was 67.83%. MAP decreased with the addition of more factors and improved up to the fifth factor. In addition to these results, considering that the measurement included several conceptual domains such as "difficulties in information seeking," "support and consultation for care," "relationship with the physician," "trust in information from medical institutions and specialized websites," and "trust in information from outside medical institutions," a 5-factor solution was adopted.

Factor 1 included items reflecting difficulties encountered in the process of seeking cancer-related information (factor loadings: 0.54-0.9) and was named "difficulties in information seeking." Factor 2 included items related to continued employment before treatment, concerns about appearance, convalescence, precautions in daily life, and workplace considerations (factor loadings: 0.42-0.74) and was interpreted as "evaluation of support for cancer care." Factor 3 included items on the primary physician's knowledge and experience, ease of consultation, and trust in information provided by physicians (factor loadings: 0.67-0.81) and was named "evaluation of the relationship with the physician." Factor 4 included trust in information from cancer-specialized or hospital or pharmaceutical company websites (factor loadings: 0.73-0.85) and was named "evaluation of the reliability of information from medical institutions and cancer-related websites." Factor 5 included trust in sources of information outside medical institutions, such as family and acquaintances, other patients, and social networking sites (factor loadings: 0.47-0.76) and was named "assessment of the reliability of information from nonmedical institutions."

The item "I had to visit several websites" loaded on both Factor 1 (0.60) and Factor 4 (0.33), and "the most reliable source is information from the physician" loaded on both Factor 3 (0.67) and Factor 4 (0.32). The internal consistency of the factors ranged from α=.75 to .9 and ω=0.78 to 0.92, with small interfactor correlations, the highest being between Factors 2 and 3 (r=0.47), justifying the treatment of the 5 factors as independent subconcepts (Table 3).

XSL•FO
RenderX

Table . Factor structure of information-seeking behavior and evaluation of information sources and psychosocial support services among survivors with cancer.

Item	Mean (SD)	Skewness	Kurtosis	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	H^2
Factor 1: difficulties in information seeking									
I didn't know where to look for it	2.88 (1.18)	0.08	−1.01	0.9	0.1	0	−0.15	0.01	0.82
I didn't know who to ask	2.81 (1.18)	0.15	−1.03	0.88	0.04	−0.04	−0.09	0.05	0.8
I didn't know what to look for	2.92 (1.17)	0.13	−0.96	0.88	0.05	0.03	−0.23	0.04	0.81
I couldn't find the information I wanted right away	3.11 (1.14)	−0.16	−0.9	0.77	−0.01	0	0.12	−0.12	0.61
I didn't know which information I could trust	3.04 (1.08)	−0.07	−0.8	0.71	−0.09	0.07	−0.01	0.05	0.51
I had to go to several websites	3.31 (1.09)	−0.43	−0.64	0.6	0	−0.04	0.33	−0.03	0.47
I couldn't ask the doctor a question	2.25 (1.12)	0.73	−0.28	0.54	−0.1	−0.39	−0.01	0.06	0.57
Factor 2: support and consultation for care									
I think a medical professional spoke to me about continuing to work before starting treatment	3.15 (1.1)	−0.2	−0.54	−0.02	0.74	−0.06	0.1	0.02	0.55
I felt I could talk to a medical professional about my concerns about changes in appearance	3.07 (1)	−0.1	−0.31	0	0.73	−0.06	0.03	0.04	0.51
I think I was able to talk to someone about my illness and medical treatment	3.27 (0.97)	−0.24	−0.4	0.05	0.71	0.13	−0.03	−0.01	0.59

XSL•FO
RenderX

Item	Mean (SD)	Skewness	Kurtosis	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	H^2
I think there were medical professionals other than my primary care physician whom I felt comfortable talking to.	3.58 (0.93)	−0.54	0.14	−0.04	0.62	0.29	0.06	−0.06	0.66
I think my workplace and work colleagues have made arrangements to allow me to continue both my treatment and work	3.02 (1.13)	−0.1	−0.55	−0.01	0.5	−0.09	0.08	0.08	0.26
I think I was able to get information from medical professionals about important points to keep in mind in daily life	3.38 (1.01)	−0.23	−0.66	0.03	0.48	0.17	−0.1	−0.09	0.31
I believe information about the patient's treatment was shared among the medical professionals involved in the treatment	3.64 (0.86)	−0.34	−0.11	0.02	0.42	0.41	0.03	−0.02	0.51
Factor 3: relationship with the physician									
The primary physician has sufficient knowledge and experience about the patient's cancer.	4.11 (0.88)	−0.9	0.43	−0.07	0.08	0.81	0.08	0.01	0.79
I think the primary physician who treated my cancer was easy to talk to.	3.97 (1)	−1.03	0.71	−0.05	0.16	0.8	−0.09	0.06	0.77

XSL•FO
RenderX

Item	Mean (SD)	Skewness	Kurtosis	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	H^2
A reliable source of information is a doctor's information.	4.21 (0.77)	−0.94	1.26	0.03	0	0.67	0.32	0	0.64
Factor 4: trust in information from medical institutions and specialized websites									
Reliable sources of information include information from cancer-specific websites such as cancer information services.	3.68 (0.75)	−0.54	0.95	0.11	0.11	0.03	0.85	−0.01	0.79
Reliable sources of information include information from hospital websites.	3.63 (0.74)	−0.42	0.61	0.03	0.06	0.26	0.75	0.06	0.76
Reliable sources of information include information from pharmaceutical company websites.	3.29 (0.74)	−0.18	0.71	−0.04	0.13	0.01	0.73	0.09	0.63
Factor 5: trust in information from outside medical institutions									
Reliable sources of information include exchanging information with family and acquaintances.	2.99 (0.74)	−0.1	0.76	0	0	0.2	−0.12	0.76	0.57
Reliable sources of information include information shared among patients.	3.08 (0.76)	0.02	0.72	0.06	−0.09	−0.01	0.18	0.74	0.62
Reliable sources of information include information from lectures.	3.04 (0.71)	−0.34	1.44	0.05	0	0.01	0.36	0.61	0.58

XSL•FO
RenderX

Item	Mean (SD)	Skewness	Kurtosis	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	H^2
Reliable sources of information included social media (Facebook and Instagram [Meta Platforms, Inc], X [formerly Twitter; X Corp], and YouTube [Google LLC]).	2.6 (0.78)	−0.45	−0.02	−0.04	0.1	−0.12	−0.09	0.47	0.25
Factor contribution	—[a]	—	—	4.32	4.26	4.03	3.09	2.03	—
Internal validity									
Crombach α	—	—	—	0.9	0.84	0.87	0.88	0.75	—
Macdonald ω	—	—	—	0.92	0.86	0.89	0.89	0.78	—
Factor correlation									
Factor 1	—	—	—	—	—	—	—	—	—
Factor 2	—	—	—	−0.08	—	—	—	—	—
Factor 3	—	—	—	−0.16	0.47	—	—	—	—
Factor 4	—	—	—	0.01	0.21	0.21	—	—	—
Factor 5	—	—	—	0.04	0.16	−0.07	0.18	—	—

[a]Not available.

LCA Results

In determining the number of latent classes, models with 2-6 classes were estimated, and BIC was used as the primary index with comparisons to aBIC and cAIC. The 3-class model had the lowest BIC (5734.74) and cAIC (5805.74) and demonstrated high classification accuracy with an entropy of 0.95. Although the 4-class model had the lowest aBIC (5465) and similarly high entropy (0.96), both BIC and cAIC worsened (BIC=5766.37 and cAIC=5861.37). For models with 5-6 classes, the information criteria further increased, suggesting model overfitting. In finalizing the model, it was confirmed that each class had a sample size of n>50 or >5% of the total and that the class profiles were clinically interpretable. On the basis of minimum BIC and cAIC values and high classification accuracy, the 3-class model was selected as the final model (Table 4).

Table . Model selection criteria.

Model	Log-likelihood	Residual df	BIC[a]	aBIC[b]	cAIC[c]	Likelihood ratio	Entropy
Model 1	—[d]	—	—	—	—	—	—
Model 2	−2778.75	303	5832.82	5683.72	5879.82	1697.5	1
Model 3	−2659.41	279	5734.74	5509.5	5805.74	1458.82	0.95
Model 4	−2604.93	255	5766.37	5465	5861.37	1349.87	0.96
Model 5	−2582.65	231	5862.39	5484.88	5981.39	1305.3	0.94
Model 6	−2563.22	207	5964.12	5510.47	6107.12	1266.43	0.94

[a]BIC: Bayesian information criterion.

[b]aBIC: adjusted Bayesian information criterion.

[c]cAIC: consistent Akaike information criterion.

[d]Not available.

XSL•FO
RenderX

After fitting the model, the results for the 3 classes identified by the LCA were as follows. Class 1 comprised 95 individuals, all of whom were female (95/95, 100%). Most patients in Class 1 were between the ages of 50 and 69 years, with 35 of 95 (36.8%) aged 50 - 59 years and 31 of 95 (32.6%) aged 60 - 69 years. Regarding treatment status, 74 of 95 (77.9%) patients were undergoing regular follow-up after completing treatment, while 21 of 95 (22.1%) were in the follow-up stage without needing to visit a medical facility. On the basis of these treatment characteristics, this class was categorized as the "group of women under observation." Breast cancer was the most common cancer type (52/95, 54.7%), followed by gynecological cancers (19/95, 20%). In terms of cancer stage, more than half of the patients were classified as stage 1 (56/95, 58.9%). Of the 95 patients, recurrence was observed in 8 (8.4%) and metastasis in 4 (4.2%). Regarding employment status, 51 of 95 (53.7%) patients were employed. For information-seeking behavior, 43 of 95 (45.3%) patients answered, "I do not use devices to gather information about cancer." Among those who gathered information, smartphones were used by 39 of 95 (41.1%) patients; however, overall, information-seeking behavior was limited compared with that of patients in the other classes.

Class 2 comprised 163 individuals, the majority of whom were male (141/163, 86.5%). Most patients in Class 2 were aged 60 years or older, with 144 of 163 (88.3%) patients in this range; 76 of 163 (46.6%) patients were aged 60 - 69 years, and 68 of 163 (41.7%) patients were ≥70 years. Regarding treatment status, 128 of 163 (78.5%) patients were under regular observation after completing treatment, while 35 of 163 (21.5%) were in the follow-up stage without requiring hospital visits. These features led to this class being categorized as the "group of men under observation." The most common cancer was gastrointestinal cancer (69/163, 42.3%), followed by prostate cancer (29/163, 17.8%). Most cases were classified as early stage, with 76 of 163 (46.6%) patients at stage 1 and 37 of 163 (22.7%) patients at stage 2. Recurrence occurred in 21 of 163 (12.9%) patients, and metastasis occurred in 14 of 163 (8.6%) patients. These proportions were higher than those in Class 1 but lower than those in Class 3. Regarding employment, 92 of 163 (56.4%) patients were not employed. The device most commonly used for information seeking was a smartphone (104/163, 63.8%), and 31 of 163 (19%) patients answered "I do not use devices to gather information about cancer"; this proportion was lower than that among participants in the other classes.

Class 3 comprised 92 individuals, all of whom were undergoing treatment (92/92, 100%). There were 59 (64.1%) females and 33 (35.9%) males, with the majority aged between 50 and 69 years (50 - 59 years: 25/92, 27.2% and 60 - 69 years: 31/92, 33.7%). Breast cancer was the most common cancer (41/92, 44.6%), followed by gastrointestinal (10/92, 10.9%) and prostate cancers (10/92, 10.9%). This class had a higher proportion of advanced cancers, with stage 4 accounting for 25 of 92 (27.2%) patients. Recurrence was observed in 33 of 92 (35.9%) patients and metastasis in 37 of 92 (40.2%) patients. Both proportions were higher in Class 3 than in the other 2 classes. Regarding information-seeking behavior, 50 of 92 (54.4%) patients used smartphones, but 31 of 92 (33.7%) patients answered "I do not use devices to gather information about cancer." Given these features, this class was defined as the "group under treatment" (Figure 2 and Multimedia Appendix 2).

XSL•FO
RenderX

Figure 2. Class-specific conditional probabilities of sociodemographic and clinical characteristics identified by latent class analysis.



Results of the Comparison Among the 3 Groups

All results were obtained from post hoc analyses. Regarding "use of hospital-recommended apps and services for counseling support and information gathering," no significant differences were found. As shown in Figure 3, regarding "things I've researched about cancer," the results revealed that "information about other patients" was significantly lower in Class 2 than in the other classes (χ^2_2 =12.22; *P*=.002, post hoc test: Class 1 > Class 2; *P*=.009, and Class 2 < Class 3; *P*=.001). Regarding "life during treatment" (χ^2_2 =7.22; *P*=.03 and post hoc test: Class 2 < Class 3; *P*=.01), Class 3 was significantly higher than Class 2. Regarding "sources of cancer information," "Cancer Consultation Support Center" (χ^2_2 =17.31; *P*<.001 and post hoc test: Class 2 < Class 3; *P*<.001) and "National Cancer Center website" (χ^2_2 =7.59; *P*=.02 and post hoc test: Class 2 < Class 3; *P*=.01) were significantly higher in Class 3 compared with Class 2. Regarding "SNS and apps used to gather cancer information," "Instagram" (χ^2_2 =12.57; *P*=.002 and post hoc test: Class 2< Class 3; *P*=.001) and "X (Twitter)" (χ^2_2 =9.55; *P*=.008 and post hoc test: Class 2 < Class 3; *P*=.002) were significantly higher in Class 3 compared with Class 2. Regarding "blog" (χ^2_2 =5.82; *P*=.054), no significant differences were found between classes after the post hoc test. For "benefits of using psychosocial support and information-gathering services recommended by health care institutions," regarding "facilitate communication with doctors and medical staff" (χ^2_2 =12.57; *P*=.02 and post hoc test: Class 1 < Class 2; *P*=.006), Class 2 was significantly higher than Class 1 (Figure 3 and Multimedia Appendix 2).

According to Figure 4, for "Factor 1: difficulties in information seeking" (χ^2_2 =11.9; *P*=.003 and post hoc test: Class 2 < Class 3; *P*=.007), Class 3 showed a significantly higher value than

XSL•FO
RenderX

Class 2. For "Factor 3: evaluation of the relationship with the physician" (χ^2_2 =12.2; *P*=.002 and post hoc test: Class 2 > Class 3; *P*=.002), Class 2 had a significantly higher value than Class 3. For "Factor 5: assessment of reliability of information from nonmedical institutions" (χ^2_2 =12.3; *P*=.002 and post hoc test: Class 1 > Class 2; *P*=.002), Class 1 showed a significantly higher score than Class 2. Regarding the TIPI-J, agreeableness (χ^2_2 =6.7; *P*=.04) and neuroticism (χ^2_2 =8.8; *P*=.01) did not show significant differences after the post hoc test (Figure 4 and Multimedia Appendix 2).

Figure 3. Results of multiple comparison analysis using the independent chi-square and post hoc tests.

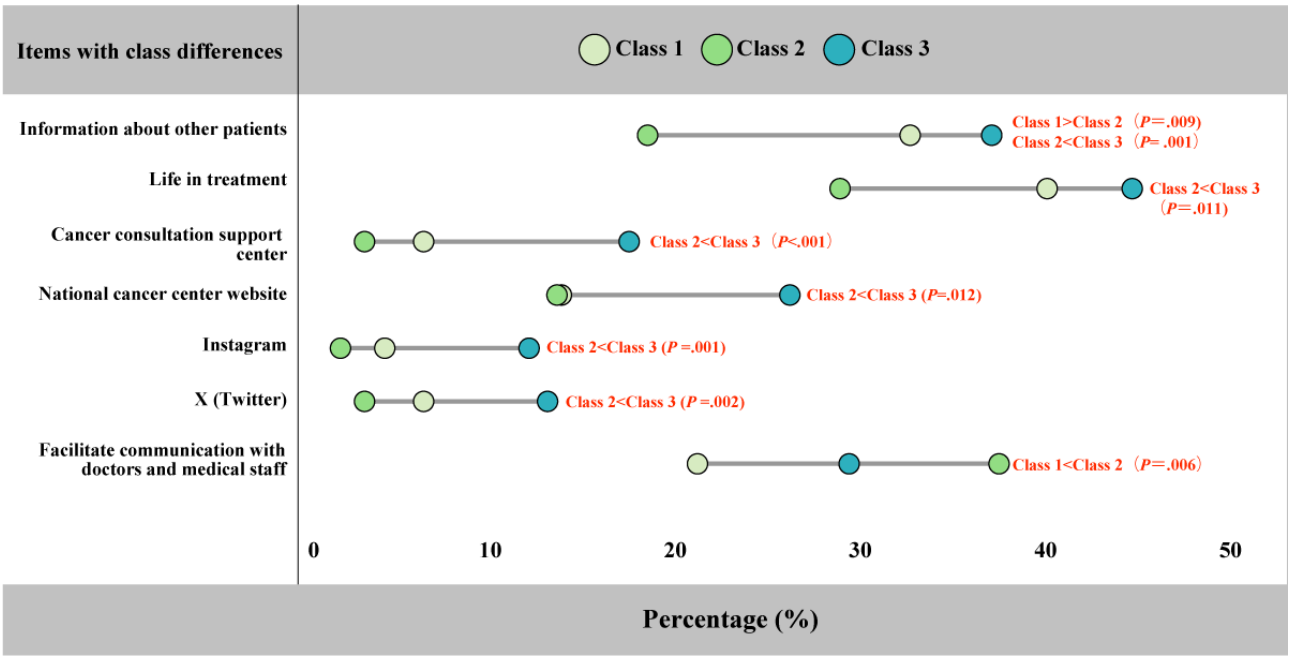
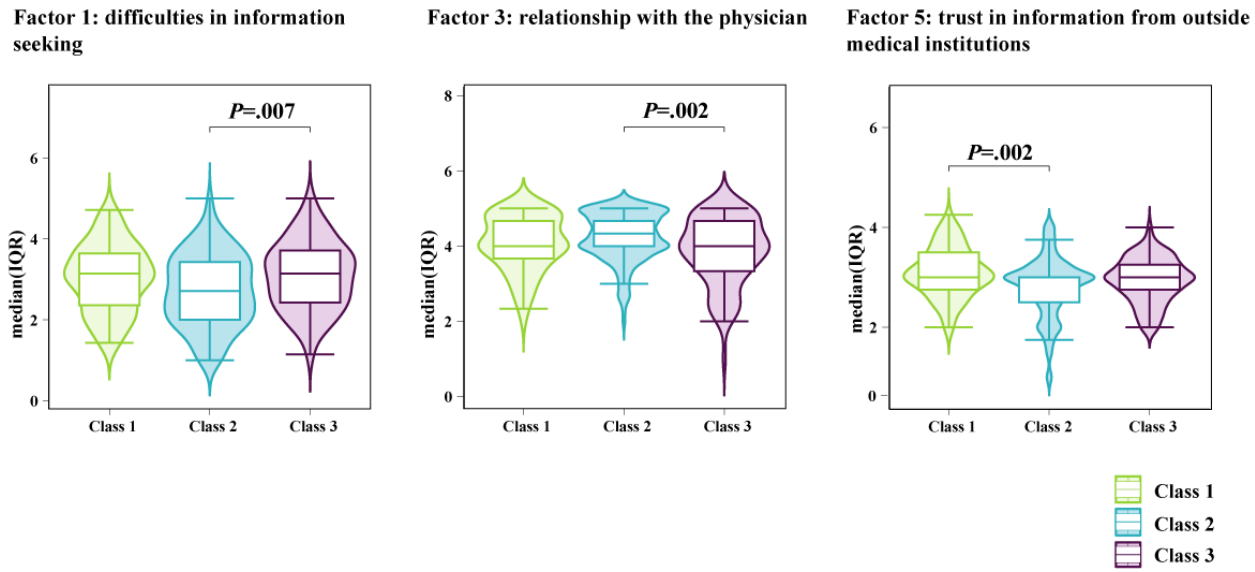


Figure 4. Results of multiple comparison analysis using the Kruskal-Wallis test and the Wilcoxon rank-sum test.



Discussion

Principal Findings

The results revealed that age distribution and cancer type composition among the respondents in this study were generally consistent with trends in Japan's national cancer statistics [31]. In addition, more than 80% of respondents indicated that they would like to use psychosocial support services or apps if recommended by medical institutions. Furthermore, in the results of the LCA, the "female under observation" group showed a high level of trust in "Factor 5: assessment of the reliability of information from nonmedical institutions," whereas the "male under observation" group was less proactive in gathering information but placed a high value on "Factor 3: evaluation of the relationship with the physician" and "facilitate communication with doctors and medical staff," indicating that they prioritize relationships with health care professionals. The "under treatment" group was the most active in collecting

XSL•FO
RenderX

information, frequently using "information about other patients," "life during treatment," "Cancer Consultation Support Center," and "National Cancer Center website." However, the "Factor 1: difficulties in information seeking" score was also high, suggesting that treatment status and gender may influence barriers to gathering information and differences in trusted sources.

Comparison With Previous Work

This study suggests that integrating the 3 factors from the TPB (ie, "attitude," "subjective norms," and "perceived behavioral control") with the "relationship function" and "task-oriented function" from the patient-provider relationship model is associated with a stronger intention to use psychosocial support services and hospital-recommended apps. Within the TPB framework, hospital recommendations may consistently shape behavioral intentions as a form of subjective norm. Previous studies have also reported that recommendations from health care providers strongly influence information-seeking behavior and support service usage among patients with cancer [51], and the results of this study are consistent with these findings.

Looking at each group, the male group under follow-up showed a high evaluation of their relationship with physicians and a strong trust in physician-provided information, with subjective norms in the TPB being particularly dependent on physicians. Previous research has also shown that male patients with cancer place a high value on trustworthiness and sincerity in their interactions with health care providers [52]. Furthermore, in the male follow-up group, the relationship function based on trust in physicians appears to play a central role, possibly serving as a prerequisite for accepting psychosocial support services and information provision.

In contrast, the female group under follow-up showed relatively higher trust in "information from other patients" and nonmedical sources, suggesting that peer-derived experiential knowledge plays an important role in forming attitudes in the TPB. Earlier studies, particularly those focusing on patients with breast cancer, have reported that peer support contributes to acquiring information and alleviating anxiety about recurrence [53-55], which is consistent with the current findings. For this group, it is also possible that task-oriented functions, such as managing uncertainty and decision support, are supplemented outside health care institutions. By providing opportunities for exchange with other patients or sharing experiential information, health care providers may help translate behavioral intentions into action.

The "under treatment" group displayed the highest level of active information-seeking behavior among the 3 groups, but also the greatest difficulty in seeking information. In terms of the TPB framework, while "behavioral intention" to seek information or use psychosocial services was high, "perceived behavioral control" appeared to be low. From the perspective of the patient-provider relationship model, it was suggested that the task-oriented functions for exchanging information and managing uncertainty, and the relationship functions providing reassurance and emotional support, may not be sufficiently aligned with patients' needs. Survivors with cancer undergoing treatment have strong information needs and clear intentions to seek information but may lack sufficient self-efficacy to understand the information and make decisions, leading to increased difficulty. Prior studies have also reported that patients with cancer in advanced or treatment phases experience heightened anxiety from online information [56], lack of social support beyond the internet, and negative treatment experiences [57]. Such negative experiences can decrease patients with cancer's self-efficacy and trust in their digital skills, making information seeking a heavy burden. Therefore, for survivors with cancer undergoing treatment, clarifying reliable information sources and having health care providers organize and supplement information is crucial for alleviating the psychological burden associated with information-seeking.

Strengths and Limitations

The strength of this study lies in its analysis of information-seeking behaviors of survivors with cancer and usage of psychosocial support services, grounded in the 3 components of TPB, as well as the communication functions of relationship-oriented and task-oriented functions in the patient-provider relationship model. By using EFA to identify factors and subsequently using LCA to classify multiple subgroups, this study systematically captured the behavioral intentions underlying information-seeking activities. This approach identified diverse needs among subgroups—needs that are often difficult to capture by a simple comparison of attributes—regarding trusted sources of information and communication tendencies among survivors with cancer. In particular, differences in the levels of trust in information sources between men and women during the follow-up period highlight the importance of personalized information provision, offering practical insights for future patient support service design. However, this study involved several limitations that should be considered. First, although the target sample size was achieved and trends were found to partially align with national cancer incidence statistics, the sampling was not stratified by age or cancer stage. Consequently, there was a bias toward survivors in stages 1 and 2, whereas the proportion of survivors in stages 3 and 4, or those who experienced recurrence or metastasis, was relatively low. Given the imbalances in cancer type, age group, and stage, caution is warranted regarding the generalization of the study findings to all survivors with cancer. Second, this was a cross-sectional study aimed primarily at identifying tendencies in types of information-seeking behaviors and thus cannot establish causal relationships related to various factors and backgrounds associated with each survey item. Detailed factor analyses and longitudinal studies are necessary to elucidate how information-seeking behaviors and perceptions evolve over time. Third, because this study focused on differences by treatment status and gender, socioeconomic factors such as income, education, and digital literacy were not included in the survey. However, these factors may also influence information-seeking behaviors and the use of support services, suggesting that future research should comprehensively examine these variables. Fourth, the questionnaire used in this study was developed on the basis of previous research and theoretical frameworks. However, because no formal pilot testing or test-retest reliability verification was conducted before the survey, the temporal stability of the measurement could not

XSL•FO
RenderX

be assessed. This represents a limitation, and caution is required when interpreting the results. In the future, a comprehensive psychometric evaluation, including test-retest reliability and measurement invariance, will be necessary.

Conclusions

This study clarified how information-seeking behaviors among survivors with cancer differ depending on treatment status and sex. By applying the TPB and patient-provider relationship models and using EFA, LCA, and multiple comparison analyses, we identified subgroups of patients with cancer with different behavioral characteristics. Our findings provide valuable insights for designing necessary approaches to promote personalized psychosocial support services and information provision for patients, as well as for improving the user interfaces of products and services that support patients psychologically and socially. Future research should focus on the realities of information provision and communication between patients with cancer and health care professionals, as they relate to individualized support for middle-aged and older survivors with cancer. Further exploration from multifaceted perspectives, including health care providers' recognition of information delivery and patients' acceptance of information, is expected to lead to the development and improvement of personalized and patient-centered communication approaches in the future.

Acknowledgments

We thank Benjamin Knight and Sarina Iwabuchi from Edanz for editing a draft of this paper. The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT taxonomy, the following tasks were delegated to GAI tools under full human supervision: translation and quality assessment. The GAI tool used was Paperpal. Responsibility for the final paper lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the outcomes. The declaration was submitted by the authors.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this paper.

Data Availability

All data generated or analyzed during this study are included in this published paper and Multimedia Appendix 2.

Authors' Contributions

Conceptualization, investigation, formal analysis, visualization, and paper preparation were performed by MH. DK and KT were responsible for reviewing and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire survey form.
[DOCX File, 1589 KB - cancer_v12i1e79065_app1.docx]

Multimedia Appendix 2
Datasets generated and analyzed during the study.
[XLSX File, 64 KB - cancer_v12i1e79065_app2.xlsx]

Checklist 1
CHERRIES checklist.
[DOCX File, 23 KB - cancer_v12i1e79065_app3.docx]

References

1. Data visualization tools for exploring the global cancer burden in 2022. Global Cancer Observatory: Cancer Today. URL: https://gco.iarc.who.int/today [accessed 2025-09-17]
2. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]

XSL•FO
RenderX

3. Nguyen PT, Hori M, Matsuda T, Katanoda K. Cancer prevalence projections in Japan and decomposition analysis of changes in Cancer burden, 2020-2050: a statistical modeling study. Cancer Epidemiol Biomarkers Prev 2023 Dec 1;32(12):1756-1770. [doi: 10.1158/1055-9965.EPI-23-0754] [Medline: 37756571]
4. Mullan F. Seasons of survival: reflections of a physician with cancer. N Engl J Med 1985 Jul 25;313(4):270-273. [doi: 10.1056/NEJM198507253130421] [Medline: 4010738]
5. Schouten B, Avau B, Bekkering GTE, et al. Systematic screening and assessment of psychosocial well-being and care needs of people with cancer. Cochrane Database Syst Rev 2019 Mar 26;3(3):CD012387. [doi: 10.1002/14651858.CD012387.pub2] [Medline: 30909317]
6. Goerling U, Faller H, Hornemann B, et al. Information needs in cancer patients across the disease trajectory. A prospective study. Patient Educ Couns 2020 Jan;103(1):120-126. [doi: 10.1016/j.pec.2019.08.011] [Medline: 31474389]
7. Hack TF, Ruether JD, Pickles T, Bultz BD, Chateau D, Degner LF. Behind closed doors II: systematic analysis of prostate cancer patients' primary treatment consultations with radiation oncologists and predictors of satisfaction with communication. Psychooncology 2012 Aug;21(8):809-817. [doi: 10.1002/pon.1984] [Medline: 21557385]
8. Ha JF, Longnecker N. Doctor-patient communication: a review. Ochsner J 2010;10(1):38-43. [Medline: 21603354]
9. Chen X, Siu LL. Impact of the media and the internet on oncology: survey of cancer patients and oncologists in Canada. J Clin Oncol 2001 Dec 1;19(23):4291-4297. [doi: 10.1200/JCO.2001.19.23.4291] [Medline: 11731511]
10. Roach AR, Lykins ELB, Gochett CG, Brechting EH, Graue LO, Andrykowski MA. Differences in cancer information-seeking behavior, preferences, and awareness between cancer survivors and healthy controls: a national, population-based survey. J Cancer Educ 2009;24(1):73-79. [doi: 10.1080/08858190802664784] [Medline: 19259869]
11. Osowiecka K, Rucińska M, Abe T, et al. Cancer related information that cancer patients need. Sci Rep 2025 May 6;15(1):15811. [doi: 10.1038/s41598-025-99498-4] [Medline: 40328851]
12. Blanch-Hartigan D, Chawla N, Moser RP, Finney Rutten LJ, Hesse BW, Arora NK. Trends in cancer survivors' experience of patient-centered communication: results from the Health Information National Trends Survey (HINTS). J Cancer Surviv 2016 Dec;10(6):1067-1077. [doi: 10.1007/s11764-016-0550-7] [Medline: 27193357]
13. Moore P, Vargas A, Núñez S, Macchiavello S. A study of hospital complaints and the role of the doctor-patient communication. Revista médica de Chile 2011 Jul;139(7):880-885. [doi: 10.4067/S0034-98872011000700008] [Medline: 22051825]
14. Akbolat M, Amarat M, Ünal Ö, Şantaş G. A survey of health information seeking by cancer patients indicates some problems over medical explanations and terminology. Health Info Libraries J 2023 Mar;40(1):29-41. [doi: 10.1111/hir.12387]
15. Hack TF, Degner LF, Parker PA, SCRN Communication Team. The communication goals and needs of cancer patients: a review. Psychooncology 2005 Oct;14(10):831-845. [doi: 10.1002/pon.949] [Medline: 16200519]
16. McCormack LA, Treiman K, Rupert D, et al. Measuring patient-centered communication in cancer care: a literature review and the development of a systematic approach. Soc Sci Med 2011 Apr;72(7):1085-1095. [doi: 10.1016/j.socscimed.2011.01.020]
17. Hirayama M, Kasai D, Tabei KI. Development of a behavioral model for the use of psychosocial support services by cancer survivors: a grounded theory pilot study. medRxiv. Preprint posted online on Jul 31, 2025. [doi: 10.1101/2025.07.31.25332471]
18. Paul B, Kirubakaran R, Isaac R, et al. A systematic review of the theory of planned behaviour interventions for chronic diseases in low health-literacy settings. J Glob Health 2023 Sep 8;13:04079. [doi: 10.7189/jogh.13.04079] [Medline: 37681679]
19. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
20. Ishikawa H. Introduction to Health Communication for Healthcare Professionals: Taishukan Shoten; 2022.
21. Duggan A, Street Jr RL. Interpersonal communication in health and illness. In: Health Behavior: Theory, Research, and Practice, 5th edition: Jossey-Bass; 2015:243-267.
22. Niu Z, Bhurosy T, Heckman C. Cancer survivors' emotional well-being: roles of internet information seeking, patient-centered communication, and social support. J Health Commun 2021 Jul 3;26(7):514-522. [doi: 10.1080/10810730.2021.1966685] [Medline: 34435927]
23. Budenz A, Sleight AG, Klein WMP. A qualitative study of online information-seeking preferences among cancer survivors. J Cancer Surviv 2022 Aug;16(4):892-903. [doi: 10.1007/s11764-021-01082-y] [Medline: 34505207]
24. Elkefi S, Matthews AK. Exploring health information-seeking behavior and information source preferences among a diverse sample of cancer survivors: implications for patient education. J Cancer Educ 2024 Dec;39(6):650-662. [doi: 10.1007/s13187-024-02448-3] [Medline: 38807002]
25. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
26. AstraZeneca conducted a web-based survey on ovarian cancer to understand how ovarian cancer patients gather information and to grasp the general public's awareness and understanding of ovarian cancer. AstraZeneca. 2021. URL: https://www.astrazeneca.co.jp/news/press-releases1/2021/2021102101.html# [accessed 2025-04-24]

XSL•FO
RenderX

27. Patient Experience Survey, FY2023. : National Cancer Center Japan; 2024 URL: https://www.ncc.go.jp/jp/icc/policy-evaluation/project/010/2023/index.html [accessed 2025-04-24]
28. Oshio A, Abe S, Cutrone P. Development, reliability, and validity of the Japanese Version of Ten Item Personality Inventory (TIPI-J). Jpn J Personal 2012;21(1):40-52. [doi: 10.2132/personality.21.40]
29. Bunevicius A. Personality traits, patient-centered health status and prognosis of brain tumor patients. J Neurooncol 2018 May;137(3):593-600. [doi: 10.1007/s11060-018-2751-6] [Medline: 29327173]
30. Glaser AW, Fraser LK, Corner J, et al. Patient-reported outcomes of cancer survivors in England 1-5 years after diagnosis: a cross-sectional survey. BMJ Open 2013;3(4):e002317. [doi: 10.1136/bmjopen-2012-002317] [Medline: 23578682]
31. Projected cancer statistics. Cancer Information Service (Ganjoho). 2025. URL: https://ganjoho.jp/en/public/statistics/short_pred.html [accessed 2026-01-14]
32. Nylund-Gibson K, Choi AY. Ten frequently asked questions about latent class analysis. Transl Issues Psychol Sci 2018;4(4):440-461. [doi: 10.1037/tps0000176]
33. Watkins MW. Exploratory factor analysis: a guide to best practice. J Black Psychol 2018 Apr;44(3):219-246. [doi: 10.1177/0095798418771807]
34. Bartlett MS. A note on the multiplying factors for various χ2 approximations. J Roy Statist Soc Ser B 1954 Jul 1;16(2):296-298. [doi: 10.1111/j.2517-6161.1954.tb00174.x]
35. Kaiser HF. An index of factorial simplicity. Psychometrika 1974 Mar;39(1):31-36. [doi: 10.1007/BF02291575]
36. Velicer WF. Determining the number of components from the matrix of partial correlations. Psychometrika 1976 Sep;41(3):321-327. [doi: 10.1007/BF02293557]
37. Mardia KV. Measures of multivariate skewness and kurtosis with applications. Biometrika 1970;57(3):519-530. [doi: 10.1093/biomet/57.3.519]
38. Bandalos DL, Gerstner JJ. Using factor analysis in test construction. In: Schweizer K, DiStefano C, editors. Principles and Methods of Test Construction: Standards and Recent Advances: Hogrefe Publishing; 2016:26-51.
39. Lloret S, Ferreres A, Hernández A, Tomás I. Exploratory factor analysis of the items: analysis guided by empirical data and software. Anales de Psicologia 2017;33(2):417. [doi: 10.6018/analesps.33.2.270211]
40. Child D. The Essentials of Factor Analysis, 3rd edition: Continuum; 2006.
41. Norman GR, Streiner DL. Biostatistics: The Bare Essentials, 4th edition: People's Medical Publishing House; 2014.
42. Jaju A, Crask MR. The perfect design: optimization between reliability, validity, redundancy in scale items and response rates. In: Menon A, Sharma A, editors. 1999 Presented at: AMA Winter Educators' Conference: Marketing Theory & Application; Chicago, IL p. 127-131 URL: https://www.proquest.com/openview/b26e21dd92211bec765015f5be8dab4a/1?pq-origsite=gscholar&cbl=35292 [accessed 2026-05-20]
43. Hayes AF, Coutts JJ. Use Omega rather than Cronbach's alpha for estimating reliability. but…. Commun Methods Meas 2020 Jan 2;14(1):1-24. [doi: 10.1080/19312458.2020.1718629]
44. Psych: procedures for psychological, psychometric, and personality research. CRAN (Comprehensive R Archive Network). 2025. URL: https://cran.r-project.org/web/packages/psych/index.html [accessed 2026-05-20]
45. Hagenaars JA, McCutcheon AL. Applied Latent Class Analysis: Cambridge University Press; 2002. [doi: 10.1017/CBO9780511499531]
46. Kwon JY, Johnson KL, Haase KR, Newton L, Fitch M, Sawatzky R. Patterns of social support among older adults with cancer and associations with patient-reported outcomes: a latent class analysis. J Geriatr Oncol 2025 Jan;16(1):102157. [doi: 10.1016/j.jgo.2024.102157] [Medline: 39616862]
47. Weller BE, Bowen NK, Faubert SJ. Latent class analysis: a guide to best practice. J Black Psychol 2020 May;46(4):287-311. [doi: 10.1177/0095798420930932]
48. Muthén LK, Muthén BO. How to use a Monte Carlo study to decide on sample size and determine power. Struct Equ Modeling 2002 Oct;9(4):599-620. [doi: 10.1207/S15328007SEM0904_8]
49. Linzer DA, Lewis JB. poLCA: an R package for polytomous variable latent class analysis. J Stat Softw 2011;42(10):1-29. [doi: 10.18637/jss.v042.i10]
50. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965 Dec;52(3/4):591. [doi: 10.2307/2333709]
51. Chua GP, Tan HK, Gandhi M. Information sources and online information seeking behaviours of cancer patients in Singapore. Ecancermedicalscience 2018;12:880. [doi: 10.3332/ecancer.2018.880] [Medline: 30483360]
52. Kandasamy S, Khalid AF, Majid U, Vanstone M. Prostate cancer patient perspectives on the use of information in treatment decision-making: a systematic review and qualitative meta-synthesis. Ont Health Technol Assess Ser 2017;17(7):1-32. [Medline: 28572868]
53. Ghazavi-Khorasgani Z, Ashrafi-Rizi H, Mokarian F, Afshar M. Health information seeking behavior of female breast cancer patients. J Educ Health Promot 2018;7:138. [doi: 10.4103/jehp.jehp_17_18] [Medline: 30596110]
54. Shahsavar Y, Choudhury A. Examining influential factors in newly diagnosed cancer patients and survivors: emphasizing distress, self-care ability, peer support, health perception, daily life activity, and the role of time since diagnosis. PLoS One 2023;18(9):e0291064. [doi: 10.1371/journal.pone.0291064] [Medline: 37656716]

XSL•FO
RenderX

55. Ruiz-Romeo M, Ciria-Suarez L, Medina JC, et al. Empowerment among breast cancer survivors using an online peer support community. Support Care Cancer 2024 Dec 28;33(1):56. [doi: 10.1007/s00520-024-09119-5] [Medline: 39730856]
56. Yli-Uotila T, Rantanen A, Suominen T. Motives of cancer patients for using the Internet to seek social support. Eur J Cancer Care (Engl) 2013 Mar;22(2):261-271. [doi: 10.1111/ecc.12025] [Medline: 23320398]
57. Ferraris G, Monzani D, Coppini V, et al. Barriers to and facilitators of online health information-seeking behaviours among cancer patients: a systematic review. Digit Health 2023;9:20552076231210663. [doi: 10.1177/20552076231210663] [Medline: 38107979]

Abbreviations

aBIC: adjusted Bayesian information criterion
BIC: Bayesian information criterion
cAIC: consistent Akaike information criterion
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
EFA: exploratory factor analysis
LCA: latent class analysis
MAP: minimum average partial
SNS: social networking service
TIPI-J: Japanese version of the 10-item Personality Inventory
TPB: theory of planned behavior

Edited by M Balcarras, N Cahill; submitted 15.Jun.2025; peer-reviewed by CN Lee, M Beltrão; revised version received 25.Jan.2026; accepted 26.Jan.2026; published 28.May.2026.

Please cite as:
Hirayama M, Kasai D, Tabei KI
Information-Seeking Patterns and Communication Preferences Among Japanese Survivors With Cancer: Cross-Sectional Analysis
JMIR Cancer 2026;12:e79065
URL: https://cancer.jmir.org/2026/1/e79065
doi:10.2196/79065

© Manae Hirayama, Daisuke Kasai, Ken-ichi Tabei. Originally published in JMIR Cancer (https://cancer.jmir.org), 28.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Extracting Quality of Life Information of Patients Diagnosed With Breast Cancer From Health Care Online Forum Posts: Data Feasibility Study

David Maria Schmidt[1], MSc; Raoul Schubert[1], BA; Brian Po-Han Chen[2], MSc; Deborah Kuk[2], MSc; Valmeek Kudesia[2], MD; Andreas Hinz[3], PhD; Philipp Cimiano[1,4], PhD

[1]Center for Cognitive Interaction Technology, Faculty of Technology, Bielefeld University, Inspiration 1, Bielefeld, Germany

[2]Inspire, Arlington, VA, United States

[3]Department of Medical Psychology and Medical Sociology, University of Leipzig, Leipzig, Germany

[4]Semalytix GmbH, Bielefeld, Germany

Corresponding Author:
David Maria Schmidt, MSc
Center for Cognitive Interaction Technology, Faculty of Technology, Bielefeld University, Inspiration 1, Bielefeld, Germany

Abstract

Background: Quality of life (QoL) questionnaires are used in many disease areas to measure the burden that a disease causes for patients, which help provide insights into disease impact, identify unmet medical needs, and inform patient-centered drug development and value assessment for treatments. The collection of data imposes both a significant burden on patients as well as effort on health care personnel, thus incurring high costs for the health care system. Given that patients share detailed information about their condition and treatment experiences on social media and patient forums, an important research question is to what extent information about QoL can be obtained from patients' online forum posts to potentially complement information obtained from questionnaires.

Objective: This study aimed to assess how much QoL information can be gained from the analysis of posts by patients in online health care communities and whether this information is rich enough to estimate individual patient's QoL based on their posts. We conducted this feasibility study in the context of breast cancer as it is the most prevalent cancer in the female population.

Methods: We recruited 134 female patients diagnosed with breast cancer on the Inspire patient online forum, who voluntarily participated in our feasibility study. They filled in the EORTC (European Organisation for Research and Treatment of Cancer) QLQ-C30 and QLQ-BR23 questionnaires consisting of 30 general questions and 23 additional breast cancer–specific questions and provided consent to analyze their posts and comments on the online forum (756 posts and 19,478 comments). Posts were coded manually to identify parts of the text providing answers to 1 of the aforementioned 53 questions.

Results: The data annotation yielded a substantial agreement (mean Fleiss κ of 0.5, SD 0.28). Overall, we found answers in the coded data for 50 out of 53 EORTC QLQ-C30 and QLQ-BR23 questions. The information coded in the posts reliably predicted the answers given in the questionnaires (F_1-score=0.7), with even better results when grouping similar questions (F_1-score=0.8 for fine-grained and 0.9 for coarse-grained grouping). The 5 questions that were most frequently answered on the basis of the coded posts were "Did you feel ill or unwell?" (304 of 2683 annotated posts and comments), "Did you worry?" (105 posts and comments), "Have you had pain?" (104 posts and comments), "Did you feel tense?" (85 posts and comments), and "Were you limited in doing either your work or other daily activities?" (77 posts and comments).

Conclusions: Our feasibility study shows that there is valuable QoL-related information in posts of online patient communities, which can potentially serve as an innovative low-burden QoL monitoring approach. Future research should consider how these insights can be used to complement existing QoL instruments and whether the process of extracting QoL-related information can be automated.

(JMIR Cancer 2026;12:e76044) doi:10.2196/76044

KEYWORDS

quality of life; social media listening; patient-focused drug development; health care communities; breast cancer

XSL•FO
RenderX

Introduction

Quality of life (QoL), as defined by the World Health Organization, encompasses physical, mental, and social well-being beyond the mere absence of disease [1]. Assessing QoL is essential in health care, providing insights into disease impact, identifying unmet medical needs, and informing patient-centered drug development and value assessment for treatments. Traditionally, QoL is measured using patient-reported outcome questionnaires, such as the widely used EORTC (European Organisation for Research and Treatment of Cancer) QLQ-C30 for patients with cancer [2]. These standardized tools facilitate comparisons across patient groups and with the general population.

However, QoL assessments are often time-consuming and administratively burdensome for both patients and health care providers [3-7]. Efforts to streamline data collection have explored digital and computer-supported approaches [6-8]. Given the increasing use of social media and online patient communities for sharing health experiences, a key question arises: Can patient-generated online content serve as a proxy or complementary source for QoL data collection? While not a replacement for structured surveys, social media data may provide real-time, cost-effective insights into patient well-being, preferences, and treatment experiences.

Previous research has leveraged social media to assess disease burden and treatment experiences across conditions, such as Parkinson disease [9,10], Sjögren disease [11], chronic obstructive pulmonary disease [12], and mental disorders [13,14]. Some related work also analyzed patient posts compared to a specific QoL ground truth, but on an aggregated level. For example, Zivanovic et al [15] compared Twitter data with official QoL data of the city of Bristol. Similarly, Cotté et al [16] also investigated cancer-specific forums and QoL data, but with a focus on immune checkpoint inhibitors and not the QoL impact for individual users. Therefore, studies have yet to systematically evaluate whether individual QoL data can be reliably extracted from online patient discussions.

To explore this, we conducted a study with patients diagnosed with breast cancer from the Inspire patient online forum, assessing whether online forum posts contain relevant QoL information. Breast cancer remains the most commonly diagnosed cancer in women, with over 2.3 million new cases in 2020 [17]. QoL is a critical concern for patients diagnosed with breast cancer and survivors, who frequently experience significant impairments [18-21].

In our study, patients diagnosed with breast cancer completed standard QoL questionnaires (EORTC QLQ-C30 and QLQ-BR23), while their online posts and comments were manually coded to determine the presence of QoL-related content. We then assessed the prevalence of QoL discussions and their correlation with questionnaire responses. Our research focuses on the following questions:

- RQ1: Do patient online forum posts contain QoL-related information?
- RQ2: Which QoL questions from EORTC QLQ-C30 and QLQ-BR23 are most frequently addressed in online posts?
- RQ3: How well do these online discussions align with individual patients' QoL questionnaire responses?

This study aims to determine whether patient-generated content from online health communities can complement traditional QoL assessments, potentially offering a scalable, low-burden alternative for understanding patient well-being.

Methods

Overall Structure

This feasibility study aimed to assess the presence of QoL information in online discussions by patients diagnosed with breast cancer. Breast cancer was selected due to its high prevalence, and data were collected from patient-generated content on the Inspire patient online forum. Patient communities on this forum are organized by condition, including "breast cancer" and "advanced breast cancer," where each thread of natural peer-to-peer conversation consists of an initial post and subsequent comments. An overview of the workflow followed in this study is given in Figure 1.

To compare patient-reported QoL in online posts with structured assessments, we (the authors of this study) administered the QLQ-C30 questionnaire by the EORTC [2] and its breast cancer–specific module, QLQ-BR23. These validated instruments provided a reference (ground truth) against which the content of online discussions could be analyzed.

Participants were recruited from the "Breast Cancer" and "Advanced Breast Cancer" communities on Inspire. Eligibility criteria included female patients diagnosed with breast cancer, age 18 years or older, residency in the United States, and at least 1 post or comment in the respective Inspire communities.

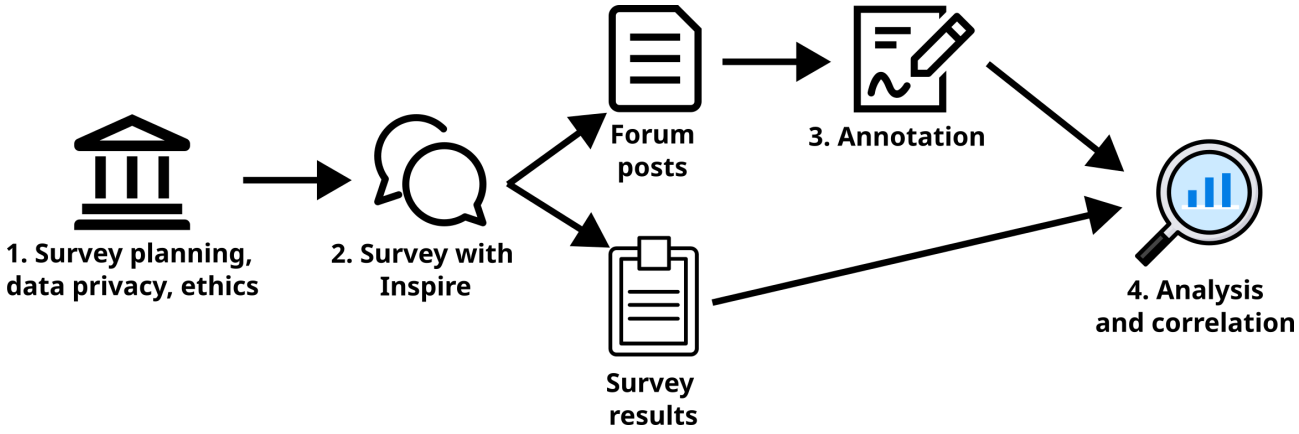
Residency in the United States was included due to Inspire being based in the United States.

The EORTC QLQ-C30 questionnaire was used to assess general QoL in patients with cancer. It comprises 30 items grouped into 5 functioning scales, a global health status/QoL scale, 3 symptom scales, and 6 single-item measures. Scores range from 0 to 100, with higher functioning scores and lower symptom scores indicating better QoL. Normative values for this instrument have been previously established [22,23]. The breast cancer–specific module, EORTC QLQ-BR23, consists of 23 items categorized into 8 scales: systemic therapy side effects, hair loss, arm symptoms, breast symptoms, body image, future perspective, sexual functioning, and sexual enjoyment. Response options align with those of the EORTC QLQ-C30.

The final survey also included demographic questions, such as year of birth, US state of residence, gender, sex at birth, insurance type, and residential setting (urban, suburban, rural, or other). The survey required approximately 5 to 10 minutes to complete and was conducted between May and June 2024.

XSL•FO
RenderX

Figure 1. Overview of the study workflow from planning survey conduction to data analysis. The online survey using the EORTC QLQ-C30 and QLQ-BR23 questionnaires was conducted between May and June 2024 with patients diagnosed with breast cancer posting in the Inspire patient online forums. EORTC: European Organisation for Research and Treatment of Cancer.



Ethical Considerations

Our study has been approved by the Ethics Committee of Bielefeld University under application number 2023 - 216-W1. A filled CHERRIES (Checklist for Reporting Results of Internet E-Surveys) [24] checklist can be found in Checklist 1. Informed consent was obtained from all participants, granting permission to analyze their posts and comments for this study. The collected data, especially the survey responses, are not publicly shared and only available to the authors of the study. Additionally, the data are stored in encrypted form only. All eligible community members were invited to participate (ie, convenience sampling), with the first 100 respondents receiving a US $20 gift card as compensation.

Annotation and Coding Process

The objective of the coding process was to identify text spans within collected posts and comments that could correspond to responses in the EORTC QoL questionnaires. To achieve this, we implemented an annotation schema similar to named-entity recognition, where each of the 53 QoL questions served as a label. Unlike traditional named-entity recognition, which identifies entities, such as persons or organizations [25,26], our approach used QoL-related labels to tag relevant text segments.

In addition to assigning labels corresponding to each QoL question, we introduced supplementary labels to capture both reported presence and absence of symptoms: “positive” or “negative.” Coders applied these labels alongside question-specific tags to indicate whether a response reflected a favorable or unfavorable QoL experience. Furthermore, an “Exact match” label was used to denote text spans that provided precise or highly relevant answers, while standard labels without this addition were applied to responses that were partial or vague.

The annotation schema and guidelines were developed in a 2-phase process. In the first phase, 3 annotators labeled 30 posts using a preliminary schema. After completing this pilot phase, the results were reviewed, and minor refinements were made. For instance, the label “Did you feel ill or unwell?” was designated for general QoL descriptions that did not fit into other, more specific categories. The annotators were explicitly trained to only use the fallback label when they could not find any other label that fits better, and this was also added to the annotation guidelines (Multimedia Appendix 1). Compliance with these rules was ensured by closely checking fallback annotations during the beginning of the annotation process of the full data. These pilot annotations were subsequently excluded from the final dataset.

In the second phase, the full annotation process commenced, with no further modifications needed. Additionally, posts and comments from the last 6 months (relative to each user’s participation date) were labeled to indicate whether they referred to past events. This distinction was critical for evaluating the correlation between forum data and survey responses, as older event descriptions may not accurately reflect a patient’s current QoL. Filtering out such posts allowed for a more precise alignment between forum-derived insights and survey-based QoL assessments.

Measuring Information Consistency Between Online Posts and Survey Responses

To assess the alignment between information shared in online posts and responses in the structured questionnaires, we applied classification metrics commonly used in machine learning. The coding scheme and questionnaire responses were simplified into a binary format, mapping each piece of information from posts and survey answers to either a “positive” or “negative” response. The intended meaning of “positive” here is that the corresponding question would be answered with “yes,” thus indicating the presence of symptoms or other issues. Analogously, “negative” corresponds to answering the question with “no,” indicating the absence of symptoms or other issues. Specifically, the answer categories “a little,” “quite a bit,” and “very much” were classified as “positive,” whereas “not at all” was categorized as “negative.” This enabled the computation of key classification metrics as follows:

- True positives (TP): Instances where both the online post and the corresponding questionnaire response indicate a “positive” answer (eg, a user stating, “I had so much pain last week!” and selecting “yes” for the survey question, “Have you had pain?”).
- False positives (FP): Cases where an online post is labeled as “positive,” but the corresponding survey response is

XSL•FO
RenderX

"negative" (eg, a user stating, "I am so worried!" but responding "no" to the question, "Did you worry?").
- True negatives (TN): Instances where both the online post and the survey response indicate a "negative" answer (eg, a user stating, "I do not have any pain!" and selecting "no" for the survey question, "Have you had pain?").
- False negatives (FN): Cases where an online post is labeled as "negative," but the corresponding survey response is "positive" (eg, a user stating, "I did not worry at all last week!" but selecting a "positive" response in the survey question, "Did you worry?").

If no labeled data existed for a specific user and QoL question, that user did not contribute to the TP, FP, TN, or FN calculations for that question. This approach allowed us to evaluate how well the available data could predict survey responses.

Using the TP, FP, TN, and FN values, we computed standard machine-learning classification metrics, including precision (TP/[TP + FP]), recall (TP/[TP + FN]), and F_1-score as harmonic mean between precision and recall, that is, 2× precision × recall/(precision + recall) [27]. Given that data sparsity increases when analyzing results at the individual user level across all 53 survey questions, we explored grouping similar survey questions based on the EORTC scoring manual. This resulted in 2 levels of categorization: (1) fine-grained grouping: based on detailed QoL question clusters (Table 1) and (2) coarse-grained grouping: functional (questions 1-7, 20-27, 39-46), symptoms (questions 8-19, 28, 31-38, 47-53), and global health (questions 29 and 30).

These groupings enabled more robust analyses by reducing sparsity and improving interpretability of the correspondence between online forum discussions and structured survey responses.

Table . Fine-grained grouping based on the EORTC[a] QLQ-C30 and QLQ-BR23 scoring manual[b].

Group	Question IDs
Global health	29, 30
Physical functioning	1, 2, 3, 4, 5
Role functioning	6, 7
Emotional functioning	21, 22, 23, 24
Cognitive functioning	20, 25
Social functioning	26, 27
Fatigue	10, 12, 18
Nausea and vomiting	14, 15
Pain	9, 19
Dyspnea	8
Insomnia	11
Appetite loss	13
Constipation	16
Diarrhea	17
Financial difficulties	28
Systemic therapy side effects	31, 32, 33, 34, 36, 37, 38
Upset by hair loss	35
Arm symptoms	47, 48, 49
Breast symptoms	50, 51, 52, 53
Body image	39, 40, 41, 42
Future perspective	43
Sexual functioning	44, 45
Sexual enjoyment	46

[a]EORTC: European Organisation for Research and Treatment of Cancer.

[b]Question IDs refer to the corresponding EORTC questionnaire question numbers. The online survey using the EORTC QLQ-C30 and QLQ-BR23 questionnaires was conducted between May and June 2024 with patients diagnosed with breast cancer posting in the Inspire patient online forums.

XSL•FO
RenderX

Results

General Statistics

The collected dataset comprises 20,204 posts and comments from 2006 to 2024, with 580 posts and comments originating within the 6 months preceding the end of the study. The distribution of posts per user follows a long-tail pattern, as illustrated in Figure 2, with absolute values presented in Table 2). As shown in Figure 2, the majority of contributions are comments rather than original posts. Table 3 shows that our data cover a broad spectrum of patients diagnosed with breast cancer, both in terms of age and distance between survey and breast cancer diagnosis.

Figure 2. Posts per user, with 123 users on the x-axis and the (log-scaled) number of posts on the y-axis, differentiating between combined posts and comments in light blue and comments only in dark blue. Users without posts or comments were disregarded. The online survey using the EORTC QLQ-C30 and QLQ-BR23 questionnaires was conducted between May and June 2024 with patients diagnosed with breast cancer posting in the Inspire patient online forums. EORTC: European Organisation for Research and Treatment of Cancer.

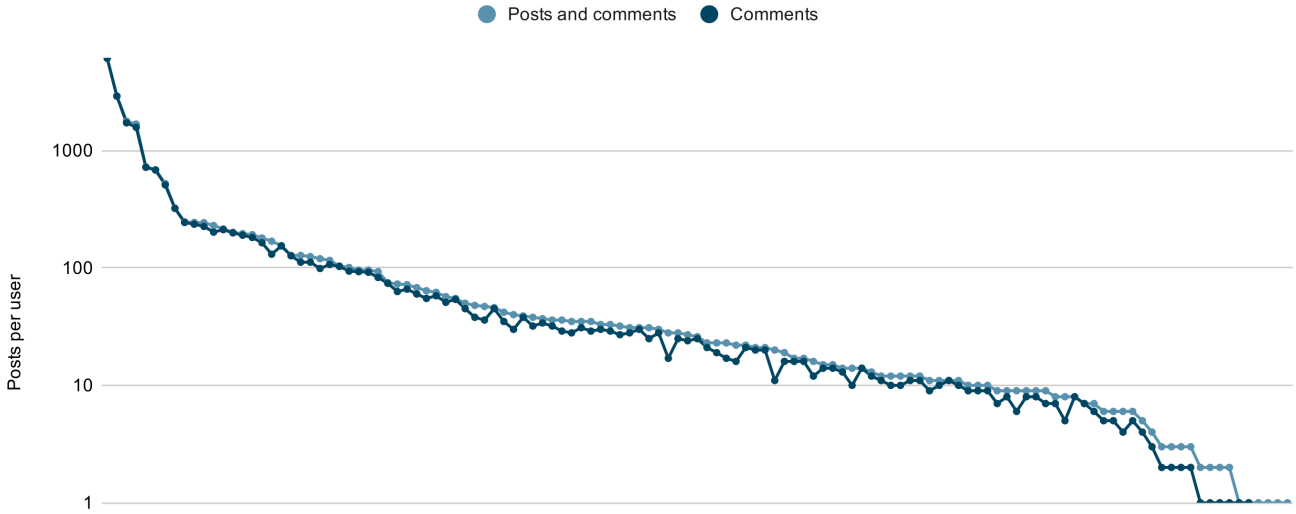


Table . Overview of the number of posts per user[a].

Number of posts and comments	Number of users
1	6
2-10	28
11-100	63
101-1000	22
>1000	4

[a]The online survey using the EORTC QLQ-C30 and QLQ-BR23 questionnaires was conducted between May and June 2024 with breast cancer patients posting in the Inspire patient online forums.

The average post length is 408.4 (SD 366.8) characters. A total of 134 participants were enrolled in the study, of whom 11 did not contribute at least 1 post or comment, resulting in 123 participants meeting the inclusion criteria.

For annotation, the last 6 months of data (580 posts and comments) were labeled by 3 annotators, with an average annotation time of 15.9 (SD 10.3) hours per annotator. The last 24 months of data (an additional 2103 posts and comments) were labeled by 2 annotators, requiring an average of 76.5 (SD 12.7) hours per annotator. In total, 2683 posts were annotated. Fourteen posts contained QoL-related information referring to someone other than the posting user. These posts were annotated but excluded from statistical evaluation.

Of the 2669 annotated posts and comments included in the final dataset, 613 (23%) contained at least 1 label corresponding to a QoL question. When considering only the last 6 months, the proportion was higher at 37.4% (217 out of 580 posts), indicating that approximately one-third to one-fourth of all posts contained QoL-relevant content.

XSL•FO
RenderX

Table . Statistics of self-reported demographic data of survey participants[a].

Data category	Mean (SD)	Minimum	Maximum
Age	68.2 (11.5)	27	116
Age at diagnosis	55.7 (12.6)	26	103
Years since diagnosis	13.1 (9.4)	0	45

[a]The online survey using the EORTC QLQ-C30 and QLQ-BR23 questionnaires was conducted between May and June 2024 with patients diagnosed with breast cancer patients posting in the Inspire patient online forums.

Among the 53 QoL questions, 50 were labeled at least once, with the most frequently occurring questions summarized in Table 4. The most commonly assigned label was "Did you feel ill or unwell?," with 301 annotations, representing 11% (n=301) of all posts and 57.3% (n=47) of users who actively posted during the study period. It has to be noted that this is used as a fallback label for the case where posts or comments contain information which is clearly QoL-relevant, but no other better-fitting label could be found. An example for this is when patients describe some very specific or very vague symptoms that do not correspond well to any other question (eg, "the symptoms increased a lot"). The next most frequently annotated questions were "Did you worry?" (105 annotations) and "Have you had pain?" (102 annotations), both capturing broad categories of patient discomfort. The overall label distribution is visualized in Multimedia Appendix 2.

Table . Top 10 most frequent labels with respect to number of posts they are annotated in for all annotated posts (most recent 24 months)[a].

Question	Number of posts	Portion of posts, %	Number of users	Portion of users, %
Did you feel ill or unwell?	301	11	47	57.3
Did you worry?	105	3.9	26	31.7
Have you had pain?	102	3.7	32	39
Did you feel tense?	84	3.1	21	25.6
Were you limited in doing either your work or other daily activities?	77	2.8	18	22
Were you tired?	55	2	18	22
Have you felt nauseated?	51	1.9	18	22
Were you worried about your health in the future?	41	1.5	16	19.5
Have you lacked appetite?	32	1.2	5	6.1
Did you feel depressed?	32	1.2	14	17.1

[a]"Portion of posts" refers to the percentage of posts with the corresponding label relative to all posts of that time period. Similarly, "Portion of users" refers to the number of users who have posts annotated with that label relative to all users who posted in that time period. The online survey using the European Organisation for Research and Treatment of Cancer (EORTC) quality of life questionnaires (QLQ) QLQ-C30 and QLQ-BR23 was conducted between May and June 2024 with patients diagnosed with breast cancer posting in the Inspire patient online forums.

Annotator Agreement

As every post and comment within the most recent 2 years was annotated by 2 or 3 annotators, the interannotator agreement can be calculated via Fleiss κ values [28]. This measure aids in determining to which degree annotators agreed on which text segments or posts answered that question at least partially. We only consider the basic labels corresponding to the questions for the calculation of interannotator agreement, ignoring the additional labels "negative" and "exact match," and handle each question separately, as this fits best with the setting assumed by the Fleiss κ metric for interannotator agreement.

We calculated the Fleiss κ for 2 different levels of annotation granularity, that is, on word and on post or comment level, respectively. The word level reflects the actual annotations, that is, the text segments created during annotation, whereas the post level is a simplification aggregating all annotations of a post such that the annotations only have to agree on whether the whole text answers a specific question or not, but not where exactly.

The top and bottom 10 Fleiss κ values for the most recent 6 months of data are provided in Table 5, together with the absolute numbers of posts in which the corresponding question has been answered.

It can be observed that many frequently occurring labels exhibit solid to high annotator agreement, for example, "Have you lost any hair?" has a Fleiss κ value of 0.7 for word units and even 1.0 for post units while being present in 26 posts in total. Even the fallback label "Did you feel ill or unwell?" with 285 labeled posts has Fleiss κ values of 0.6 and 0.7 for word and post granularity, respectively.

Most questions in the lower range of Fleiss κ values have a very low number of labeled posts. The 6 questions with the lowest

XSL•FO
RenderX

observed agreement of 0.0 together only account for 14 posts in total, so these appear to be cases where the annotators do not agree whether the label is applicable. Further investigating the data, the most disagreement is caused by vaguely formulated posts for which a label could but does not necessarily have to apply. For example, for labels, such as "Did you have any pain in your arm or shoulder?," issues can range from just vaguely described body areas to mentions of certain diseases or syndromes, which are closely connected to pain in that area, without pain being explicitly mentioned. Although we cannot give real posts as examples for data privacy reasons, we summarized prototypical posts to illustrate the reasons for disagreement. For "Did you have any pain in your arm or shoulder?," annotators disagreed for posts, such as "I take that medication for 5 years now and started to have problems with hand-foot syndrome some time ago," as hand-foot syndrome itself is not an explicit mention of pain and can occur with or without pain [29]. Similarly, annotators disagreed for posts, such as "In the last months I had cramps in different parts of my hand from time to time." Another label with low annotator agreement was "Have you had skin problems on or in the area of your affected breast (e.g., itchy, dry, flaky)?," for example, for posts like "I try to avoid foods with high folic acid as much as possible as they can cause pain in the arms for some of us. For the dryness, I use some ointment all the time," as it is not very clear what part of the body "dryness" refers to.

XSL•FO
RenderX

Table . Annotator agreement for the 10 questions with the highest agreement and 10 questions with the lowest agreement (Fleiss κ, the most recent 6 months of data before the survey) using word units and post units[a].

Question	Fleiss κ (word units)	Fleiss κ (post units)	Labeled posts
Top 10			
Did you have headaches?	1.0	1.0	3
Have you had difficulty remembering things?	0.9	1.0	3
Have you lacked appetite?	0.9	0.9	15
Did you have hot flushes?	0.8	0.8	9
Have you lost any hair?	0.7	1.0	26
To what extent were you sexually active? (with or without intercourse)	0.7	1.0	6
Did you feel ill or unwell?	0.6	0.7	285
Have you had diarrhea?	0.6	0.9	10
Do you have any trouble taking a short walk outside of the house?	0.6	0.7	19
Were you tired?	0.6	0.8	35
Bottom 10			
Have you had skin problems on or in the area of your affected breast (eg, itchy, dry, and flaky)?	0.2	0.3	11
Have you vomited?	0.2	0.5	2
Do you have any trouble taking a long walk?	0.2	0.3	7
Did you have any pain in your arm or shoulder?	0.2	0.4	11
How would you rate your overall health during the past week?	0.0	0.0	1
Did you need to rest?	0.0	0.0	1
Was it difficult to raise your arm or to move it sideways?	0.0	0.0	1
Did pain interfere with your daily activities?	0.0	0.0	3
Did food and drink taste different than usual?	0.0	0.0	6
How would you rate your overall quality of life during the past week?	0.0	0.0	2

[a]The online survey using the European Organisation for Research and Treatment of Cancer (EORTC) quality of life questionnaires (QLQ) QLQ-C30 and QLQ-BR23 was conducted between May and June 2024 with patients diagnosed with breast cancer posting in the Inspire patient online forums.

Correlation Results Between Annotated Data and Survey Data

We estimate the correlation between the information provided by the authors to QoL aspects and the data provided in the questionnaires by computing F_1-score as described in the "*Measuring Information Consistency Between Online Posts and Survey Responses*" section. Table 6 shows the F_1-scores and total annotations for the 10 most frequently annotated questions based on data originating from the most recent 6-month period and compares these numbers to the corresponding values for 24 months.

As observed from Table 6, all of the 10 most frequently answered questions can be predicted reliably with an F_1-score of 0.6 or higher for the last 6 months of data. For 7 of the 10 questions, the F_1-score is even 0.8 or higher. The lowest score of 0.6 is achieved for the most frequent question, that is, "Did you feel ill or unwell?" This specific question is also used as a fallback label for utterances, which do not fit in any other category, thus increasing prediction difficulty compared to other labels. Even when considering a much larger time period of 24 months, the F_1-scores for these questions do not drop below 0.5, with some even increasing.

XSL•FO
RenderX

Although the observed F_1-scores are high for many questions, for example, "Were you tired?" or "Were you worried about your health in the future?" with each of an F_1-score of 1.0, for some questions, the available data get sparse and the data sparsity increases even further when considering single users or shorter time spans, leading to lower F_1-scores in some cases, for example, for "Have you vomited?" (F_1-score of 0.2 with 2 labeled posts in 6 months of data). One way of reducing data sparsity is to consider a larger time span. However, this comes at the cost of posts and comments with a larger temporal distance from the survey date being potentially less relevant with respect to the QoL at that time. However, in Table 6, despite the inclusion of older posts, we can observe solid to high F_1-scores (0.6-1.0, mean 0.8, SD 0.1) for all 10 most frequently answered questions. This shows that including posts from dates outside the time period asked for in the QoL surveys can still be useful.

Table . F_1-scores for the 10 most frequently annotated questions, describing how well annotations predict the actual answers given in the survey, considering the last 6 and 24 months of data, which are not filtered by time deviation[a].

Question	F_1-score (6 mo)	Total annotations (6 mo)	F_1-score (24 mo)	Total annotations (24 mo)
Did you feel ill or unwell?	0.6	127	0.7	222
Did you worry?	0.8	36	0.9	47
Did you feel tense?	0.9	33	0.8	40
Have you had pain?	0.8	28	0.8	74
Were you tired?	1.0	14	0.9	50
Have you felt nauseated?	0.7	14	0.5	34
Has your physical condition or medical treatment interfered with your social activities?	0.7	13	0.7	13
Has your physical condition or medical treatment caused you financial difficulties?	1.0	12	1.0	12
Were you worried about your health in the future?	1.0	11	1.0	15
Have you lost any hair?	0.8	10	0.8	29

[a]The online survey using the European Organisation for Research and Treatment of Cancer (EORTC) quality of life questionnaires (QLQ) QLQ-C30 and QLQ-BR23 was conducted between May and June 2024 with patients diagnosed with breast cancer posting in the Inspire patient online forums.

As some questions occur sparsely, we group the answers from the posts and the questionnaires into the groups "Symptomatic" and "Functional" used in the EORTC scoring manual. The results for the grouped questions are shown in Table 7.

The only group in Table 7 having annotations available and exhibiting a lower F_1-score than 0.7 for the last 6 months is "cognitive functioning," which comprises only 2 questions, that is "Have you had difficulty in concentrating on things, like reading a newspaper or watching television?" and "Have you had difficulty remembering things?" All other groups show very promising results, with over half of the groups having scores of 0.9 or higher. Groups of questions for which there were no annotations in the data, that is "Appetite loss," "Global health," "Insomnia," "Sexual enjoyment," "Sexual functioning," and "Upset by hair loss," have been omitted in the table. Even when considering much older data from the last 24 months as well, F_1-scores do not drop below 0.5, with many scores even increasing, underlining the relevance of older data.

For 6 months of data, a total of 31 users have shared posts containing QoL information with on average 10.8% (SD 9.8%) of the 53 QoL questions answered. For the full 24 months of annotated data, there are 46 users with QoL-relevant posts, with 10.8% (SD 10.1%) of the questionnaire questions answered on average.

XSL•FO
RenderX

Table . F_1-scores describing how well annotations predict the actual answers given in the survey, considering the last 6 and 24 months of data, using fine-grained and coarse-grained groups, which are filtered by time deviation[a].

Question	F_1-score (6 mo)	Total annotations (6 mo)	F_1-score (24 mo)	Total annotations (24 mo)
Fine grouping				
Fatigue	1.0	5	0.9	44
Future perspective	1.0	5	1.0	9
Arm symptoms	1.0	4	1.0	6
Diarrhea	1.0	3	0.6	16
Dyspnea	1.0	2	0.9	14
Nausea and vomiting	1.0	2	0.6	25
Body image	1.0	1	1.0	1
Financial difficulties	0.9	9	0.9	9
Emotional functioning	0.9	38	0.9	70
Social functioning	0.9	11	0.9	11
Physical functioning	0.9	9	0.8	12
Breast symptoms	0.9	7	0.9	11
Role functioning	0.8	10	0.7	21
Systemic therapy side effects	0.7	59	0.9	189
Pain	0.7	15	0.8	61
Constipation	0.7	2	0.8	10
Cognitive functioning	0.0	1	0.5	3
Insomnia	—[b]	0	1.0	1
Appetite loss	—	0	0.7	16
Coarse grouping				
Functional	0.9	68	0.9	120
Symptomatic	0.8	101	0.9	385

[a]The online survey using the EORTC QLQ-C30 and QLQ-BR23 questionnaires was conducted between May and June 2024 with patients diagnosed with breast cancer posting in the Inspire patient online forums.

[b]Not applicable.

Discussion

The goal of this feasibility study is to investigate whether we can reliably gain QoL information from posts of patients in online communities. To this end, we have introduced a new methodology that examines the content of posts in relation to data from QoL surveys completed by the same cohort of patients whose posts we have access to. This has allowed us to not only assess how frequently users post information that is relevant to QoL, but also measure the degree to which this information corresponds to the answers to the corresponding questions in the questionnaires. This methodology is novel, as it has not previously been applied to analyzing whether information shared by patients online can reliably estimate individual QoL data. We have provided evidence that this methodology is sound, showing that the information can be coded at good levels of agreement and that there is some objective "ground truth" in the information shared by patients online that can be consistently interpreted as providing answers to the survey questions. With regard to RQ1, we can say that patient online forum posts indeed contain QoL-related information. The questions most frequently addressed in those posts (RQ2) are "Did you worry?" "Did you feel tense?" and "Have you had pain?" Moreover, the online posts align well with the actual QoL questionnaire responses (RQ3), achieving an F_1-score of 0.7, with even better results when grouping similar questions (F_1-score =0.8 for fine-grained and 0.9 for coarse-grained grouping).

In particular, Fleiss κ values were nonnegative with "Did you have headaches?" achieving the highest score of 1.0. Only 6 out of 53 questions showed agreement values of 0, with the mean Fleiss κ value being 0.5 (SD 0.28), indicating moderate to solid agreement among the annotators, especially considering the complexity of the task.

We have further quantified the degree to which patients share information that can be mapped to questions from the EORTC QLQ-C30 and QLQ-BR23 questionnaires.

XSL•FO
RenderX

From the total of 2669 annotated posts and comments, 613 (23%) answered at least 1 question from the used EORTC questionnaire. For the last 6 months (since the end of the survey), even 217 out of 580 posts (37.4%) contain QoL-relevant information. As these numbers might be affected by selection bias (indicated by older posts containing less QoL-relevant information), they should be seen more as upper bounds compared to scenarios where the data of an arbitrary patient are investigated.

We have further shown that there is a substantial correlation between the information that participants have provided in their online posts and the survey responses. We have proposed to rely on standard evaluation metrics as used in machine-learning research to estimate the correctness of the information provided by participants in their online posts compared to their answers to corresponding questions in the questionnaires.

The average score for the most recent 6 months of data across all 33 questionnaire questions that had answers in the respective data was an F_1-score of 0.7, indicating that posts and comments containing QoL-relevant information can reliably predict questionnaire answers. Similarly, from the patient level, an average of 63.4% (SD 33.5%) of the estimated answers to the questions is correct. This indicates a high predictive power of forum posts which contain QoL information.

When grouping similar questions together, we even get a mean F_1-score of 0.8 (SD 0.154) for fine-grained grouping and a mean F_1-score of 0.9 (SD 0.01) for coarse-grained grouping when applying the time deviation filtering, excluding posts and comments talking about events and experiences from an unspecified or distant past. Without filtering, we get mean F_1-scores of 0.7 (SD 0.353) without grouping, 0.8 (0.158) with fine-grained grouping, and 0.9 (SD 0.051) with coarse-grained grouping.

The most frequently answered questions (RQ2) include “Did you worry?,” “Did you feel tense?,” “Have you had pain?,” “Were you limited in doing either your work or other daily activities?,” “Were you tired?,” “Have you felt nauseated?,” “Were you worried about your health in the future?,” “Have you lacked appetite?,” and “Did you feel depressed?” These questions are not particularly specific to breast cancer, thus indicating patients typically discuss their QoL online more on a general level or in such a specialized manner that none of the EORTC questions fit well as shown by the high number of “Did you feel ill or unwell?” occurrences.

One inherent problem of the collection and analysis of the questionnaire data is related to missing data, as often patients do not fill out all information. In our case, the rate of completeness for the survey data was 100%, as the completion of all questionnaire items was required. Two questions of the form “If you were/had …” allowed choosing “no answer,” which, however, we do not consider missing data but instead a different way to answer the question.

However, not all questions were answered in posts of all patients. For patients who have shared QoL-relevant information in their posts, only about 10.8% (SD 9.8%) of the questions could be answered by information given in the posts. This means that the available data are typically not sufficient to answer all questions of the EORTC QoL questionnaire, but at least some parts can be reliably extracted for each patient. We also investigated ways to deal with data sparsity by grouping questions and considering longer time periods, which help to get more insights into the QoL without losing too much reliability.

Overall, our results show that users of online communities, such as Inspire, indeed post information about their QoL, providing important information that can be consistently and objectively mapped to the categories of standard questionnaires.

Our results have important implications. First, the results suggest that we could potentially seek answers to questions that are typically not included in standardized QoL questionnaires. Furthermore, there are nuances in the posts of patients that are not directly captured by the questionnaires. On one hand, the topics that patients do or do not bring up in online discussions may indicate which topics are currently the most important for that patient, potentially referring to specific aspects of the relatively broad areas covered by the QoL questionnaires. On the other hand, this can provide insights into areas, which are not covered by the QLQ because they are unique to that patient (eg, due to certain living circumstances). This hints at the potential of our methodology to complement data obtained from standard QoL questionnaires.

Comparison With Prior Work

In recent years, there have been numerous studies investigating whether QoL-related information can be gained from unstructured data sources. However, our work differs from related approaches in substantial ways.

For one, our approach leverages different types of data considered. We considered social media posts from dedicated health care online communities or more precisely the breast cancer forum of Inspire. Other work has focused on clinical reports [30,31] or health records [31,32], which, in spite of also having unstructured elements, considerably differ in terms of content and style from social media data. Among similar approaches considering social media, many approaches have sourced Twitter data [13,15,33-37], in contrast to the specific health care online community data that we used. An exception is the work by Leung et al [38,39] who have used data from cancer support groups. However, their data originate from chats moderated by health care professionals instead of posts from organic, forum-like online communities. Their study had a narrower focus by studying signs of emotional distress in patients participating in the online discussions instead of the broader focus of capturing QoL-related information in our study. An overview of how social media contents can be analyzed for health care is given by Fu et al [40].

Our study has aimed to gain insights into the QoL of individual patients based on their posts in health care communities. In contrast, the goal of many other studies is to discover aggregate trends for a certain geographic segment of the target population [13,15,36]. Other studies are interested in population-wide trends and effects related to how a disease, intervention, or treatment

XSL•FO
RenderX

affects patients in general, instead of estimating values for specific users [9,11,12,37,41-52]. Such population-wide analyses can support the creation of conceptual models of a disease [9,42], whereas other studies have focused on general and individual health [33], recognition of discussions about health services [53], or emotional distress in participants of online chats [38,39].

Some studies have been conducted in relation to existing QoL questionnaires. For example, Tapi Nzali et al [54] have also used the EORTC QLQ-C30 and QLQ-BR23 questionnaires, aiming at identifying general topics that patients diagnosed with breast cancer discuss in relation to EORTC questionnaires. Gries and Fastenau [52] have relied on items in the EORTC QLQ-C30 and QLQ-MY20 questionnaires to investigate concepts of interest, symptoms, and impacts discussed by patients with multiple myeloma on a specific health care social media platform. These studies analyze general trends rather than predicting QoL for individual patients.

Renner et al [55] annotated data from 19 health-related forums using the Short Form Health Survey SF-36 [56] and Euro Quality of Life 5 Dimensions [57], but they did not compare and contrast survey responses with the information in patient posts as we do.

There are few studies that have analyzed the information posted by patients in relation to ground truth QoL data as available in the existing QoL questionnaire data. For example, Zivanovic et al [15] have compared Twitter data originating from the city of Bristol with official QoL data for this city. They analyzed the data on an aggregate level for the whole city rather than for individual patients as we do. Twitter has also been used by Sarma et al [35] who compared information obtained from Tweets with ground truth data of Centers for Disease Control Healthy Days Questionnaire [58]. The questionnaire used is much shorter than the one used in our study, providing a less detailed view on patients' QoL. Cotté et al [16] have compared data from cancer-specific forums with data obtained from the Functional Assessment of Cancer Therapy-General [59] and EORTC QLQ-C30 questionnaires. Their main focus has been on investigating the impact of immune checkpoint inhibitors, but not investigating the QoL impact for individual users.

Although this work focuses on breast cancer, related work covers a wide range of diseases. These diseases include Parkinson disease [9,10], amblyopia [41], presbyopia [42,49], cancer [16,30,48,51], dry eye disease [43], chronic ocular pain [44], Sjögren disease [11], chronic obstructive pulmonary disease [12], systemic lupus erythematosus [45,60], immunoglobulin A nephropathy [46], complement 3 glomerulopathy [47], atopic dermatitis [50,61], obesity [36], depression [62], and (general) mental health [13,14].

In summary, our work advances the field by paving the way for a new option to monitor patients' QoL. As elaborated in the previous section, analyzing forum posts with respect to QoL questions can be used to complement existing instruments for assessing QoL, helping to focus on the most relevant aspects of patients' QoL as well as to discover areas potentially not covered by existing instruments in detail. Additionally, possible ways to automate the process of identifying QoL-relevant information in online posts using, for example, artificial intelligence methods, pose an interesting path of future research.

Limitations

Despite carefully annotating the data and collecting ground truth data through a survey, our work has multiple limitations, which are briefly discussed in this section.

In our study, the investigated population has been limited to those active on the Inspire platform. However, many other online patient communities exist, and both the frequency of QoL-relevant utterances as well as their predictive power might be different.

Similarly, the fact that older posts in our dataset discuss QoL-relevant topics less often indicates that selection bias might be present in our data, as patients currently interested in QoL might be more likely to have participated in our study. Thus, our results have to be considered more like an upper bound to the possible insights about patients' QoL from posts in online forums.

Moreover, we also found signs of recall bias in the data. For example, we found that some patients talked about how much they worried about different topics shortly before participating in the study, but then answered that they did not worry at all. The occurrence of such phenomena has not been systematically investigated in this work and might thus have an impact on the results.

Certain highly intimate topics, such as aspects related to sexuality or body image, were rarely discussed in the forum posts. Given the public nature of the data, which is to be expected, however, it represents a limitation in terms of depth and scope of insights that can be obtained from such a dataset. Compared to one-on-one interviews, forum data may provide a less comprehensive picture of patients' concerns in these areas.

Finally, data sparsity and unequal distribution of posts and comments among forum users are typical problems when considering social media. As illustrated by the long-tail distribution in Figure 2, there are a few users with many posts and comments in our dataset, but the majority of users have much fewer posts available. The degree to which we can obtain QoL information is thus dependent on the posting behavior of a particular user. Statistics and numbers presented in this work thus have to be interpreted as upper bounds for the potential to extract QoL information from a randomly chosen user of the Inspire communities. Similarly, due to the general data sparsity, labels that are scarcely covered by our data sample might be present more prominently for other users, which have not been part of our study, and analogously, labels with many occurrences in our data might be more rare to find for other users.

Conclusions

This study has investigated the feasibility of gaining insights about QoL from posts of patients in online forums. Toward this goal, we have recruited patients diagnosed with breast cancer from Inspire. These patients have filled the EORTC Quality of Life Questionnaire and allowed us to analyze their posts and comments in relation to their answers given in the questionnaire. By this, we have been able to compare the information provided

XSL•FO
RenderX

in their posts with the answers provided in the questionnaires. Our key findings are that between 23% and 37.4% of patients' posts contain information that can be mapped to at least one of the questions of the questionnaire. Furthermore, we could show that the QoL information in posts is reliable, corresponding to a large extent to the answers given in the questionnaire. Our findings thus show that the information shared by patients in online posts can contribute to a better understanding of how the disease affects their QoL.

While the information provided by patients in posts shared in online forums cannot replace the collection of data via questionnaires, it represents a low-cost, low-burden alternative to collect insights about QoL effectively. Most importantly, the information in posts can help us to discover aspects related to disease burden and QoL that are not captured by existing questionnaires. This could even provide guidance on how survey instruments could be extended to cover those aspects that are most burdensome to patients.

Future work should investigate how exactly the information provided by patients online can complement existing QoL assessment methodologies. A further line of research relates to how the information can be extracted automatically, in order to avoid having to scan and read thousands of posts. Given the current state of research on information extraction, this seems to be a promising avenue to investigate. Additionally, first experiments conducted in this direction yielded promising results.

All in all, our work contributes to the foundation for a novel low-burden alternative for monitoring patients' QoL and complementing existing QoL instruments. Moreover, this opens up interesting new avenues of research on how to automate the process including using artificial intelligence methods and thus further lowering the burden for both patients and clinical staff.

Acknowledgments

The authors thank the European Organisation for Research and Treatment of Cancer for giving them the permission to use their questionnaires QLQ-C30 and BR-23 (request IDs 96009 and 102328). They also thank Antonia Frank for her help with annotating the data and developing the annotation guidelines. No generative artificial intelligence tools were used to write this paper.

Funding

This work is partially funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia under grant number NW21-059A (SAIL). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to data privacy reasons.

Authors' Contributions

DMS is the main author of all sections of this paper. He is the main author of the annotation guidelines and annotated the data, supervised the annotation process, and conducted all experiments and results presented in this paper. RS contributed to the *Methods* and *Limitations* sections as well as the annotation guideline development and annotated the data. BPHC, DK, and VK implemented the survey as well as contributed to all sections of the paper. AH helped with the survey construction as well as contributed to the *Introduction* and *Methods* sections. PC supervised the work and contributed to all sections. All authors reviewed and approved the manuscript.

Conflicts of Interest

BR-HC, DK, and VK are employees of Inspire, the online community where we conducted the survey and collected the forum data. PC is a shareholder of Semalytix, a company providing social media listening services to the life science industry. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Annotation guidelines and insights from annotation.
[DOCX File, 22 KB - cancer_v12i1e76044_app1.docx]

Multimedia Appendix 2
Distribution of annotations between labeled questions.
[DOCX File, 54 KB - cancer_v12i1e76044_app2.docx]

Checklist 1
CHERRIES checklist.

XSL•FO
RenderX

[PDF File, 73 KB - cancer_v12i1e76044_app3.pdf]

References

1. Constitution of the World Health Organization. : World Health Organization; 1948 URL: https://www.globalhealthrights.org/wp-content/uploads/2013/10/Constitution-of-the-World-Health-Organization-WHO.pdf [accessed 2026-04-06]
2. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 3;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
3. Jacobsen PB, Davis K, Cella D. Assessing quality of life in research and clinical practice. Oncology (Williston Park) 2002 Sep;16(9 Suppl 10):133-139. [Medline: 12380963]
4. Detmar SB, Aaronson NK. Quality of life assessment in daily clinical oncology practice: a feasibility study. Eur J Cancer 1998 Jul;34(8):1181-1186. [doi: 10.1016/s0959-8049(98)00018-5] [Medline: 9849476]
5. Bezjak A, Ng P, Skeel R, Depetrillo AD, Comis R, Taylor KM. Oncologists' use of quality of life information: results of a survey of Eastern Cooperative Oncology Group physicians. Qual Life Res 2001;10(1):1-13. [doi: 10.1023/a:1016692804023] [Medline: 11508471]
6. Buxton J, White M, Osoba D. Patients' experiences using a computerized program with a touch-sensitive video monitor for the assessment of health-related quality of life. Qual Life Res 1998 Aug;7(6):513-519. [doi: 10.1023/a:1008826408328] [Medline: 9737141]
7. Velikova G, Brown JM, Smith AB, Selby PJ. Computer-based quality of life questionnaires may contribute to doctor-patient interactions in oncology. Br J Cancer 2002 Jan 7;86(1):51-59. [doi: 10.1038/sj.bjc.6600001] [Medline: 11857011]
8. Giesinger J, Kemmler G, Meraner V, et al. Towards the implementation of quality of life monitoring in daily clinical routine: methodological issues and clinical implication. Breast Care (Basel) 2009;4(3):148-154. [doi: 10.1159/000224158] [Medline: 20847874]
9. Staunton H, Kelly K, Newton L, et al. A patient-centered conceptual model of symptoms and their impact in early Parkinson's disease: a qualitative study. J Parkinsons Dis 2022;12(1):137-151. [doi: 10.3233/JPD-202457] [Medline: 34657850]
10. Javidnia M, Arbatti L, Hosamath A, Eberly SW, Oakes D, Shoulson I. Predictive value of verbatim Parkinson's disease patient-reported symptoms of postural instability and falling. J Parkinsons Dis 2021;11(4):1957-1964. [doi: 10.3233/JPD-212636] [Medline: 34250951]
11. Perella C, Steenackers M, Robbins B, et al. Patient experience of Sjögren's disease and its multifaceted impact on patients' lives. Rheumatol Ther 2023 Jun;10(3):601-614. [doi: 10.1007/s40744-023-00531-7] [Medline: 36797434]
12. Cook NS, Kostikas K, Gruenberger JB, et al. Patients' perspectives on COPD: findings from a social media listening study. ERJ Open Res 2019 Feb;5(1):00128-2018. [doi: 10.1183/23120541.00128-2018] [Medline: 30775374]
13. McClellan C, Ali MM, Mutter R, Kroutil L, Landwehr J. Using social media to monitor mental health discussions - evidence from Twitter. J Am Med Inform Assoc 2017 May 1;24(3):496-502. [doi: 10.1093/jamia/ocw133] [Medline: 27707822]
14. Karmen C, Hsiung RC, Wetter T. Screening Internet forum participants for depression symptoms by assembling and enhancing multiple NLP methods. Comput Methods Programs Biomed 2015 Jun;120(1):27-36. [doi: 10.1016/j.cmpb.2015.03.008] [Medline: 25891366]
15. Zivanovic S, Martinez J, Verplanke J. Capturing and mapping quality of life using Twitter data. GeoJournal 2020 Feb;85(1):237-255. [doi: 10.1007/s10708-018-9960-6]
16. Cotté FE, Voillot P, Bennett B, et al. Exploring the health-related quality of life of patients treated with immune checkpoint inhibitors: social media study. J Med Internet Res 2020 Sep 11;22(9):e19694. [doi: 10.2196/19694] [Medline: 32915159]
17. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
18. Arraras JI, Manterola A, Asin G, et al. Quality of life in elderly patients with localized breast cancer treated with radiotherapy. A prospective study. Breast 2016 Apr;26:46-53. [doi: 10.1016/j.breast.2015.12.008] [Medline: 27017241]
19. Carreira H, Williams R, Dempsey H, Stanway S, Smeeth L, Bhaskaran K. Quality of life and mental health in breast cancer survivors compared with non-cancer controls: a study of patient-reported outcomes in the United Kingdom. J Cancer Surviv 2021 Aug;15(4):564-575. [doi: 10.1007/s11764-020-00950-3] [Medline: 33089480]
20. Mols F, Vingerhoets AJJM, Coebergh JW, van de Poll-Franse LV. Quality of life among long-term breast cancer survivors: a systematic review. Eur J Cancer 2005 Nov;41(17):2613-2619. [doi: 10.1016/j.ejca.2005.05.017] [Medline: 16226458]
21. Montazeri A. Health-related quality of life in breast cancer patients: a bibliographic review of the literature from 1974 to 2007. J Exp Clin Cancer Res 2008 Aug 29;27(1):32. [doi: 10.1186/1756-9966-27-32] [Medline: 18759983]
22. Nolte S, Liegl G, Petersen MA, et al. General population normative data for the EORTC QLQ-C30 health-related quality of life questionnaire based on 15,386 persons across 13 European countries, Canada and the Unites States. Eur J Cancer 2019 Jan;107:153-163. [doi: 10.1016/j.ejca.2018.11.024] [Medline: 30576971]
23. Hinz A, Singer S, Brähler E. European reference values for the quality of life questionnaire EORTC QLQ-C30: results of a German investigation and a summarizing analysis of six European general population normative studies. Acta Oncol 2014 Jul;53(7):958-965. [doi: 10.3109/0284186X.2013.879998] [Medline: 24456505]

XSL•FO
RenderX

24. Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet e-surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
25. Grishman R, Sundheim B. Message understanding conference - 6: a brief history. Presented at: Proceedings of the 16th International Conference on Computational Linguistics (COLING 1996); Aug 5-9, 1996. [doi: 10.3115/992628.992709]
26. Li J, Sun A, Han J, Li C. A survey on deep learning for named entity recognition. IEEE Trans Knowl Data Eng 2022;34(1):50-70. [doi: 10.1109/TKDE.2020.2981314]
27. Advanced Data Mining Techniques: Springer; 2008. [doi: 10.1007/978-3-540-76917-0]
28. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull 1971 Nov;76(5):378-382. [doi: 10.1037/h0031619]
29. Hsu YH, Shen WC, Wang CH, Lin YF, Chen SC. Hand-foot syndrome and its impact on daily activities in breast cancer patients receiving docetaxel-based chemotherapy. Eur J Oncol Nurs 2019 Dec;43:101670. [doi: 10.1016/j.ejon.2019.09.011] [Medline: 31586645]
30. Naseri H, Kafi K, Skamene S, et al. Development of a generalizable natural language processing pipeline to extract physician-reported pain from clinical reports: generated using publicly-available datasets and tested on institutional clinical reports for cancer patients with bone metastases. J Biomed Inform 2021 Aug;120:103864. [doi: 10.1016/j.jbi.2021.103864] [Medline: 34265451]
31. Forsyth AW, Barzilay R, Hughes KS, et al. Machine learning methods to extract documentation of breast cancer symptoms from electronic health records. J Pain Symptom Manage 2018 Jun;55(6):1492-1499. [doi: 10.1016/j.jpainsymman.2018.02.016] [Medline: 29496537]
32. Pakhomov SVS, Shah ND, Van Houten HK, Hanson PL, Smith SA. The role of the electronic medical record in the assessment of health related quality of life. AMIA Annu Symp Proc 2011;2011:1080-1088. [Medline: 22195169]
33. Yin Z, Fabbri D, Rosenbloom ST, Malin B. A scalable framework to detect personal health mentions on Twitter. J Med Internet Res 2015 Jun 5;17(6):e138. [doi: 10.2196/jmir.4305] [Medline: 26048075]
34. Riga M, Stocker M. Atmospheric environment and quality of life information extraction from Twitter with the use of self-organizing maps. J Environ Inform 2015;26. [doi: 10.3808/jei.201500311]
35. Sarma KV, Spiegel BMR, Reid MW, et al. Estimating the health-related quality of life of Twitter users using semantic processing. Stud Health Technol Inform 2019 Aug 21;264:1065-1069. [doi: 10.3233/SHTI190388] [Medline: 31438088]
36. Zhang X, Athanasiadou R, Razavian N. Tracing state-level obesity prevalence from sentence embeddings of tweets: a feasibility study. Presented at: Heterogeneous Data Management, Polystores, and Analytics for Healthcare (DMAH 2020, Poly 2020); Aug 31 to Sep 4, 2020. [doi: 10.1007/978-3-030-71055-2_12]
37. Shankar R, Xu Q, Bundele A. Patient voices in dialysis care: sentiment analysis and topic modeling study of social media discourse. J Med Internet Res 2025 May 15;27:e70128. [doi: 10.2196/70128] [Medline: 40372782]
38. Leung YW, Wouterloot E, Adikari A, et al. Natural language processing-based virtual cofacilitator for online cancer support groups: protocol for an algorithm development and validation study. JMIR Res Protoc 2021 Jan 7;10(1):e21453. [doi: 10.2196/21453] [Medline: 33410754]
39. Leung YW, Park B, Heo R, et al. Providing care beyond therapy sessions with a natural language processing-based recommender system that identifies cancer patients who experience psychosocial challenges and provides self-care support: pilot study. JMIR Cancer 2022 Jul 29;8(3):e35893. [doi: 10.2196/35893] [Medline: 35904877]
40. Fu J, Li C, Zhou C, et al. Methods for analyzing the contents of social media for health care: scoping review. J Med Internet Res 2023 Jun 26;25:e43349. [doi: 10.2196/43349] [Medline: 37358900]
41. Syntosi A, Felizzi F, Bouchet C. A social media listening study to understand the unmet needs and quality of life in adult and pediatric amblyopia patients. Ophthalmol Ther 2022 Dec;11(6):2183-2196. [doi: 10.1007/s40123-022-00571-2] [Medline: 36175822]
42. Findley A, Sharma G, Bentley S, et al. Comparison of literature review, social media listening, and qualitative interview research methods in generating patient-reported symptom and functional impact concepts of presbyopia. Ophthalmol Ther 2023 Feb;12(1):501-516. [doi: 10.1007/s40123-022-00620-w] [Medline: 36502495]
43. Cook N, Mullins A, Gautam R, et al. Evaluating patient experiences in dry eye disease through social media listening research. Ophthalmol Ther 2019 Sep;8(3):407-420. [doi: 10.1007/s40123-019-0188-4] [Medline: 31161531]
44. Goel K, Parashar N, Aasaithambi S, Verma H, O'Brien P, Sloesen B. P33 patient experiences and insights on chronic ocular pain from a social media listening study. Value Health 2022 Jan;25(1):S8. [doi: 10.1016/j.jval.2021.11.032]
45. Spies E, Andreu T, Koelling J, Hartung M, Kamudoni P, Park J. PCR251 retrospective social listening study of patients living with systemic lupus erythematosus (SLE): understanding the patient experience. Value Health 2022 Dec;25(12):S438. [doi: 10.1016/j.jval.2022.09.2184]
46. Tyagi N, Aasaithambi S, Chauhan J, George A, Zaour N. PUK32 patient insights for immunoglobulin A nephropathy (IgAN) using social media listening. Value Health 2019 Nov;22:S919. [doi: 10.1016/j.jval.2019.09.2716]
47. Tyagi N, Aasaithambi S, Chauhan J, George A, Zaour N. PUK33 patient insights for complement 3 glomerulopathy (C3G) using social media listening. Value Health 2019 Nov;22:S919. [doi: 10.1016/j.jval.2019.09.2717]
48. Tadmouri A, Alivon M, Andreu T, et al. RWD141 exploration of melanoma patient-generated real-world data using an AI-based social listening approach. Value Health 2022 Dec;25(12):S476-S477. [doi: 10.1016/j.jval.2022.09.2366]

XSL•FO
RenderX

49. Wolffsohn JS, Leteneux-Pantais C, Chiva-Razavi S, et al. Social media listening to understand the lived experience of presbyopia: systematic search and content analysis study. J Med Internet Res 2020 Sep 21;22(9):e18306. [doi: 10.2196/18306] [Medline: 32955443]
50. Voillot P, Riche B, Portafax M, et al. Social media platforms listening study on atopic dermatitis: quantitative and qualitative findings. J Med Internet Res 2022 Jan 28;24(1):e31140. [doi: 10.2196/31140] [Medline: 35089160]
51. Rodrigues A, Chauhan J, Sagkriotis A, Aasaithambi S, Montrone M. Understanding the lived experience of lung cancer: a European social media listening study. BMC Cancer 2022 Apr 30;22(1):475. [doi: 10.1186/s12885-022-09505-4] [Medline: 35490223]
52. Gries KS, Fastenau J. Using a digital patient powered research network to identify outcomes of importance to patients with multiple myeloma. J Patient Rep Outcomes 2020 Sep 1;4(1):74. [doi: 10.1186/s41687-020-00242-5] [Medline: 32870420]
53. Yang YC, Al-Garadi MA, Bremer W, Zhu JM, Grande D, Sarker A. Developing an automatic system for classifying chatter about health services on Twitter: case study for Medicaid. J Med Internet Res 2021 May 3;23(5):e26616. [doi: 10.2196/26616] [Medline: 33938807]
54. Tapi Nzali MD, Bringay S, Lavergne C, Mollevi C, Opitz T. What patients can tell us: topic analysis for social media on breast cancer. JMIR Med Inform 2017 Jul 31;5(3):e23. [doi: 10.2196/medinform.7779] [Medline: 28760725]
55. Renner S, Marty T, Khadhar M, et al. A new method to extract health-related quality of life data from social media testimonies: algorithm development and validation. J Med Internet Res 2022 Jan 28;24(1):e31528. [doi: 10.2196/31528] [Medline: 35089152]
56. Ware Jr JE. SF-36 health survey. In: The Use of Psychological Testing for Treatment Planning and Outcomes Assessment, 2nd edition: Lawrence Erlbaum Associates; 1999:1227-1246.
57. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
58. Taylor VR. Measuring healthy days: population assessment of health-related quality of life. : Centers for Disease Control and Prevention (CDC); 2000 URL: https://stacks.cdc.gov/view/cdc/6406
59. Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol 1993 Mar;11(3):570-579. [doi: 10.1200/JCO.1993.11.3.570] [Medline: 8445433]
60. Matsuda S, Ohtomo T, Tomizawa S, et al. Incorporating unstructured patient narratives and health insurance claims data in pharmacovigilance: natural language processing analysis of patient-generated texts about systemic lupus erythematosus. JMIR Public Health Surveill 2021 Jun 29;7(6):e29238. [doi: 10.2196/29238] [Medline: 34255719]
61. Falissard B, Simpson EL, Guttman-Yassky E, et al. Qualitative assessment of adult patients' perception of atopic dermatitis using natural language processing analysis in a cross-sectional study. Dermatol Ther (Heidelb) 2020 Apr;10(2):297-305. [doi: 10.1007/s13555-020-00356-0] [Medline: 32006346]
62. Tejaswini V, Sathya Babu K, Sahoo B. Depression detection from social media text analysis using natural language processing techniques and hybrid deep learning model. ACM Trans Asian Low Resour Lang Inf Process 2024 Jan 31;23(1):1-20. [doi: 10.1145/3569580]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
EORTC: European Organisation for Research and Treatment of Cancer
FN: false negatives
FP: false positives
QoL: quality of life
TN: true negatives
TP: true positives

Edited by M Balcarras; submitted 15.Apr.2025; peer-reviewed by A AL-Asadi, Á García-Barragán, T Escudier; revised version received 27.Jan.2026; accepted 23.Feb.2026; published 30.Apr.2026.

Please cite as:
Schmidt DM, Schubert R, Chen BPH, Kuk D, Kudesia V, Hinz A, Cimiano P
Extracting Quality of Life Information of Patients Diagnosed With Breast Cancer From Health Care Online Forum Posts: Data Feasibility Study
JMIR Cancer 2026;12:e76044
URL: https://cancer.jmir.org/2026/1/e76044
doi:10.2196/76044

XSL•FO
RenderX

© David Maria Schmidt, Raoul Schubert, Brian Po-Han Chen, Deborah Kuk, Valmeek Kudesia, Andreas Hinz, Philipp Cimiano. Originally published in JMIR Cancer (https://cancer.jmir.org), 30.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Feasibility of Implementing an Electronic Patient-Reported Outcome-Based Digital Health Platform in Pediatric Oncology and Palliative Care: Insights From the MyPal4Kids Multicenter Observational Study

Marcel Meyerheim[1], MSc; Lenka Radová[2], PhD; Anna Burns-Gebhart[3], MD; Marvin Mergen[1], Dr med; Pantelis Natsiavas[4], PhD; Panos Bonotis[4], MSc; Christina Karamanidou[4], PhD; Sheila Payne[5], Prof PhD; Julia Downing[6], Prof PhD; Julie Ling[7], PhD; Annette Sander[3], Dr med; Kasra Mirzaie[3], MEd; Susan Dokutur[3]; Jana Vaculíková[8], MSc; Petr Lokaj[8], MD; Ondřej Rohleder[8], MD; Norbert Graf[1], Prof Dr med

[1]Department of Pediatric Oncology and Hematology, Faculty of Medicine, Saarland University, Building 9, Homburg, Germany

[2]Center of Molecular Medicine, Central European Institute of Technology, Masaryk University, Brno, Czech Republic

[3]Department for Pediatric Oncology and Hematology, Hannover Medical School, Hannover, Germany

[4]Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki, Greece

[5]International Observatory on End of Life Care, Division of Health Research, Lancaster University, Lancaster, United Kingdom

[6]International Children's Palliative Care Network, Bristol, United Kingdom

[7]European Association for Palliative Care, Vilvoorde, Belgium

[8]Department of Pediatric Oncology, University Hospital Brno, Brno, Czech Republic

Corresponding Author:
Marcel Meyerheim, MSc
Department of Pediatric Oncology and Hematology, Faculty of Medicine, Saarland University, Building 9, Homburg, Germany

Abstract

Background: Digital health solutions that incorporate electronic patient-reported outcomes (ePROs) hold promise for enhancing communication and patient engagement in pediatric oncology and palliative care. While ePROs are increasingly used in adult populations, their use in children, especially when combined with gamification, remains underexplored.

Objective: This MyPal4Kids study aimed to assess the feasibility, acceptability, and user engagement of a digital health platform that integrates ePROs and a serious game, designed for pediatric oncology patients aged 6 - 17 years and their caregivers.

Methods: A multicenter observational feasibility study was conducted between December 2020 and September 2022 at 3 clinical sites in Germany and the Czech Republic. A total of 83 children and adolescents with cancer, their parent or legal guardian, as well as 10 health care professionals (HCPs) participated. Primary outcomes included platform acceptability and engagement, assessed through recruitment, participation, and attrition rates, adherence, and usability ratings. Secondary outcomes focused on the feasibility of ePRO-based data capture and the perceived impact on HCP workflows and communication. A mixed methods design was used, combining in-app data and surveys with qualitative insights from focus groups.

Results: The recruitment rate was 55%, with an attrition rate of 18%. Usability was rated positively, with most users finding the platform intuitive. Sustained engagement declined over time, particularly among older children who found the serious game insufficiently engaging. Adherence varied by age group and HCP involvement. Participants reported increased motivation when symptom reports were acknowledged by HCPs, although the impact on communication was perceived inconsistently.

Conclusions: The developed platform shows potential for use in clinical care and research. However, maintaining long-term engagement remains a challenge. Tailoring content to different age groups and strengthening feedback mechanisms from HCPs are critical for improving the user experience. Successful implementation in routine clinical practice will require integration into existing workflows and digital infrastructure, along with ongoing user-driven development.

Trial Registration: DRKS DRKS00021458; https://drks.de/search/en/trial/DRKS00021458/details and ClinicalTrials.gov NCT04381221; https://clinicaltrials.gov/study/NCT04381221 and Universal Trial Number: U1111-1251-0043

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2020-045226

(JMIR Cancer 2026;12:e80718) doi:10.2196/80718

XSL•FO
RenderX

KEYWORDS

electronic patient-reported outcomes; digital health; serious game; gamification; pediatric oncology; palliative care; patient engagement; mHealth; mobile health

Introduction

The illness trajectory of children and adolescents with cancer presents a heavy burden for patients and their families, marked by diagnosis, the difficult treatment process, and fear of relapse or death. Physical and psychological challenges affect patients, relatives, and peers [1-3]. This underscores the need for comprehensive support, including palliative care, which should begin at diagnosis, not just at the end of life [4].

Recognizing this, pediatric oncology has shifted toward investigating patient-reported outcomes, typically collected via standardized questionnaires. Electronic patient-reported outcomes (ePROs) are digital versions, often available through mobile apps [5-7]. This patient-centered approach allows young patients with cancer to self-report experiences, preferences, concerns, and symptoms in real time [8], without relying on health care professionals (HCPs) [9]. Such patient-centered approaches aim to improve self-efficacy and health-related quality of life (QoL) for young patients with cancer and their relatives [9,10]. They also aim to improve communication among all stakeholders [11], addressing time constraints during appointments and providing continuous, nuanced information on patients' well-being [12].

While ePROs have demonstrated utility as outcome measures in clinical trials involving childhood cancer and palliative care [6,13-17], their implementation in daily clinical practice remains limited and underexplored [13,18-22]. Expert guidance also highlights the need for carefully designed development and implementation processes for pediatric ePRO systems to ensure usability, clinical relevance, and successful integration into routine care [23-25].

Approaches such as serious games [26] and gamification [27-29] have gained increasing attention as strategies to enhance engagement with digital health interventions, particularly in younger populations [30-33]. Gamification elements (eg, rewards, narratives, or interactive tasks) may foster motivation, adherence, and sustained use of health apps, thereby potentially improving data completeness and the overall effectiveness of ePRO systems in supporting coping, communication, and self-expression. However, the use of gamification to engage young patients is still poorly understood [18,34-40].

At the same time, the application of digital health tools in pediatric oncology requires careful consideration of the heterogeneity of the target population. Children and adolescents differ widely in terms of cognitive development, emotional needs, and digital literacy, and their experiences of cancer vary depending on diagnosis, treatment intensity, and prognosis [1-3]. Moreover, the involvement of parents or caregivers plays a central role in treatment-related decision-making, distinguishing pediatric settings from adult care, while also being associated with increased levels of distress among caregivers [41,42]. These factors highlight the importance of designing flexible, developmentally appropriate platforms that can accommodate both self-reporting by patients and proxy reporting by caregivers, while remaining sensitive to the diverse clinical and psychosocial contexts within pediatric oncology.

The European Union-funded MyPal project (grant 825872 [43]) addressed these challenges by developing and evaluating a novel digital health platform under real-world conditions. The MyPal digital health platform and its components were developed by the MyPal project partners Foundation for Research and Technology-Hellas and Promotion Software GmbH. It encompassed two clinical studies: MyPal Adults (adult patients with cancer) [44] and MyPal4Kids (children and adolescents with cancer) [45]. MyPal4Kids aimed to evaluate the acceptability, appropriateness, and impact of an ePRO-based platform for pediatric oncology and palliative care, as perceived by children and adolescents with cancer, parents, and HCPs. The study also sought to identify barriers and facilitators relevant to the platform's integration into routine care. An interdisciplinary team, including experts from pediatric oncology, palliative care, psychology, IT, and software development, oversaw all stages from design to implementation and evaluation.

Previous publications from the MyPal consortium have addressed key feasibility aspects [46], including a systematic review on ePRO interventions in palliative care [47], ethical and legal considerations [48], implementation frameworks [49], technical development [50], psycho-oncological aspects [51], and policy implications [52]. This paper presents the core findings of the MyPal4Kids clinical study, combining quantitative and qualitative data to evaluate whether integrating ePROs via a digital platform can support more personalized and sustainable care in pediatric oncology and palliative care.

Methods

Overview

The multicenter observational feasibility study, MyPal4Kids, was conducted from December 2020 to September 2022 at Saarland University (USAAR), Hannover Medical School (MHH; Germany), and University Hospital Brno (FNBRNO; Czech Republic). Recruitment closed on March 31, 2022 (last patient, first visit).

Study Design

An observational study design was chosen to evaluate the feasibility of the platform in a real-world setting without randomization, enabling a naturalistic assessment of acceptability and engagement.

Setting and Participants

The study took place in pediatric oncology units at all 3 participating hospitals, which provide inpatient and outpatient care to about 290 new pediatric oncology patients annually, supported by multidisciplinary teams. Eligible participants were children and adolescents with cancer aged 6-17 years diagnosed

XSL•FO
RenderX

with solid tumors or leukemia and treated at one of the sites within the previous 12 months.

Recruitment

Trained clinical staff recruited participants. Eligible patients and at least 1 parent or legal guardian were invited to participate in the 6-month study during an informational session. Written informed consent was obtained from patients and their parents. Study staff created user accounts, provided leaflets with QR codes for app installation and baseline questionnaires, and offered support as needed. Details of recruitment and eligibility are available in the protocol [45].

Intervention: The MyPal Digital Health Platform

In addition to standard care, participants used the ePRO-based platform during their 6-month study participation. The platform had 3 components (Table 1) [50], which were available in German and Czech with a "read-aloud" functionality for younger users. A video demonstrating the MyPal platform and its apps is available online [53]. The platform was implemented through these 3 components and was offered for use within clinical workflows across user-selected devices. Participation was voluntary, and no mandatory actions or response pathways were established for HCPs. Data were securely transmitted and made available to HCPs to support clinical decision-making.

Table . A description of the 3 components of the MyPal digital health platform evaluated in MyPal4Kids.

Component	Description
MyPal-Child App	A mobile video game designed for young patients with cancer that integrates ePRO[a]-based symptom self-reporting and diary entries into the serious game "AquaScouts." [50] The game immerses players in an underwater adventure in which infected areas are progressively restored through the gameplay. Players collect tokens while responding to prioritized symptom-related questions presented both inside and outside the game environment. Diary entries could be optionally shared with treating HCPs[b] through the App. Gamification elements included obstacle avoidance mechanics associated with point deductions, selectable movement speeds, token-based rewards (eg, artifacts), daily high-score lists among study participants, and avatar customization intended to support motivation and engagement.
MyPal-Carer App	Enables parents to proxy-report their child's symptoms and access relevant web links. Both this and the MyPal-Child App could be installed on mobile devices and were also used to complete study questionnaires.
Browser-based HCP interface	A web-accessible tool (via smartphone, tablet, or desktop) that allows real-time review of reported symptoms and diary entries through interactive visualizations. HCPs could add treatment protocols, symptoms, and reactions based on observed data. The study protocol did not prescribe specific HCP responses. The interface was used on a discretionary basis.

[a]ePRO: electronic patient-reported outcome.

[b]HCP: health care professional.

Participatory Design

A participatory design approach guided the development of the platform with the aim to enhance its relevance, utility, and usability for the target population. This process included a literature review, prestudy focus groups [54], and structured discussions with patients, parents, and HCPs to identify needs, preferences, potential barriers, and desired functionalities. These activities served as a prestudy assessment to inform key design decisions and ensure that the platform was aligned with user requirements prior to implementation. Development followed an iterative feedback loop and included a 2-week pilot phase with children prior to study initiation to further refine the platform, identify technical issues, evaluate the local technical infrastructure, and obtain initial insights before the feasibility study commenced [50].

Outcomes Measures

Participants were invited to use the platform and its apps for 6 months. Table 2 summarizes the study end points, outcome measures, and time points adapted from the published study protocol [45]. In addition to optional symptom reporting, patients and parents completed ePRO-based secondary end point measures via the apps. The questionnaires were selected based on validity, dissemination, digital availability, and author agreement (refer to Table 2 for end points and the assessment schedule and S5 in Multimedia Appendix 1 for questionnaire dimensions, rating scales, and scoring information).

XSL•FO
RenderX

Table . The study end points and outcome measures in MyPal4Kids.

Outcome measures and ePROs[a]	Time point of measure
Primary end points	
Acceptability of and engagement with the MyPal Platform	
Recruitment rate, participation rate, attrition rate, and adherence to the different components of the MyPal platform	M1-M6[b] continuously
System Usability (validated System Usability Scale [SUS] [55]) and a self-adapted version for children (S1 in Multimedia Appendix 1, child-adapted SUS version)	Once after M6
Focus groups with children and their parents	Once after M6
Secondary end points	
Demonstrating the feasibility of measuring	
Children's symptom burden (digital adaptation of the validated Mini-SSPedi[c] [56] and SSPedi [57]): higher ratings indicate higher burden	M1-M6 optional anytime
Children's QoL[d] (validated PedsQL[e] 3.0 Cancer Module [58,59]): higher scores indicate fewer problems and better QoL	Monthly (Baseline, M1-M6)
Parents' satisfaction with cancer care (validated EORTC PATSAT-C33[f] [60,61], adapted to assess parent's satisfaction with children's cancer care): higher scores indicate greater satisfaction with care	Monthly (Baseline, M1-M6)
Impact of pediatric illness on the family (validated IOFS[g] [62,63]): higher scores indicate greater negative impact on family	Monthly (Baseline, M1-M6)
Parents' QoL (EuroQol EQ-5D-3L [64]): higher visual analog scale scores indicate better perceived health	Monthly (Baseline, M1-M6)
Focus groups with children and their parents	Once after M6
Impact on HCPs[h] due to the integration of ePROs in care	
Web-based follow-up questionnaire (S2 in Multimedia Appendix 1, questionnaire for HCPs)	Once after M6
Focus groups with project-internal and external HCPs involved in the study	Once after M6

[a]ePRO: electronic patient-reported outcome.

[b]M[X]: Month X of study participation.

[c]SSPedi: Symptom Screening in Pediatrics.

[d]QoL: quality of life.

[e]PedsQL: Pediatric Quality of Life Inventory.

[f]EORTC PATSAT-C33: European Organisation for Research and Treatment of Cancer Satisfaction with Cancer Care Core Questionnaire.

[g]IOFS: Impact on Family Scale.

[h]HCP: health care professional.

Per protocol, the recruitment rate was defined as the proportion of recruited patients among eligible patients. The participation rate referred to recruited participants completing at least 70% of the 28 monthly questionnaires (before imputation). The attrition rate included patients who did not complete the 6-month study due to withdrawal or death. Reasons for missed assessments (multiple-choice and free-text responses) were documented when available.

Adherence was assessed by user group. For patients, adherence included the number of optional monthly symptom reports, diary entries, and in-game personalization activities. For parents, adherence included the number of monthly optional proxy symptom reports. For HCPs, adherence included the number of monthly treatment plans, reaction protocols, and proxy symptom reports. Engagement was defined as the frequency and consistency of platform use over time. Adherence referred to the completion of scheduled questionnaires, while participation described overall involvement in study activities.

Scoring and missing data handling followed the respective questionnaire guidelines (refer to Table 2). Additional details, including the full questionnaires, interview guides, coding schemes, paraphrased statements, and extended statistical outputs, are provided in Multimedia Appendix 1.

Follow-Up Focus Groups

Focus groups were conducted at all three sites with (1) children and adolescents with cancer, their parents and (2) HCPs to obtain

XSL•FO
RenderX

qualitative feedback on the platform (eg, experiences, challenges, barriers, and facilitators).

Semistructured interview guides, based on the mHealth App Usability Questionnaire [65], included prioritized, neutral, open-ended, unambiguous questions (refer to S3 in Multimedia Appendix 1 for patient and parent focus groups and S4 in Multimedia Appendix 1 for focus groups with HCPs). Patients and parents were approached post study. HCPs were divided into two subgroups: (1) project-internal HCPs who provided feedback during the platform's development and (2) project-external HCPs who were involved only in the study.

Written informed consent was obtained before interviews (conducted in person or online and in the local language). Sessions lasted 45 - 60 minutes and were moderated by trained local researchers. Audio recordings were transcribed verbatim, pseudonymized, and translated into English for analysis.

Data and Statistical Analysis

Statistical analyses were conducted in R (version 4.3.2, R Core Team, R Foundation for Statistical Computing) [66]. Descriptive statistics summarized recruitment and baseline characteristics (frequency, mean [SD], or median [IQR], as appropriate). Tests were chosen based on variable distribution. Two-sided *P* values <.05 were considered significant.

Missing data were imputed according to questionnaire guidelines. For symptom reports, the highest monthly symptom burden per patient was considered. Missing responses (≤2 per measure) were replaced by the nearest previous value. The target sample size was 100 participants [45]. To accommodate missing data and the small sample size, repeated measures (baseline and M1-M6) were grouped into 4 time points (baseline, M1-M2, M3-M4, and M5-M6), ensuring no significant differences between merged time points via the two 1-sided tests, with odds ratios calculated as effect sizes.

The aligned rank transformation (ARTool; developed by Kay et al. [67]) enabled a nonparametric 2-factor repeated measures ANOVA for the factors of time, group, and their interaction effects [68]. Participants were grouped by (1) birth sex (female or male), (2) age (6-9, 10-13, and 14-17 years), (3) diagnosis (leukemia or solid tumor), and (4) Karnofsky Performance Status (KPS) [69,70] scale (below or above the cohort median). Partial eta squared (η_p^2) was classified as small ($\eta_p^2 \geq 0.01$), medium ($\eta_p^2 \geq 0.06$), and large effect size ($\eta_p^2 \geq 0.16$). Post hoc comparisons used Wilcoxon signed-rank tests (time points) and Mann-Whitney *U* tests (group comparisons), both with Holm-Bonferroni correction. For significant interactions, rank-transformed time-point differences between groups were compared using Mann-Whitney *U* tests with Holm-Bonferroni correction.

Spearman rank correlation (ρ) assessed monotonic associations between patient self-reports and parent proxy reports of symptom burden per month (no imputation). Effect sizes were classified as weak ($|\rho| \geq 0.1$), moderate ($|\rho| \geq 0.3$), and strong ($|\rho| \geq 0.5$).

Qualitative focus group data were analyzed in MAXQDA 2022 (VERBI Software) using deductive and inductive approaches [71]. Two researchers independently reviewed transcripts and developed a consensus coding system for thematic analysis. Statements were cross-coded when applicable (eg, positive, neutral, and negative) and paraphrased after joint review.

Missing data handling, subgroup analyses, and model specifications were performed as described in this section. No formal adjustment for site-level differences was performed because of the limited sample size at each site.

Ethical Considerations

The MyPal4Kids study protocol has obtained ethical approval from the ethics committee competent for the Saarland University (Ärztekammer des Saarlandes; reference: Ha 23/20; March 16, 2020), the ethics committee of the Hannover Medical School (reference: 9095_BO_K_2020; February 12, 2020), and the ethics committee of the University of Brno (reference: 01 - 1 20 220/EK; May 11, 2020). This study was performed in line with the principles of the Declaration of Helsinki. Ethical and General Data Protection Regulation (GDPR) compliance considerations ensured data safety for this vulnerable participant group [48], as reflected in the protocol through age-appropriate information sheets and informed consent forms [45]. Informed consent was obtained from all participants and from parents or legal guardians of children or adolescents.

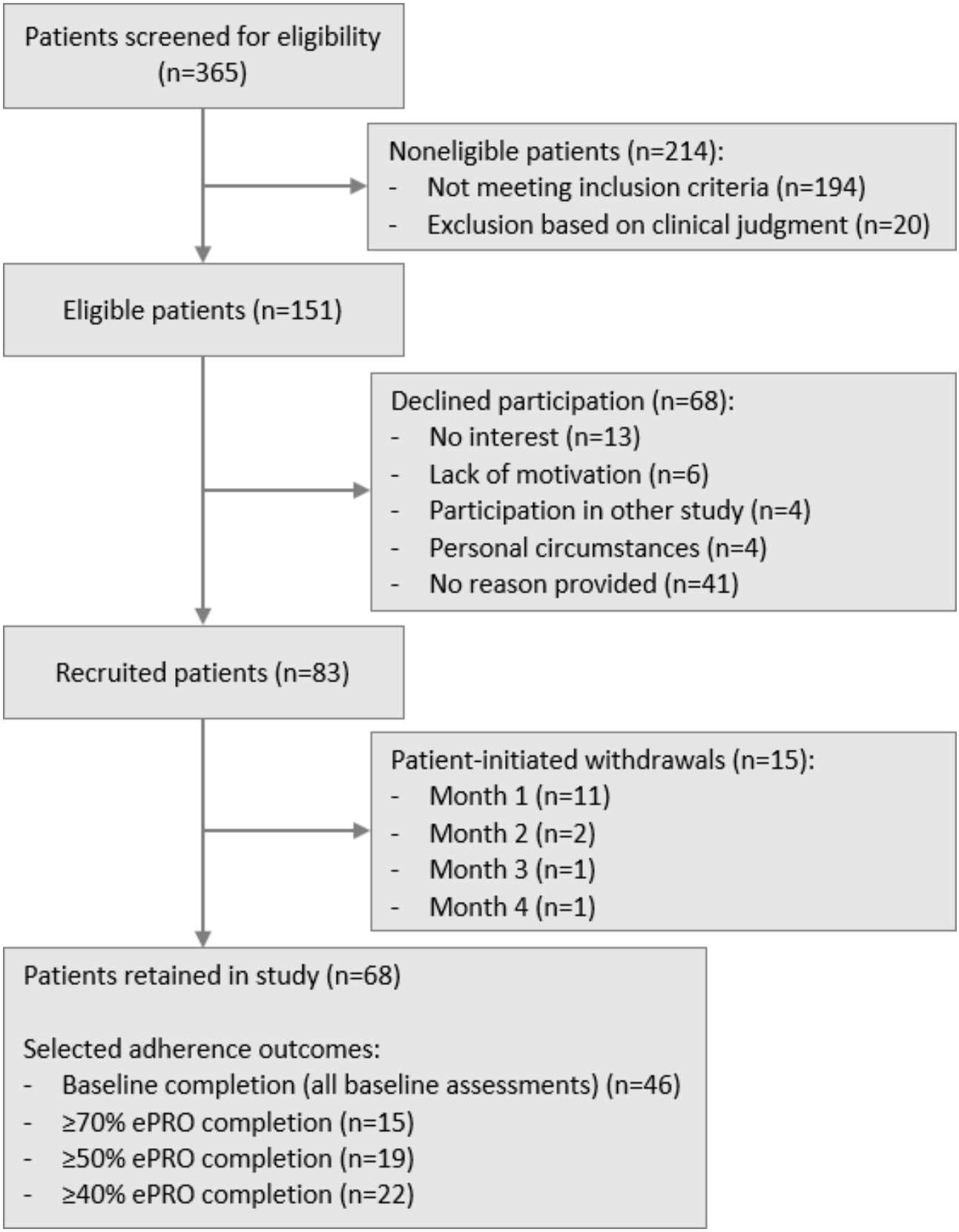
Results

Recruitment

Recruitment took place over 15 months (December 2020-March 2022), including a 6-month COVID-19 extension. A total of 83 participants were enrolled (refer to Figure 1).

XSL•FO
RenderX

Figure 1. Flow diagram on screening, eligibility, and recruitment in MyPal4Kids. ePRO: electronic patient-reported outcome.



Enrollment Characteristics

Table 3 summarizes the characteristics of the enrolled patients. Birth sex distribution and treatment status were nearly balanced. Almost half of the patients were aged 14 - 17 years. Solid tumors accounted for nearly half of the cases, with brain tumors listed separately due to their potential neurological effects [72]. Most patients were newly diagnosed (70/83, 84%) and 16% (13/83) of patients had relapsed.

XSL•FO
RenderX

Table . Patient information collected at baseline.

Baseline characteristic	Study cohort (N=83), n (%)
Birth sex	
Male	46 (55)
Female	37 (45)
Age (years)	
6 - 9	18 (22)
10 - 13	25 (30)
14 - 17	40 (48)
Diagnosis	
Brain tumor	14 (17)
Leukemia	29 (35)
Solid tumor	40 (48)
Stage	
New diagnosis	70 (84)
Relapse	13 (16)
On treatment	
Yes	45 (54)
No	38 (46)
KPS[a]	
Documented[b,c]	55 (66)
Not documented[d]	28 (34)

[a]KPS: Karnofsky Performance Status.

[b]Median (IQR): 80 (80-90).

[c]Range (min-max): 50-100.

[d]Nonmandatory information.

Acceptability and Engagement With the MyPal Platform

Attrition, Participation, and Adherence Patterns

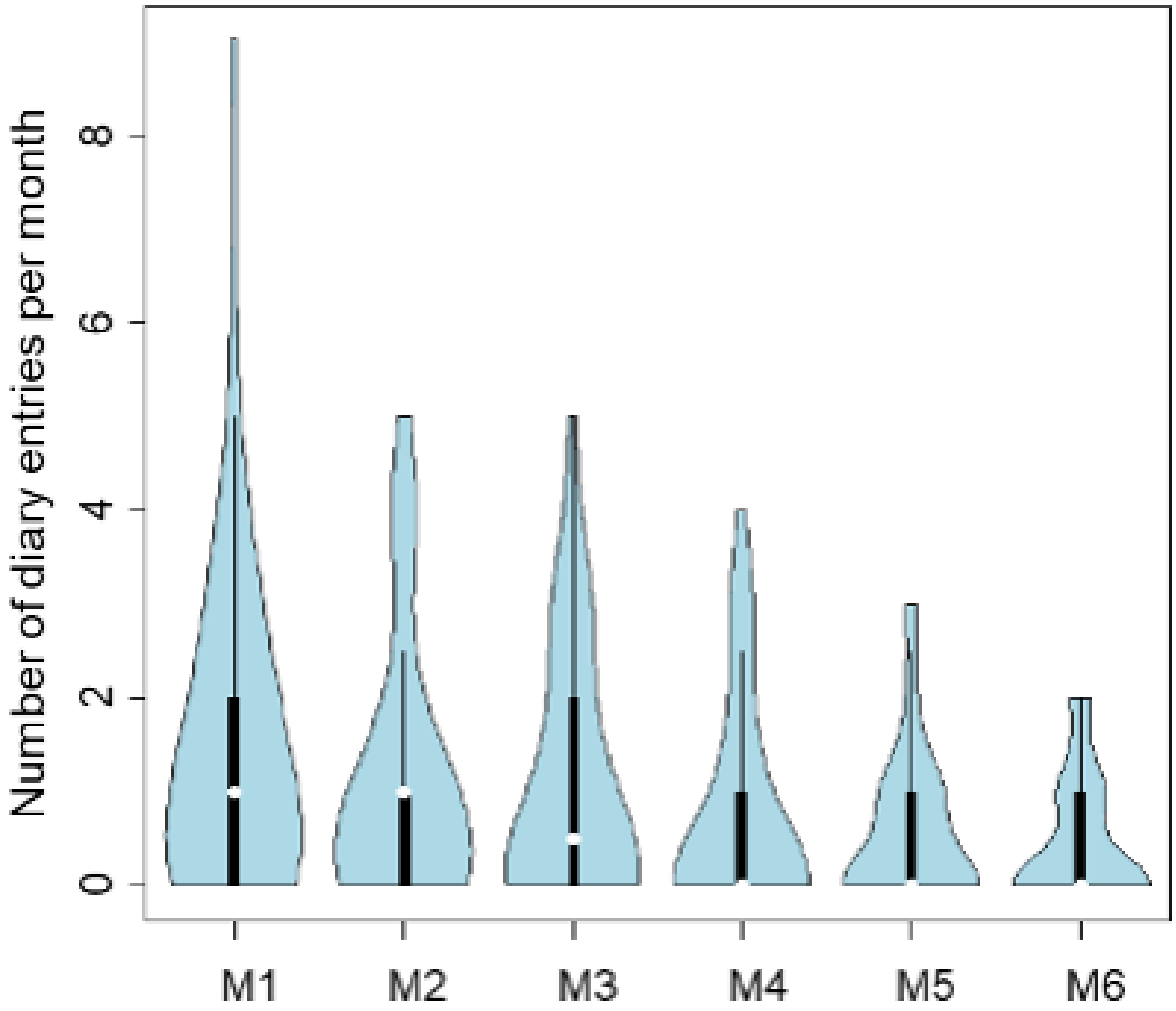
The attrition rate was 18% (15/83), with 73% of withdrawals at participants' request (11/15) occurring in the first month (M1). Withdrawn participants included 8 male and 7 female participants; 8 had solid tumors, 6 had leukemia, and 1 had a brain tumor. The age distribution among those who discontinued participation was 10 aged 14 - 17 years, 4 aged 10 - 13 years, and 1 aged 6 - 9 years. Most were newly diagnosed (n=14) and receiving treatment (n=9). Reasons for withdrawal (documented in 7 cases; multiple choice) included disease progression (n=4) and/or patient reluctance (n=6). No deaths occurred during participation.

Per protocol, 15 of 83 (18%) participants completed ≥70% of all monthly questionnaires, 19 (23%) completed ≥50%, and 22 (27%) completed ≥40%. Overall, 46 (55%) participants completed all baseline assessments (refer to Figure 1). Because these adherence thresholds were cumulative, the categories were not mutually exclusive. HCPs documented reasons for missed assessments in 32 cases (39%): patient reluctance (n=17) and disease progression (n=15).

Of 83 participants, 62 (75%) patients assessed each symptom at least once. Assessment frequency declined over time. Overall, 42 (51%) patients created at least one diary entry, with the frequency declining over time (refer to Figure 2). In total, 64 (77%) patients customized their game avatar.

XSL•FO
RenderX

Figure 2. Violin plot depicting monthly frequencies of diary entries created by patients with at least one diary entry (n=42, 51%). In all violin plots, the width of the violin reflects the frequency of the respective measure, with white dots representing the median values, thick black bars indicating the IQR, and thin black bars indicating 1.5 x the IQR for each measure. Each violin represents the distribution across the entire study cohort for the respective month.



Regarding proxy reports on symptom burden, at least 53 (64%) parents assessed each symptom at least once. HCPs created overall 98 treatment plans for 29 (35%) patients, ranging from 1 to 8 plans per patient. HCPs documented symptoms through the HCP interface 15 times for 13 (16%) patients. HCPs added reactions to reported symptoms 24 times for 15 (18%) patients.

Usability of the MyPal Platform

Sixty of 83 patients and parents (72%) completed the respective version of the System Usability Scale (SUS). Among responding patients, 30% (18/60) were younger than 10 years, whereas 70% (42/60) were 10 years or older. Younger children largely found the game enjoyable and easy, whereas older patients reported more variable ratings (refer to Table 4). A significant association with a small effect size was found between age group and game enjoyment (Q1; χ^2_1=4.4; *P*=.04; ϕ=−0.27).

XSL•FO
RenderX

Table . Answers to the adapted SUS[a] evaluating the usability of the MyPal-Child app were received from 60 of 83 patients.

Questionnaire items	Yes, n (%)	No, n (%)	Indecisive, n (%)
Adapted SUS[a] for children <10 years (n=18)			
Q1: Did you like the game?	15 (83)	3 (17)	0 (0)
Q2: Was the game easy to play?	15 (83)	2 (11)	1 (6)
Q3: Did you like to answer questions in the game?	10 (56)	8 (44)	0 (0)
Adapted SUS for children≥10 years (n=42)			
Q1: Did you like the game?	23 (55)	19 (45)	0 (0)
Q2: Was the game easy to play?	39 (93)	2 (5)	1 (2)
Q3: Did you like to answer questions in the game?	21 (50)	21 (50)	0 (0)
Q4: Were the questions asked to you relevant at the time?	25 (60)	15 (36)	2 (5)
Q5: Was the frequency of the reminders to answer questions appropriate?	24 (57)	18 (43)	0 (0)

[a]SUS: System Usability Scale.

Free-text comments from 34 (57%) patients mentioned avatar customization, points collection, and graphics as enjoyable features. Some found the game too simple and repetitive and suggested increased difficulty, customization options, and additional achievements. Parents rated the MyPal-Carer app highly for usability (mean SUS 76.8, SD 15.9; median 82.5, IQR 63.1 - 90), indicating good to excellent usability [55].

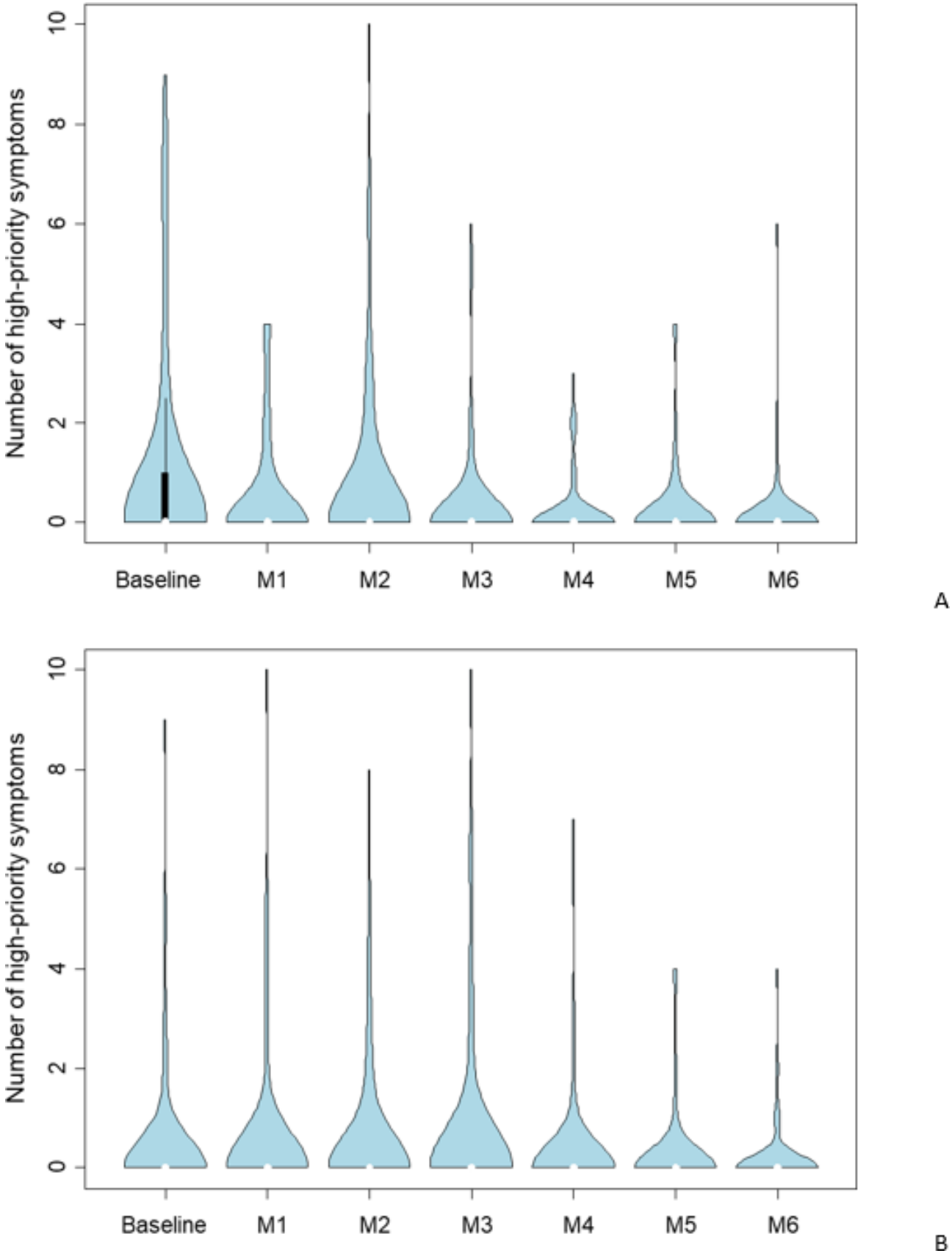
Feasibility of ePRO-Based Measures

Patient and Proxy-Reported Symptom Burden

Overall, 42 of 83 (51%) participants reported at least one symptom with the highest burden (rating 4=extremely bothered) during the 6-month period (Figure 3A). The most frequent symptoms were tiredness (29 reports), pain (other than headache; 26 reports), mouth sores (23 reports), and changes of appetite (23 reports). Less than 50% of patients reported at least one symptom with the highest burden per month.

XSL•FO
RenderX

Figure 3. Frequency distribution of monthly symptom reports with high priority (ie, the highest burden) via Symptom Screening in Pediatrics or Mini Symptom Screening in Pediatrics during study participation. Each violin represents the distribution across the entire study cohort for the respective month. (**A**) Self-reports by children and (**B**) proxy reports by parents for their children.



Overall, 34 of 83 (41%) parents submitted proxy reports (refer to Figure 3B), citing feeling disappointed or sad (24 reports), perceived changes in appearance (23 reports), and changes in appetite (23 reports) with the highest burden. Thus, some overlap existed between child and parent reports.

XSL•FO
RenderX

Spearman rank correlation across 15 symptoms over 7 time points revealed significant correlations for 8 symptoms at up to 3 time points each (refer to Table 5), ranging from ρ=0.39 to ρ=1.0 (moderate to strong), but inconsistently across time.

XSL•FO
RenderX

Table . Spearman rank correlation ρ between children's and parents' reported symptom burden via Symptom Screening in Pediatrics or Mini Symptom Screening in Pediatrics across time points.

Symptom and time point	df (n - 2)	ρ
Feeling disappointed or sad		
Baseline	24	0.18
M1	22	0.09
M2	22	0.16
M3	13	0.21
M4	13	0.14
M5	12	0.62[a]
M6	6	0.51
Feeling scared or worried		
Baseline	25	0.22
M1	21	0.48[a]
M2	22	0.46[a]
M3	12	−0.16
M4	13	0.11
M5	13	0.45
M6	6	0.59
Feeling cranky or angry		
Baseline	24	0.39[a]
M1	22	0.58[b]
M2	21	0.54[b]
M3	11	0.19
M4	14	0.16
M5	12	0.46
M6	5	0.73
Changes in how body or face look		
Baseline	22	0.14
M1	21	0.39
M2	22	0.10
M3	12	0.28
M4	14	−0.11
M5	12	0.57[a]
M6	5	0.87[a]
Feeling tired		
Baseline	21	0.22
M1	21	0.67[c]
M2	22	0.49[a]
M3	11	0.52
M4	14	0.30
M5	12	0.20

XSL•FO
RenderX

Symptom and time point	df (n - 2)	ρ
M6	4	0.67
Feeling pain		
Baseline	19	0.18
M1	20	0.47[a]
M2	21	0.26
M3	12	0.38
M4	13	0.36
M5	11	0.42
M6	4	1.0[c]
Tingly or numb hands or feet		
Baseline	19	0.20
M1	20	0.62[b]
M2	21	0.54[b]
M3	11	0.32
M4	13	0.31
M5	11	0.22
M6	4	0.50
Vomiting or nausea		
Baseline	18	0.30
M1	20	0.39
M2	20	0.28
M3	11	0.40
M4	14	−0.30
M5	11	0.71[b]
M6	4	0.72

[a]*P*<.05.
[b]*P*<.01.
[c]*P*<.001.

Subgroup and Longitudinal Differences in Symptom Burden and QoL

Table 6 summarizes results from the aligned rank transformation analysis, excluding symptoms without significant findings. KPS-based subgroups were categorized as above or below the median score of 90 (IQR 80 - 90). For example, a significant main effect of time was found for changes in taste (Figure 4A), with symptom burden decreasing from M1-M2 to M5-M6 regardless of group according to post hoc comparisons (birth sex: P=.003, η_p^2=0.151; age: P=.01, η_p^2=0.124; diagnosis: P=.007, η_p^2=0.118; KPS: P=.02, η_p^2=0.128). Regarding problems with thinking, a significant main effect of group was found, with symptom burden being higher among females than among males, regardless of time (Figure 4B).

XSL•FO
RenderX

Table . Symptoms reported via Symptom Screening in Pediatrics or Mini Symptom Screening in Pediatrics with significant results derived from the aligned rank transformation analysis.

Symptoms, groups, and subsamples	Imputed, n	Factor time		Factor group		Interaction time × group	
		P value	η_p^2	*P* value	η_p^2	*P* value	η_p^2
Feeling tired							
Birth sex	13	0.13	0.069	0.03	0.167[a]	0.74	0.016
Female (n=16)							
Male (n=12)							
More or less hungry than usual (changes in appetite)							
Birth sex	12	0.73	0.017	0.02	0.194[a]	0.74	0.019
Female (n=16)							
Male (n=11)							
Mouth sores							
Diagnosis	12	0.04	0.102[b]	0.02	0.190[a]	0.86	0.01
Leukemia (n=9)							
Solid tumor (n=18)							
Headache							
KPS[c]	10	0.004	0.183[a]	0.85	0.002	0.1	0.09
KPS ≥80 (n=11)							
KPS <80 (n=13)							
Tingly or numb hands/feet							
Diagnosis	12	0.01	0.136[b]	0.36	0.033	0.31	0.046
Leukemia (n=9)							
Solid tumor (n=18)							
Changes in how body or face look							
Birth sex	13	0.02	0.109[b]	0.29	0.041	0.41	0.035
Female (n=17)							
Male (n=12)							
Age group (y)	13	0.01	0.134[b]	0.85	0.012	0.99	0.009

XSL•FO
RenderX

Symptoms, groups, and sub-samples	Imputed, n	Factor time		Factor group		Interaction time × group	
		P value	p^2	*P* value	p^2	*P* value	p^2
6 - 9 (n=11)							
10 - 13 (n=8)							
14 - 17 (n=10)							
Diagnosis	13	0.003	0.160[a]	0.5	0.017	0.32	0.042
Leukemia (n=9)							
Solid tumor (n=20)							
KPS	9	0.02	0.144[a]	0.81	0.003	0.41	0.042
KPS ≥80 (n=11)							
KPS <80 (n=13)							
Changes in taste							
Birth sex	12	<.001	0.248[a]	0.34	0.037	0.69	0.019
Female (n=16)							
Male (n=11)							
Age group (y)	12	0.001	0.198[a]	0.87	0.011	0.83	0.038
6 - 9 (n=10)							
10 - 13 (n=8)							
14 - 17 (n=9)							
Diagnosis	12	<.001	0.284[a]	0.56	0.014	0.02	0.123[b]
Leukemia (n=9)							
Solid tumor (n=18)							
KPS	10	0.001	0.223[a]	0.54	0.017	0.7	0.021
KPS ≥80 (n=11)							
KPS <80 (n=13)							
Vomiting or nausea							
Birth sex	12	0.009	0.142[a]	0.47	0.021	0.18	0.063
Female (n=16)							
Male (n=11)							

XSL•FO
RenderX

Symptoms, groups, and sub-samples	Imputed, n	Factor time		Factor group		Interaction time × group	
		P value	p^2	*P* value	p^2	*P* value	p^2
Age group (y)	12	0.003	0.175[a]	0.76	0.021	0.04	0.161[a]
6 - 9 (n=10)							
10 - 13 (n=8)							
14 - 17 (n=9)							
Diagnosis	12	0.005	0.157[a]	0.91	0.001	0.81	0.013
Leukemia (n=9)							
Solid tumor (n=18)							
Problems with thinking and re-membering things							
Birth sex	13	0.57	0.024	0.001	0.321[a]	0.56	0.025
Female (n=17)							
Male (n=12)							
KPS	9	0.95	0.005	0.15	0.092	0.03	0.123[b]
KPS ≥80 (n=11)							
KPS <80 (n=13)							
Diarrhea							
Birth sex	12	0.08	0.084	0.16	0.076	0.03	0.108[b]
Female (n=16)							
Male (n=11)							
Age group (y)	12	0.17	0.066	0.3	0.096	0.02	0.181[a]
6 - 9 (n=10)							
10 - 13 (n=8)							
14 - 17 (n=9)							
Diagnosis	12	0.002	0.181[a]	0.27	0.048	0.009	0.143[a]
Leukemia (n=9)							
Solid tumor (n=18)							
Constipation							
Birth sex	12	0.14	0.07	0.3	0.043	0.03	0.111[b]

XSL•FO
RenderX

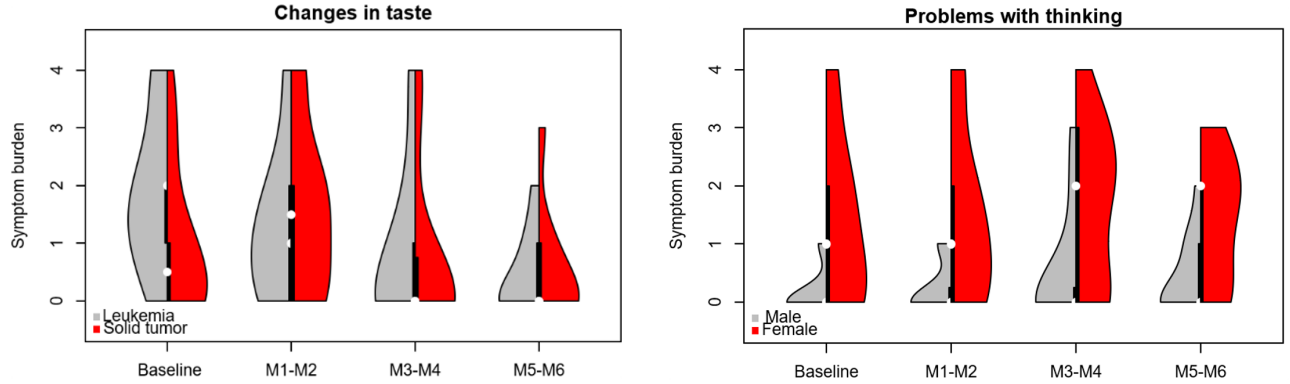
Symptoms, groups, and sub-samples	Imputed, n	Factor time		Factor group		Interaction time × group	
		P value	η_p^2	*P* value	η_p^2	*P* value	η_p^2
Female (n=17)							
Male (n=12)							

[a]Large effect size (η_p^2).

[b]Medium effect size (η_p^2).

[c]KPS: Karnofsky Performance Status.

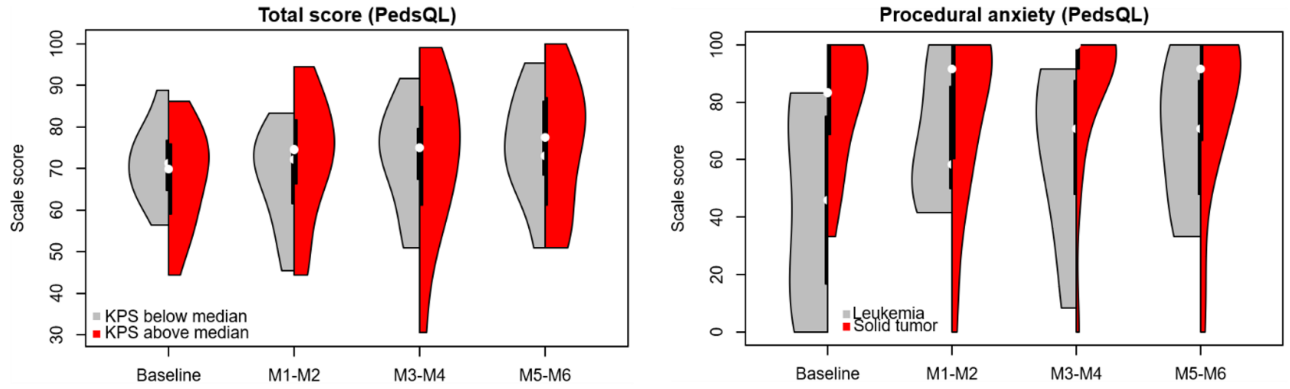
Figure 4. Violin plots depicting the distribution of maximum burden (ranging from 0=not at all bothered to 4=extremely bothered) for 2 Symptom Screening in Pediatrics or Mini Symptom Screening in Pediatrics symptoms across 4 timepoints from baseline until the final 2 months of study participation (M5-M6).



A B

Similarly, time, group, and interaction effects emerged for the other ePRO measures. For example, there was a significant main effect of time for the children's QoL total score regardless of group. Post hoc tests revealed a significant increase from baseline to M5-M6 (P=.02; η_p^2=0.126; Figure 5A). A significant main effect of group (Figure 5B) was found for procedural anxiety, regardless of time: problems with procedural anxiety were more prevalent in patients with leukemia and in the youngest age group compared with patients with solid tumors and the oldest age group (P=.009; η_p^2=0.285). Full statistical results for these ePRO measures (Pediatric Quality of Life Inventory [PedsQL], European Organisation for Research and Treatment of Cancer Satisfaction with Cancer Care Core Questionnaire [PATSAT-C33], and Impact on Family Scale [IOFS]), including subgroup analyses, are provided in S6 in Multimedia Appendix 1.

Figure 5. Violin plots depicting the distribution of 2 PedsQL dimensional scores, reported by the respective subgroups across 4 timepoints from baseline until the final 2 months of study participation (M5-M6). KPS: Karnofsky Performance Status; PedsQL: Pediatric Quality of Life Inventory.



A B

XSL•FO
RenderX

Focus Groups With Children and Their Parents

Ten parent-child pairs participated: 3 at USAAR, 4 at MHH, and 3 at FNBRNO, across all age groups: 2 aged 6 - 9 years, 2 aged 10 - 13 years, and 6 aged 14 - 17 years. The coding system linking 18 categories to statements and their frequencies is shown in S7 in Multimedia Appendix 1. Most comments concerned the categories "Recommendations," followed by "Game Design" and "ePROs." Key findings are provided in Table 7, while the full coding scheme, category frequencies, and paraphrased statements are provided in S7 in Multimedia Appendix 1.

Table . Key findings from the focus groups with children and their parents.

Code category	Summary
Game Design	Children frequently mentioned enjoying the customization features and underwater theme. Comments on gameplay included perceptions of repetitiveness and limited challenge.
ePROs[a]	Parents appreciated symptom reporting but noted some children underreported, often due to fatigue. Both parents and children found the questions understandable, but some described the questions as repetitive or less relevant during symptom-free periods.
Consequences of App Usage	The app felt most beneficial at the treatment's onset. However, engagement decreased over time. It was useful for self-reflection and documenting symptoms, particularly when direct communication with HCPs[b] was limited.
Communication	The app was not commonly perceived as improving communication but was reported as enabling some children to express feelings outside face-to-face interactions. However, infrequent feedback from HCPs in response to symptom reports led to frustration.
Recommendations	Participants suggested expanding gameplay variety, customization options, tailoring the App for different age groups, and adding multiplayer modes and peer support as social interaction features, as well as improving HCP feedback to enhance engagement.

[a]ePRO: electronic patient-reported outcome.

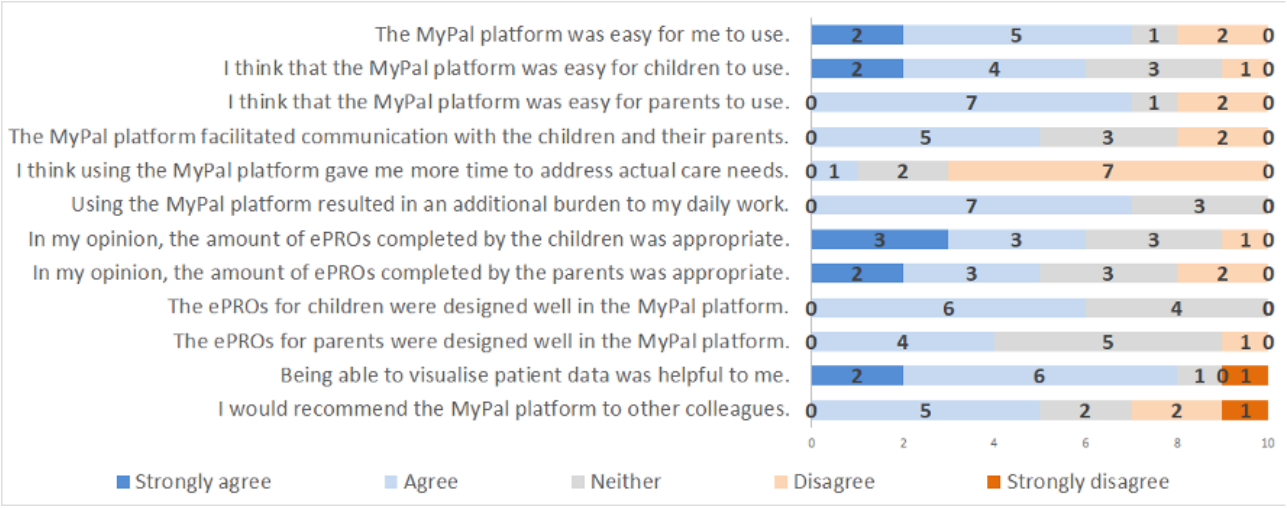
[b]HCP: health care professional.

Impact on HCPs Due to the Integration of ePROs in Care

Follow-Up Questionnaire for HCPs

Ten HCPs (5 project-internal and 5 project-external) completed the follow-up questionnaire (refer to Figure 6). No significant group differences were found. Half added free-text comments, praising online communication when in-person appointments were not possible. They valued ePROs and the idea of the platform but stressed the importance of personal contact. Concerns included missing data from disengaged participants. Recommendations included better integration with existing clinical systems, enhanced data visualization, and improved game design.

Figure 6. Distribution of the absolute frequencies of the health care professionals' responses to the items of the web-based follow-up questionnaire (n=10). ePRO: electronic patient-reported outcome.



XSL•FO
RenderX

Focus Groups With HCPs

Nine HCPs joined focus groups: 4 physicians (2 from MHH, 1 from FNBRNO, and 1 from USAAR), 1 nurse (MHH), and 1 psychologist (MHH) as project-internal, along with 1 physician (MHH), 1 psychologist (MHH), and 1 physician (FNBRNO) as project-external. The coding mirrored parent-child pair focus groups, but "Game Design" and "ePROs" were addressed less, while "Influences on Motivation," "HCP-Patient Communication," and "Compliance" had higher coverage. Compared to internal HCPs, external HCPs provided fewer statements coded as negative during focus group discussions and discussed user experiences and suggestions more than internal HCPs. Tables 8 and 9 summarize the internal and external HCP focus groups, respectively. The full coding schemes, category frequencies, and paraphrased statements are provided in S7 in Multimedia Appendix 1.

Table . Key findings from the focus groups with project-internal HCPs[a].

Code category	Summary
Game Design	HCPs[a] observed that older children were less engaged, mentioning limited variety and a perceived lack of competitiveness in the game.
Expectations	Initial expectations ranged from mixed curiosity to skepticism regarding the platform's ability to sustain engagement and seamlessly integrate into clinical workflows.
Consequences of App Usage	HCPs noted that some patients seemed to avoid the app as it reminded them of their illness, leading to disengagement, especially at home.
Communication	The app served as an additional communication channel, particularly for teenagers who preferred expressing themselves digitally. However, its impact on team communication seemed to highly depend on the strength of existing communication between HCPs, patients, and families. In cases of weaker communication, digital tools were seen as helpful for capturing and preserving important information.
Influences on Motivation	Initial curiosity waned due to limited feedback. HCPs reported more consistent use among girls and follow-up patients compared to boys or those in acute care.
Recommendations	HCPs recommended incorporating automated data collection (eg, wearables) and enhancing the HCP interface for better notification management. However, they also highlighted the need to address data protection and institutional barriers in general before effective implementation is possible.

[a]HCPs: health care professionals.

XSL•FO
RenderX

Table . Key findings from the focus groups with project-external HCPs[a].

Code category	Summary
Game Design	External HCPs[a] found the game's long-term engagement insufficient, citing competition with highly stimulating commercial games as a major hurdle.
Expectations	Initially, HCPs were optimistic about the app providing a constructive, meaningful alternative to passive media consumption for children.
Consequences of App Usage	HCPs described limited changes in daily workflows, although some noted occasional use for symptom tracking. Families sometimes used it to recall symptoms during appointments, but it did not seem to substantially alter HCP working routines.
Communication	The app aided communication for children more comfortable with digital tools, though HCPs emphasized that digital reporting should complement, not replace, in-person interactions.
Influences on Motivation	HCPs observed reduced engagement over time, particularly as patients' health improved. Parents of younger children were more engaged in symptom tracking, while older children primarily used the diary function for information sharing. Participants described engagement as influenced by usage habits, perceived utility, and occasional technical challenges (eg, login issues).
Impression of eHealth[b] Solutions	HCPs were generally optimistic about the growing role of eHealth but stressed that digital tools must complement traditional care. Challenges included administrative barriers and insufficient digital infrastructure.
Recommendations	External HCPs suggested additional personalized game features to increase long-term engagement and enhance the app interface for HCPs, including a priority system to highlight urgent issues.

[a]HCPs: health care professionals.

[b]eHealth: electronic health solution.

Discussion

Principal Findings

MyPal4Kids assessed the feasibility and acceptability of an ePRO-based platform in pediatric oncology and palliative care. Patient engagement was strongly influenced by perceived system utility, linked to HCP feedback frequency and the game's ability to sustain motivation. The core hypothesis that gamification would make self-reporting more engaging for young patients was partially supported: initial interest was high, but long-term engagement declined. These findings align with research emphasizing bidirectional communication in digital health [73-75], while also highlighting implementation challenges.

A key challenge was a disengagement cycle: fewer patient reports led to fewer platform reviews by HCPs, reducing feedback and further discouraging engagement. Similar patterns have been observed in other studies [76], likely amplified by the emotional burden on young patients with cancer and their families, along with competing HCP priorities. Usability did not limit patient engagement in our study but was more relevant for HCP adoption. These findings suggest that all groups may require additional support, such as proactive feedback loops and tighter integration into clinical workflows and systems, to maintain engagement. Engagement decreased over time, which represents an important finding of this study. This pattern was analyzed using descriptive longitudinal summaries rather than formal subgroup modeling. Engagement reflects a behavioral usage process characterized by heterogeneous exposure durations, irregular participation intervals, and varying initiation and discontinuation points across participants. This differs from structured repeated-measures clinical outcome assessments collected at predefined time points, which are more suitable for subgroup comparisons within the feasibility framework. To complement the quantitative description, we also further provide stratified descriptive observations (eg, age-related differences in usability and engagement patterns) alongside qualitative findings that help contextualize these trends, including differences in motivation, disease status, and HCP feedback.

Potential explanations for the observed decline include limited novelty effects, insufficient personalization of gamification elements, and varying needs across age groups. Additionally, the degree of feedback from HCPs may have influenced sustained use. These findings highlight the need for adaptive and user-centered design strategies to maintain long-term engagement. In addition to the COVID-19 pandemic, institutional factors such as varying levels of digital infrastructure, staffing resources, and staff engagement may have influenced implementation processes and platform use across study sites.

Similar engagement trajectories have been reported in pediatric digital health and ePRO-based interventions, where engagement is typically high at initiation but declines over time despite generally acceptable usability. Prior pediatric ePRO studies have similarly described decreases in reporting frequency in longitudinal implementations [77-79], as well as in broader mHealth and game-based interventions [80,81]. These findings

XSL•FO
RenderX

are consistent with evidence that gamification strategies often produce short-term engagement effects but are insufficient to sustain long-term use [30]. In addition, the effectiveness of social or competitive design elements may depend on a sufficiently large and continuously active user base, which is often lacking in pediatric feasibility settings [82]. Together, these observations suggest that sustained engagement requires more than initial usability and gamification and depends on continued perceived relevance, feedback mechanisms, and integration into clinical workflows.

Platform Utility for Clinical Care and Research

Importantly, this study was designed as a feasibility study and was not intended to evaluate the effectiveness of the intervention. Therefore, any observed changes in ePROs should be interpreted descriptively and not as evidence of causal effects. The platform proved feasible for tracking individual symptom burden over time, contingent on patient engagement in reporting symptoms. Clinicians could benefit from reviewing individual symptom trajectories. Statistical analyses showed significant main and interaction effects of time and group. Larger datasets could enable more nuanced comparisons, enhancing the platform's potential for research applications, such as comparing treatments in terms of effectiveness or side effects. This dual capability, supporting individualized patient care while generating population-level insights, underscores the platform's potential for both clinical practice and research. While symptom burden was the primary focus, secondary measures like children's QoL also showed improvement from baseline until the end of the study. Interpretation of these findings should remain cautious, as broader pandemic-related disruptions may have influenced reporting behaviors, engagement patterns, and ePROs during the study period [83-85].

Contextual Factors Influencing Implementation

Underlying factors may include age, digital literacy, emotional strain, and institutional support for digital health tools. Assessed symptoms largely reflected expected chemotherapy side effects, about which patients and parents had prior information. Consequently, HCPs sometimes gave limited direct feedback. Still, reports of symptoms such as tingling or numbness in the extremities are clinically relevant, potentially guiding treatment adjustments of medication doses. Our findings reinforce concerns that insufficiently integrated digital health tools can increase rather than reduce workload for HCPs, particularly when clear integration pathways into routine clinical workflows and responsibilities are lacking [75,86,87]. In the present study, the platform was deliberately not embedded into mandatory clinical workflows, which may have contributed to this effect. Existing commercial clinical systems offer limited interoperability with digital health platforms, including electronic health record integration, application programming interfaces and authentication infrastructures (eg, different user accounts and password systems). This complicates seamless integration into routine clinical workflows and requires the use of parallel systems. Dedicated staff to review digital health data could alleviate this burden, an approach supported in the literature [12]. Timely symptom data are crucial for responsive care, particularly between infrequent visits [88]. The observed decline in symptom reporting mirrors similar studies, possibly due to symptom relief or engagement fatigue. While gamification sparked initial interest, it alone was insufficient to sustain engagement, especially in older children. The game's design was constrained by requirements for nonviolence, gender neutrality, and broad age group appeal. While these specifications ensured accessibility, they may have resulted in an overly cautious design. Concerns about overuse did not materialize. Established gamification elements such as rewards and competitive features were less effective without a critical mass of simultaneous active users. Perspective differences between internal and external HCPs should be considered. Internal HCPs, involved in platform development, focused on long-term potential, while external HCPs emphasized practical challenges like workflow integration. Both groups agreed that technology should complement, not replace, direct care [89].

Strengths and Limitations

Interdisciplinary collaboration and testing across multiple clinical sites strengthened the study. Patient, parent, and HCP involvement in the design and evaluation enhanced the relevance of the findings. Ethical considerations were central to platform development and study implementation. Integration of ePROs into a gamified mobile app for young patients with cancer was innovative, aligning with evidence that children prefer visually engaging digital measures with short recall periods [90]. A strength of the presented platform is its flexibility for various ePROs and gamification strategies, enabling implementation across varied contexts.

The study period overlapped with the COVID-19 pandemic, which likely affected recruitment, clinical workflows, and sustained platform engagement. Reduced face-to-face interactions, restrictions during clinical visits, and increased reliance on digital communication may have influenced onboarding, participant retention, and reporting behavior. Prolonged exposure to digital communication and online platforms during the pandemic may also have contributed to digital fatigue and declining engagement over time. Similar pandemic-related disruptions affecting recruitment, ePRO-based research, and health-related QoL assessments have been described in pediatric and digital health research [83,85,91].

Further limitations include discrepancies between child and parent symptom reports, possibly due to recall issues or differing perceptions. While proxy reports can increase data completeness, they may introduce bias, highlighting the importance of careful interpretation and, where possible, cross-checking [92]. Some parents may have prioritized immediate care over proxy reports. Dual reporting was optional, limiting analyses to overlapping cases. Still, at certain time points, some symptoms demonstrated strong parent-child correlations, supporting limited proxy validity. MyPal was not designed as an emergency tool, as stated in the consent. Legal and ethical concerns, such as liability for missed reports, implicated that HCPs were not expected to monitor reports constantly. This limited the platform's immediate clinical utility but matched its intended use as a supplemental communication channel. The small sample size limits the robustness and generalizability of the findings. Given the limited sample size and reduced numbers in subgroup

XSL•FO
RenderX

analyses, these findings should be interpreted with caution. The results are exploratory and intended to inform future studies rather than provide definitive conclusions. The 6-month study period may not reflect long-term patterns. Given the drop in engagement, longer studies may require design refinements. The sample was limited in terms of diversity (eg, age and disease types). Selection bias is possible, as families with higher digital literacy or strong clinical engagement may have been more likely to participate. The overrepresentation of adolescents (aged 14 - 17 years) does not reflect the general pediatric cancer population [93]. Younger children younger than 4 years have the highest cancer incidence but were excluded due to literacy barriers. Treatment plan creation in the platform was optional for HCPs, limiting analysis of symptom-treatment correlations. Recruitment and HCP availability were likely affected by the pandemic, though the impacts are hard to quantify. Emotional burden during early illness phases also influenced recruitment. Subjective feedback from participants within the clinical setting may have shaped both recruitment and engagement. Measures such as flexible baseline and monthly assessment windows, as well as offline functionality, were implemented to support compliance, though their success was limited. Secondary end point measures (eg, patients' and parents' QoL, care satisfaction, and family impact) were not available to HCPs during the study. Access to these data could increase HCP engagement.

Taken together, these factors may limit the generalizability of the findings beyond European pediatric oncology settings, particularly in contexts with different health care infrastructures, cultural expectations, and levels of digital literacy.

Implications and Future Directions

Although more digital health solutions are emerging for specific illness phases, creating comprehensive platforms that address diverse individual needs remains challenging. Concerns persist about using research-oriented ePRO measures in clinical decision-making [8]. Pediatric-specific clinical ePRO guidelines are lacking [94], underscoring the need for evidence to support standardization and broader clinical adoption. Future versions should consider structured HCP training and targeted family support. Embedding ePROs into standard care rather than offering them as optional tools may help sustain use. Refinements should reflect user feedback and improve age-specific adaptability to enable scalability across clinical settings. Hybrid models, combining automated feedback with periodic HCP review, may boost engagement and long-term utility. Digital health solutions may not suit all patients; tailoring them to specific populations is key. Longitudinal randomized controlled trials are needed to assess effectiveness. Future research should explore the use of the platform in different patient populations, such as long-term survivors, who may demonstrate different supportive care needs and engagement patterns. In addition, studies conducted in postpandemic settings may help clarify the extent to which engagement patterns, recruitment dynamics, and sustained platform use were influenced by pandemic-related contextual factors. Evaluating the platform in larger cohorts and in more controlled study settings will be important next steps.

In conclusion, integrating digital health tools into pediatric oncology and palliative care requires both technological innovation and thoughtful workflow integration. Sustaining participation and utility demands consistent HCP engagement and compelling gamification. Large-scale implementation will require dedicated resources, including staff time for data monitoring, integration into existing hospital IT infrastructures and application programming interfaces, authentication systems, and sustainable allocation of clinical responsibility for digital symptom review. Our findings highlight key areas for refinement to improve ePRO-based approaches before clinical use is feasible.

Acknowledgments

MyPal was coordinated by Kostas Stamatopoulos of the Centre for Research and Technology Hellas, Greece. Other partners were Fraunhofer Institute for Biomedical Engineering, Germany; Foundation for Research and Technology Hellas, Greece; International Observatory on End of Life Care, Lancaster University, United Kingdom; Central European Institute of Technology, Masaryk University, Czech Republic; Karolinska Institute, Sweden; Vita-Salute San Raffaele University, Italy; University Hospital of Heraklion, Greece; Hannover Medical School, Germany; University Hospital Brno, Czech Republic; Saarland University, Germany; Promotion Software GMBH, Germany; Atlantis Healthcare, United Kingdom; European Association for Palliative Care; International Children's Palliative Care Network; National School of Public Health, Greece. The authors are grateful to the clinicians at the clinical study sites who were essential to this study through patient recruitment, symptom assessment and management, and data collection. This work is dedicated to Vassilis Koutkias.

Funding

MyPal: Fostering Palliative Care of Adults and Children with Cancer Through Advanced Patient-Reported Outcome Systems was funded by the European Union's Horizon 2020 Framework Programme under grant 825872. The funder had no role in the study design, data collection, data analysis, data interpretation, or writing of this paper.

Data Availability

The datasets generated and/or analyzed during this study are not publicly available due to the sensitive nature of pediatric patient data and restrictions imposed by ethical approval and data protection regulations. However, anonymized data may be available

XSL•FO
RenderX

from the corresponding author upon reasonable request and subject to institutional and ethical approval. Further information on the MyPal project, including newsletters and general project outcomes, is publicly accessible on the project website [43].

Authors' Contributions

MM created the initial draft of the paper. LR carried out the statistical analyses. LR, AB-G, MM, PN, PB, CK, SP, JD, JL, AS, KM, SD, JV, PL, OR, and NG contributed to critical revisions of the paper. NG, AS, and PL were local principal investigators and took overall responsibility for all aspects of the study design, the protocol and the study conduct at the involved clinical study sites.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional study materials, including questionnaires, focus group guides, detailed information on outcome measures, and extended quantitative and qualitative results.
[PDF File, 321 KB - cancer_v12i1e80718_app1.pdf]

References

1. Zebrack BJ, Butler M. Context for understanding psychosocial outcomes and behavior among adolescents and young adults with cancer. J Natl Compr Canc Netw 2012 Sep;10(9):1151-1156. [doi: 10.6004/jnccn.2012.0118]
2. Zebrack BJ, Corbett V, Embry L, et al. Psychological distress and unsatisfied need for psychosocial support in adolescent and young adult cancer patients during the first year following diagnosis. Psychooncology 2014 Nov;23(11):1267-1275. [doi: 10.1002/pon.3533] [Medline: 24664958]
3. Smrke A, Leung B, Srikanthan A, McDonald M, Bates A, Ho C. Distinct features of psychosocial distress of adolescents and young adults with cancer compared to adults at diagnosis: patient-reported domains of concern. J Adolesc Young Adult Oncol 2020 Aug;9(4):540-545. [doi: 10.1089/jayao.2019.0157] [Medline: 32255694]
4. Pyke-Grimm KA, Fisher B, Haskamp A, Bell CJ, Newman AR. Providing palliative and hospice care to children, adolescents and young adults with cancer. Semin Oncol Nurs 2021 Jun;37(3):151166. [doi: 10.1016/j.soncn.2021.151166] [Medline: 34175165]
5. Montgomery KE, Raybin JL, Ward J, et al. Using patient-reported outcomes to measure symptoms in children with advanced cancer. Cancer Nurs 2020;43(4):281-289. [doi: 10.1097/NCC.0000000000000721] [Medline: 32559050]
6. Reeve BB, Greenzang KA, Sung L. Patient-reported outcomes in pediatric patients with cancer. Am Soc Clin Oncol Educ Book 2023 May;43(43):e390272. [doi: 10.1200/EDBK_390272] [Medline: 37172266]
7. Wolfe J, Orellana L, Ullrich C, et al. Symptoms and distress in children with advanced cancer: prospective patient-reported outcomes from the PediQUEST study. J Clin Oncol 2015 Jun 10;33(17):1928-1935. [doi: 10.1200/JCO.2014.59.1222] [Medline: 25918277]
8. Churruca K, Pomare C, Ellis LA, et al. Patient-reported outcome measures (PROMs): a review of generic and condition-specific measures and a discussion of trends and issues. Health Expect 2021 Aug;24(4):1015-1024. [doi: 10.1111/hex.13254] [Medline: 33949755]
9. Rothmund M, Sodergren S, Rohde G, et al. Updating our understanding of health-related quality of life issues in children with cancer: a systematic review of patient-reported outcome measures and qualitative studies. Qual Life Res 2023 Apr;32(4):965-976. [doi: 10.1007/s11136-022-03259-z]
10. Bravo L, Killela MK, Reyes BL, et al. Self-management, self-efficacy, and health-related quality of life in children with chronic illness and medical complexity. J Pediatr Health Care 2020;34(4):304-314. [doi: 10.1016/j.pedhc.2019.11.009] [Medline: 32107073]
11. Yang LY, Manhas DS, Howard AF, Olson RA. Patient-reported outcome use in oncology: a systematic review of the impact on patient-clinician communication. Support Care Cancer 2018 Jan;26(1):41-60. [doi: 10.1007/s00520-017-3865-7] [Medline: 28849277]
12. Di Maio M, Basch E, Denis F, et al. The role of patient-reported outcome measures in the continuum of cancer clinical care: ESMO Clinical Practice Guideline. Ann Oncol 2022 Sep;33(9):878-892. [doi: 10.1016/j.annonc.2022.04.007] [Medline: 35462007]
13. Horan MR, Sim JA, Krull KR, Baker JN, Huang IC. A review of patient-reported outcome measures in childhood cancer. Children (Basel) 2022 Sep 30;9(10):1497. [doi: 10.3390/children9101497] [Medline: 36291433]
14. Holmen H, Winger A, Steindal SA, et al. Patient-reported outcome measures in children, adolescents, and young adults with palliative care needs-a scoping review. BMC Palliat Care 2023 Oct 6;22(1):148. [doi: 10.1186/s12904-023-01271-9] [Medline: 37798706]

XSL•FO
RenderX

15. Rusconi D, Basile I, Rampichini F, et al. Electronic patient reported outcomes measures (e-PROMs) in pediatric palliative oncology care: a scoping review. J Palliat Care 2024 Oct;39(4):298-315. [doi: 10.1177/08258597241274027] [Medline: 39295504]
16. Riedl D, Hubel N, Endres A, Vetrano C, Meryk A, Crazzolara R. Enhancing pediatric oncology clinical trials through patient reported outcomes (PROs). EJC Paediatr Oncol 2024 Dec;4:100199. [doi: 10.1016/j.ejcped.2024.100199]
17. Riedl D, Rothmund M, Darlington AS, et al. Rare use of patient-reported outcomes in childhood cancer clinical trials - a systematic review of clinical trial registries. Eur J Cancer 2021 Jul;152:90-99. [doi: 10.1016/j.ejca.2021.04.023] [Medline: 34090144]
18. Leahy AB, Feudtner C, Basch E. Symptom monitoring in pediatric oncology using patient-reported outcomes: why, how, and where next. Patient 2018 Apr;11(2):147-153. [doi: 10.1007/s40271-017-0279-z] [Medline: 29071524]
19. Greenzang KA, Montgomery KE, DuVall A, et al. Incorporation of patient-reported outcomes in pediatric cancer clinical trials: design, implementation, and dissemination. J Natl Cancer Inst 2025 Sep 1;117(9):1764-1773. [doi: 10.1093/jnci/djaf083] [Medline: 40163691]
20. Aanes SG, Wiig S, Nieder C, Haukland EC. Implementing digital patient-reported outcomes in routine cancer care: barriers and facilitators. ESMO Real World Data Digit Oncol 2024 Dec;6:100088. [doi: 10.1016/j.esmorw.2024.100088] [Medline: 41646091]
21. Schepers SA, Haverman L, Zadeh S, Grootenhuis MA, Wiener L. Healthcare professionals' preferences and perceived barriers for routine assessment of patient-reported outcomes in pediatric oncology practice: moving toward international processes of change. Pediatr Blood Cancer 2016 Dec;63(12):2181-2188. [doi: 10.1002/pbc.26135] [Medline: 27511830]
22. Rothmund M, Lehmann J, Moser W, et al. Patient-reported outcomes are under-utilised in evaluating supportive therapies in paediatric oncology - a systematic review of clinical trial registries. Crit Rev Oncol Hematol 2022 Aug;176:103755. [doi: 10.1016/j.critrevonc.2022.103755] [Medline: 35803454]
23. Zettl A, Gsell H, Maślak K, et al. How are you feeling today? – Gaps and opportunities in capturing patient reported outcomes (PROs) in pediatric cancer care. EJC Paediatric Oncology 2025 Dec;6:100306. [doi: 10.1016/j.ejcped.2025.100306]
24. Matza LS, Patrick DL, Riley AW, et al. Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. Value Health 2013 Jun;16(4):461-479. [doi: 10.1016/j.jval.2013.04.004] [Medline: 23796280]
25. Horan MR, Sim JA, Krull KR, et al. Ten considerations for integrating patient-reported outcomes into clinical care for childhood cancer survivors. Cancers (Basel) 2023 Feb 6;15(4):1024. [doi: 10.3390/cancers15041024] [Medline: 36831370]
26. Michael DR, Chen SL. Serious Games: Games That Educate, Train, and Inform: Muska & Lipman/Premier-Trade; 2005.
27. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". 2011 Presented at: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; Tampere, Finland p. 9-15. [doi: 10.1145/2181037.2181040]
28. Robson K, Plangger K, Kietzmann JH, McCarthy I, Pitt L. Is it all a game? Understanding the principles of gamification. Bus Horiz 2015 Jul;58(4):411-420. [doi: 10.1016/j.bushor.2015.03.006]
29. Helmefalk M. An interdisciplinary perspective on gamification: mechanics, psychological mediators and outcomes. IJSG 2019 Mar 14;6(1):3-26. [doi: 10.17083/ijsg.v6i1.262]
30. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Jul;71:31-48. [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]
31. Pramana G, Parmanto B, Lomas J, Lindhiem O, Kendall PC, Silk J. Using mobile health gamification to facilitate cognitive behavioral therapy skills practice in child anxiety treatment: open clinical trial. JMIR Serious Games 2018 May 10;6(2):e9. [doi: 10.2196/games.8902] [Medline: 29748165]
32. Martin A, Caon M, Adorni F, et al. A mobile phone intervention to improve obesity-related health behaviors of adolescents across Europe: iterative co-design and feasibility study. JMIR Mhealth Uhealth 2020 Mar 2;8(3):e14118. [doi: 10.2196/14118] [Medline: 32130179]
33. Miller AS, Cafazzo JA, Seto E. A game plan: gamification design principles in mHealth applications for chronic disease management. Health Informatics J 2016 Jun;22(2):184-193. [doi: 10.1177/1460458214537511] [Medline: 24986104]
34. Navarro-Alamán J, Lacuesta R, Garcia-Magariño I, Gallardo J. A methodology for the design and development of gamified mobile apps for monitoring cancer survivors. J Biomed Inform 2020 Jun;106:103439. [doi: 10.1016/j.jbi.2020.103439] [Medline: 32380179]
35. Stinson JN, Jibb LA, Nguyen C, et al. Construct validity and reliability of a real-time multidimensional smartphone app to assess pain in children and adolescents with cancer. Pain 2015 Dec;156(12):2607-2615. [doi: 10.1097/j.pain.0000000000000385] [Medline: 26580680]
36. Jibb LA, Stevens BJ, Nathan PC, et al. Implementation and preliminary effectiveness of a real-time pain management smartphone app for adolescents with cancer: a multicenter pilot clinical study. Pediatr Blood Cancer 2017 Oct;64(10):e26554. [doi: 10.1002/pbc.26554] [Medline: 28423223]
37. Fortier MA, Chung WW, Martinez A, Gago-Masague S, Sender L. Pain buddy: a novel use of m-health in the management of children's cancer pain. Comput Biol Med 2016 Sep 1;76:202-214. [doi: 10.1016/j.compbiomed.2016.07.012] [Medline: 27479493]

XSL•FO
RenderX

38. Linder LA, Newman AR, Stegenga K, et al. Feasibility and acceptability of a game-based symptom-reporting app for children with cancer: perspectives of children and parents. Support Care Cancer 2021 Jan;29(1):301-310. [doi: 10.1007/s00520-020-05495-w]
39. Tark R, Metelitsa M, Akkermann K, Saks K, Mikkel S, Haljas K. Usability, acceptability, feasibility, and effectiveness of a gamified mobile health intervention (Triumf) for pediatric patients: qualitative study. JMIR Serious Games 2019 Sep 30;7(3):e13776. [doi: 10.2196/13776] [Medline: 31573904]
40. Kim S, Wilson P, Abraham O. Investigating the use of serious games for cancer control among children and adolescents: scoping review. JMIR Serious Games 2024 Jul 10;12:e58724. [doi: 10.2196/58724] [Medline: 38985502]
41. Isabel Tan XW, Mordiffi SZ, Lopez V, Leong K. Psychological distress in parents of children with cancer: a descriptive correlational study. Asia Pac J Oncol Nurs 2021;8(1):94-102. [doi: 10.4103/apjon.apjon_46_20] [Medline: 33426196]
42. Van Schoors M, De Paepe AL, Lemiere J, et al. Family adjustment when facing pediatric cancer: the role of parental psychological flexibility, dyadic coping, and network support. Front Psychol 2019;10:2740. [doi: 10.3389/fpsyg.2019.02740] [Medline: 31920799]
43. Fostering palliative care of adults and children with cancer through advanced patient reported outcome systems. MyPal. URL: http://www.mypal-project.eu/ [accessed 2026-07-23]
44. Scarfò L, Karamanidou C, Doubek M, et al. MyPal ADULT study protocol: a randomised clinical trial of the MyPal ePRO-based early palliative care system in ADULT patients with haematological malignancies. BMJ Open 2021 Nov 2;11(11):e050256. [doi: 10.1136/bmjopen-2021-050256] [Medline: 34728446]
45. Meyerheim M, Karamanidou C, Payne S, et al. MyPal-Child study protocol: an observational prospective clinical feasibility study of the MyPal ePRO-based early palliative care digital system in paediatric oncology patients. BMJ Open 2021 Apr 13;11(4):e045226. [doi: 10.1136/bmjopen-2020-045226] [Medline: 33849855]
46. Orsmond GI, Cohn ES. The distinctive features of a feasibility study. OTJR: Occupational Therapy Journal of Research 2015 Jul;35(3):169-177. [doi: 10.1177/1539449215578649]
47. Karamanidou C, Natsiavas P, Koumakis L, et al. Electronic patient-reported outcome–based interventions for palliative cancer care: a systematic and mapping review. JCO Clin Cancer Inform 2020 Jul;4(4):647-656. [doi: 10.1200/CCI.20.00015] [Medline: 32697604]
48. Garani-Papadatos T, Natsiavas P, Meyerheim M, Hoffmann S, Karamanidou C, Payne SA. Ethical principles in digital palliative care for children: the MyPal Project and experiences made in designing a trustworthy approach. Front Digit Health 2022;4:730430. [doi: 10.3389/fdgth.2022.730430] [Medline: 35373180]
49. Meyerheim M, Burns-Gebhart A, Mirzaie K, Garani-Papadatos T, Braun Y, Graf N. Challenges and pitfalls for implementing digital health solutions in clinical studies in Europe. Front Digit Health 2021;3:730680. [doi: 10.3389/fdgth.2021.730680] [Medline: 34713200]
50. Hoffmann S, Schraut R, Kröll T, et al. AquaScouts: ePROs Implemented as a serious game for children with cancer to support palliative care. Front Digit Health 2021;3(December):730948. [doi: 10.3389/fdgth.2021.730948] [Medline: 34957461]
51. Mirzaie K, Burns-Gebhart A, Meyerheim M, Sander A, Graf N. Emotion regulation flexibility and electronic patient-reported outcomes: a framework for understanding symptoms and affect dynamics in pediatric psycho-oncology. Cancers (Basel) 2022 Aug 11;14(16):3874. [doi: 10.3390/cancers14163874] [Medline: 36010870]
52. Payne S, Begovic D, Salifu Y, et al. Applying digital health in cancer and palliative care in Europe: policy recommendations from an International Expert Workshop (MyPal Project). J Palliat Med 2024 Feb;27(2):216-223. [doi: 10.1089/jpm.2023.0309] [Medline: 37738323]
53. MyPal platform: fostering palliative care for cancer patients via epros. YouTube. URL: https://youtu.be/K32nGL2R7sk [accessed 2026-07-23]
54. Kyrou D, Bonotis P, Kakalou C, et al. Perceptions and needs of stakeholders regarding MyPal Project's electronic patient-reported outcome app: cross-sectional qualitative focus group study. JMIR Cancer 2025 Aug 13;11:e57388. [doi: 10.2196/57388] [Medline: 40802976]
55. Peres SC, Pham T, Phillips R. Validation of the System Usability Scale (SUS). Proc Hum Factors Ergon Soc Annu Meet 2013 Sep;57(1):192-196. [doi: 10.1177/1541931213571043]
56. Tomlinson D, Hyslop S, Stein E, et al. Development of mini-SSPedi for children 4-7 years of age receiving cancer treatments. BMC Cancer 2019 Jan 8;19(1):32. [doi: 10.1186/s12885-018-5210-z] [Medline: 30621617]
57. Dupuis LL, Johnston DL, Baggott C, et al. Validation of the symptom screening in pediatrics tool in children receiving cancer treatments. J Natl Cancer Inst 2018 Jun 1;110(6):661-668. [doi: 10.1093/jnci/djx250] [Medline: 29272441]
58. Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQLTM in pediatric cancer. Cancer 2002 Apr;94(7):2090-2106. [doi: 10.1002/cncr.10428]
59. Robert RS, Paxton RJ, Palla SL, et al. Feasibility, reliability, and validity of the pediatric quality of life inventoryTM generic core scales, cancer module, and multidimensional fatigue scale in long - term adult survivors of pediatric cancer. Pediatr Blood Cancer 2012 Oct;59(4):703-707. [doi: 10.1002/pbc.24099] [Medline: 22302778]
60. Brédart A, Anota A, Young T, et al. Phase III study of the European Organisation for Research and Treatment of Cancer satisfaction with cancer care core questionnaire (EORTC PATSAT-C33) and specific complementary outpatient module (EORTC OUT-PATSAT7). Eur J Cancer Care 2018 Jan;27(1):e12786. [doi: 10.1111/ecc.12786]

XSL•FO
RenderX

61. Brédart A, Beaudeau A, Young T, et al. The European organization for research and treatment of cancer — satisfaction with cancer care questionnaire: revision and extended application development. Psychooncology 2017 Mar;26(3):400-404. [doi: 10.1002/pon.4127]
62. Stein REK, Riessman CK. The development of an impact-on-family scale: preliminary findings. Med Care 1980 Apr;18(4):465-472. [doi: 10.1097/00005650-198004000-00010] [Medline: 7401703]
63. Stein REK, Jessop DJ. The impact on family scale revisited: further psychometric data. J Dev Behav Pediatr 2003 Feb;24(1):9-16. [doi: 10.1097/00004703-200302000-00004] [Medline: 12584480]
64. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
65. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
66. R Core Team. R: A language and environment for statistical computing. : R Foundation for Statistical Computing; 2022 URL: https://www.R-project.org/ [accessed 2025-10-22]
67. Kay M, Elkin LA, Higgins JJ, Wobbrock JO. ARTool: aligned rank transform for nonparametric factorial ANOVAs. 2021. [doi: 10.5281/zenodo.594511]
68. Wobbrock JO, Findlater L, Gergle D, Higgins JJ. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. 2011 May 7 Presented at: CHI '11; Vancouver BC Canada p. 143-146 URL: https://dl.acm.org/doi/proceedings/10.1145/1978942 [doi: 10.1145/1978942.1978963]
69. Karnofsky DA, Abelmann WH, Craver LF, Burchenal JH. The use of the nitrogen mustards in the palliative treatment of carcinoma. With particular reference to bronchogenic carcinoma. Cancer 1948 Nov;1(4):634-656. [doi: 10.1002/1097-0142(194811)1:43.0.CO;2-L]
70. Péus D, Newcomb N, Hofer S. Appraisal of the Karnofsky Performance Status and proposal of a simple algorithmic system for its evaluation. BMC Med Inform Decis Mak 2013 Jul 19;13(1):72. [doi: 10.1186/1472-6947-13-72] [Medline: 23870327]
71. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13(1):117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
72. Udaka YT, Packer RJ. Pediatric brain tumors. Neurol Clin 2018 Aug;36(3):533-556. [doi: 10.1016/j.ncl.2018.04.009] [Medline: 30072070]
73. Alpert JM, Dyer KE, Lafata JE. Patient-centered communication in digital medical encounters. Patient Educ Couns 2017 Oct;100(10):1852-1858. [doi: 10.1016/j.pec.2017.04.019] [Medline: 28522229]
74. Jensen LWH, Ghaffari A, Rahbek O, Dinesen B, Kold S. The use of asynchronous digital two-way communication between patients and healthcare professionals after hospital discharge: A scoping review. Patient Educ Couns 2024 Nov;128:108393. [doi: 10.1016/j.pec.2024.108393] [Medline: 39146890]
75. Mair JL, Hashim J, Thai L, et al. Understanding and overcoming barriers to digital health adoption: a patient and public involvement study. Transl Behav Med 2025 Jan 16;15(1):ibaf010. [doi: 10.1093/tbm/ibaf010] [Medline: 40167046]
76. Amagai S, Pila S, Kaat AJ, Nowinski CJ, Gershon RC. Challenges in participant engagement and retention using mobile health apps: literature review. J Med Internet Res 2022 Apr 26;24(4):e35120. [doi: 10.2196/35120] [Medline: 35471414]
77. Dreher J, Meryk A, Giesinger JM, et al. Adherence to daily electronic symptom screening in pediatric cancer: associations with time, location and pain. Qual Life Res 2025 Aug;34(8):2363-2375. [doi: 10.1007/s11136-025-04002-0] [Medline: 40493152]
78. Meryk A, Kropshofer G, Hetzer B, et al. Use of daily patient-reported outcome measurements in pediatric cancer care. JAMA Netw Open 2022 Jul 1;5(7):e2223701. [doi: 10.1001/jamanetworkopen.2022.23701] [Medline: 35881395]
79. Rodgers C, Wills-Bagnato P, Sloane R, Hockenberry M. Health-related quality of life among children and adolescents during hematopoietic stem cell transplant recovery. J Pediatr Oncol Nurs 2015;32(5):329-336. [doi: 10.1177/1043454214563413] [Medline: 25592668]
80. Mustafa AS, Ali N, Dhillon JS, Alkawsi G, Baashar Y. User engagement and abandonment of mHealth: a cross-sectional survey. Healthcare (Basel) 2022 Jan 24;10(2):221. [doi: 10.3390/healthcare10020221] [Medline: 35206837]
81. Wang M, Xu J, Zhou X, Li X, Zheng Y. Effectiveness of gamification interventions to improve physical activity and sedentary behavior in children and adolescents: systematic review and meta-analysis. JMIR Serious Games 2025 Sep 18;13:e68151. [doi: 10.2196/68151] [Medline: 40966596]
82. Patel MS, Small DS, Harrison JD, et al. Effectiveness of behaviorally designed gamification interventions with social incentives for increasing physical activity among overweight and obese adults across the United States: the STEP UP randomized clinical trial. JAMA Intern Med 2019 Dec 1;179(12):1624-1632. [doi: 10.1001/jamainternmed.2019.3505] [Medline: 31498375]
83. Holch P, Turner G, Keetharuth AD, Gibbons E, Cocks K, Absolom KL. The impact of COVID-19 on PRO development, collection and implementation: views of UK and Ireland professionals. J Patient Rep Outcomes 2023 Nov 27;7(1):121. [doi: 10.1186/s41687-023-00663-y] [Medline: 38010558]

XSL•FO
RenderX

84. Nobari H, Fashi M, Eskandari A, Villafaina S, Murillo-Garcia Á, Pérez-Gómez J. Effect of COVID-19 on health-related quality of life in adolescents and children: a systematic review. Int J Environ Res Public Health 2021 Apr 25;18(9):4563. [doi: 10.3390/ijerph18094563] [Medline: 33923120]
85. Tran A, Hou SHJ, Forbes C, et al. The impact of the early COVID-19 global pandemic on children undergoing active cancer treatment and their parents. Curr Oncol 2023 Feb 17;30(2):2441-2456. [doi: 10.3390/curroncol30020186] [Medline: 36826147]
86. Wosny M, Strasser LM, Hastings J. Experience of health care professionals using digital tools in the hospital: qualitative systematic review. JMIR Hum Factors 2023 Oct 17;10:e50357. [doi: 10.2196/50357] [Medline: 37847535]
87. Keuper J, van Tuyl LHD, de Geit E, et al. The impact of eHealth use on general practice workload in the pre-COVID-19 era: a systematic review. BMC Health Serv Res 2024 Sep 19;24(1):1099. [doi: 10.1186/s12913-024-11524-9] [Medline: 39300456]
88. Coolbrandt A, Van den Heede K, Vanhove E, De Bom A, Milisen K, Wildiers H. Immediate versus delayed self-reporting of symptoms and side effects during chemotherapy: does timing matter? Eur J Oncol Nurs 2011 Apr;15(2):130-136. [doi: 10.1016/j.ejon.2010.06.010] [Medline: 20685164]
89. Coombes L, Harðardóttir D, Braybrook D, et al. Design and administration of patient-centred outcome measures: the perspectives of children and young people with life-limiting or life-threatening conditions and their family members. Patient 2023 Sep;16(5):473-483. [doi: 10.1007/s40271-023-00627-w] [Medline: 37221441]
90. Coombes L, Bristowe K, Ellis-Smith C, et al. Enhancing validity, reliability and participation in self-reported health outcome measurement for children and young people: a systematic review of recall period, response scale format, and administration modality. Qual Life Res 2021 Jul;30(7):1803-1832. [doi: 10.1007/s11136-021-02814-4] [Medline: 33738710]
91. Tran NN, Tran M, Lopez J, Woon J, Nguyen J, Brecht ML. Impact of COVID-19 on pediatric clinical research. J Pediatr Nurs 2021;61:240-244. [doi: 10.1016/j.pedn.2021.06.005] [Medline: 34171517]
92. Montgomery KE, Vos K, Raybin JL, et al. Comparison of child self-report and parent proxy-report of symptoms: results from a longitudinal symptom assessment study of children with advanced cancer. J Spec Pediatr Nurs 2021 Jul;26(3):e12316. [doi: 10.1111/jspn.12316] [Medline: 33118275]
93. Steliarova-Foucher E, Colombet M, Ries LAG, et al. International incidence of childhood cancer, 2001-10: a population-based registry study. Lancet Oncol 2017 Jun;18(6):719-731. [doi: 10.1016/S1470-2045(17)30186-9] [Medline: 28410997]
94. Cracchiolo JR, Arafat W, Atreja A, et al. Getting ready for real‐world use of electronic patient‐reported outcomes (ePROs) for patients with cancer: a National Comprehensive Cancer Network ePRO Workgroup paper. Cancer 2023 Aug 15;129(16):2441-2449. [doi: 10.1002/cncr.34844] [Medline: 37224181]

Abbreviations

EORTC PATSAT-C33: European Organisation for Research and Treatment of Cancer Satisfaction with Cancer Care Core Questionnaire
ePRO: electronic patient-reported outcome
FNBRNO: University Hospital Brno
GDPR: General Data Protection Regulation
HCP: health care professional
IOFS: Impact on Family Scale
IOFS: Impact on Family Scale
KPS: Karnofsky Performance Scale
MHH: Hannover Medical School
PedsQL: Pediatric Quality of Life Inventory
PedsQL: Pediatric Quality of Life Inventory
QoL: quality of life
SUS: System Usability Scale
USAAR: Saarland University

XSL•FO
RenderX

Edited by M Balcarras; submitted 15.Jul.2025; peer-reviewed by B Satravada, SL Coven; revised version received 29.May.2026; accepted 29.May.2026; published 11.Aug.2026.

Please cite as:
Meyerheim M, Radová L, Burns-Gebhart A, Mergen M, Natsiavas P, Bonotis P, Karamanidou C, Payne S, Downing J, Ling J, Sander A, Mirzaie K, Dokutur S, Vaculíková J, Lokaj P, Rohleder O, Graf N
Feasibility of Implementing an Electronic Patient-Reported Outcome-Based Digital Health Platform in Pediatric Oncology and Palliative Care: Insights From the MyPal4Kids Multicenter Observational Study
JMIR Cancer 2026;12:e80718
URL: https://cancer.jmir.org/2026/1/e80718
doi:10.2196/80718

© Marcel Meyerheim, Lenka Radová, Anna Burns-Gebhart, Marvin Mergen, Pantelis Natsiavas, Panos Bonotis, Christina Karamanidou, Sheila Payne, Julia Downing, Julie Ling, Annette Sander, Kasra Mirzaie, Susan Dokutur, Jana Vaculíková, Petr Lokaj, Ondřej Rohleder, Norbert Graf. Originally published in JMIR Cancer (https://cancer.jmir.org), 11.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Large Language Models for Supporting Clear Writing and Detecting Spin in Randomized Controlled Trials in Oncology: Comparative Analysis of GPT Models and Prompts

Carole Koechli[1,2], MD; Fabio Dennstädt[2], MD; Christina Schröder[1,2], MD; Daniel M Aebersold[2], MD; Robert Förster[1,2], MD; Daniel R Zwahlen[1], MD; Paul Windisch[1,2], MD

[1]Department of Radiation Oncology, Kantonsspital Winterthur, Brauerstrasse 15, Winterthur, Switzerland

[2]Department of Radiation Oncology, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

Corresponding Author:
Paul Windisch, MD
Department of Radiation Oncology, Kantonsspital Winterthur, Brauerstrasse 15, Winterthur, Switzerland

Abstract

Background: Randomized controlled trials (RCTs) are the gold standard for evaluating interventions in oncology, but reporting can be subject to "spin"—presenting results in ways that mislead readers about true efficacy.

Objective: This study aimed to investigate whether large language models (LLMs) could provide a standardized approach to detect spin, particularly in the conclusions, where it most commonly occurs.

Methods: We randomly sampled 250 two-arm, single–primary end point oncology RCTs from 7 major medical journals published between 2005 and 2023. Two authors independently annotated trials as positive or negative based on whether they met their primary end point. Three commercial LLMs (GPT-3.5 Turbo, GPT-4o, and GPT-o1) were tasked with classifying trials as positive or negative when provided with (1) conclusions only; (2) methods and conclusions; (3) methods, results, and conclusions; or (4) title and full abstract. LLM performance was evaluated against human annotations. Afterward, trials incorrectly classified as positive when the model was provided only with the conclusions but correctly classified as negative when provided with the whole abstract were analyzed for patterns that may indicate the presence of spin. Model performance was assessed using accuracy, precision, recall, and F_1-score calculated from confusion matrices.

Results: Of the 250 trials, 146 (58.4%) were positive, and 104 (41.6%) were negative. The GPT-o1 model demonstrated the highest performance across all conditions, with F_1-scores of 0.932 (conclusions only; 95% CI 0.90-0.96), 0.96 (methods and conclusions; 95% CI 0.93-0.98), 0.98 (methods, results, and conclusions; 95% CI 0.96-0.99), and 0.97 (title and abstract; 95% CI 0.95-0.99). Analysis of trials incorrectly classified as positive when the model was provided only with the conclusions revealed shared patterns, including absence of primary end point results, emphasis on subgroup improvements, or unclear distinction between primary and secondary end points. These patterns were almost never found in trials correctly classified as negative.

Conclusions: LLMs can effectively detect potential spin in oncology RCT reporting by identifying discrepancies between how trials are presented in the conclusions vs the full abstracts. This approach could serve as a supplementary tool for improving transparency in scientific reporting, although further development is needed to address more complex trial designs beyond those examined in this feasibility study.

(JMIR Cancer 2026;12:e78221) doi:10.2196/78221

KEYWORDS

spin; randomized controlled trials; large language models; data mining; natural language processing

Introduction

Randomized controlled trials (RCTs) represent the gold standard for evaluating interventions in oncology [1]. However, the reporting and interpretation of trial results can be subject to inconsistency and "spin"—the presentation of results in a way that may mislead readers about the true efficacy of interventions [2]. This can, for example, be accomplished by emphasizing secondary end points or subgroup analyses when primary end points are not met. While most research that has looked at the topic has found a substantial prevalence of spin, the exact number varies as it is not always straightforward to differentiate between what constitutes a balanced and comprehensive presentation of the results and what may be an attempt to mislead the reader [3].

The presence of spin has important implications. Clinicians, policymakers, and even patients often rely heavily on abstracts and conclusions when interpreting trial findings, as full-text

XSL•FO
RenderX

analyses are time-consuming and not always accessible. Therefore, misrepresentation of results might contribute to overly optimistic perceptions of treatment benefits, potentially influencing clinical decision-making, guideline development, and even the allocation of research funding. Given the increasing complexity of cancer care and the rapidly expanding volume of clinical trials, ensuring clarity and accuracy in scientific reporting is crucial to avoid bias in evidence synthesis and translation into practice.

The growing capabilities of large language models (LLMs) could constitute a standardized way to determine the presence of spin. If an abstract is clearly written, a state-of-the-art LLM should be able to determine whether its primary end point was met. As multiple studies have identified the conclusions as the most frequent source of spin [4,5], we hypothesized that trials which are correctly classified as negative—defined as trials that did not meet their primary end point—by an LLM when provided with the title and abstract, but incorrectly classified as positive when provided with only the conclusions, would be likely to contain some form of spin. Therefore, the aim of this study was to evaluate whether LLMs can reliably classify oncology RCTs as positive or negative and whether discrepancies between conclusion-only and full-abstract classifications can help identify patterns consistent with spin.

Methods

Overview

Randomized controlled oncology trials from 7 major medical journals (*British Medical Journal*, *Journal of the American Medical Association*, *Journal of the American Medical Association Oncology*, *Journal of Clinical Oncology*, *The Lancet*, *The Lancet Oncology*, and *The New England Journal of Medicine*) published between 2005 and 2023 were randomly sampled by downloading the available abstracts for the time frame via PubMed in a text file and parsing the abstracts using regular expressions. These 7 journals were selected because they publish a large and consistent volume of oncology RCTs and are widely regarded as leading general or oncology-specific medical journals. The 2005 to 2023 range was chosen to capture contemporary trial reporting practices while ensuring sufficient volume across all selected journals. To avoid edge cases for this feasibility study, it was decided to limit the eligible trials to designs with exactly 2 arms and 1 primary end point.

We aimed to sample 250 trials as this number ensured a sufficiently large dataset for the feasibility analysis while remaining feasible for manual dual annotation. Trials were sampled by creating a randomized list of all retrieved abstracts. Two authors (CK and PW) then started the annotation from the top of the random list and stopped after 250 two-arm, single–primary end point oncology trials had been annotated. No journal-level quotas were applied.

The purpose of the annotation was to establish the ground-truth classification—whether the trial met its primary end point—against which model predictions could be evaluated. The annotation was conducted in a 2-step process. After annotating the first 20 trials, all samples were discussed to recognize potential differences in the annotation criteria. The remaining trials were annotated separately, and discrepancies were discussed after all trials had been annotated. A third author (DRZ) would have been responsible for judging disagreements that persisted after discussion. However, this was not necessary. The annotation was performed using the Prodigy tool (version 1.13.1; Explosion), which only showed the extracted abstract as text without any additional information such as authors or institutions. Only in cases in which the abstract did not clearly state the primary end point and its results did we refer to the full publication or protocol. Three commercially available LLMs, namely, GPT-3.5 Turbo, GPT-4o, and GPT-o1 (OpenAI), were then tasked with classifying the trials as positive or negative. The 3 models were chosen to investigate whether the inherent capabilities of the models would impact their suitability for the classification task (eg, simpler models requiring more explicit language to correctly identify trials) and, thus, their performance when trying to leverage differences in classification accuracy to detect unclear writing and spin. The decision to use OpenAI models was based on the prevalent use of these models at the time as well as the convenience of application programming interface access and lack of privacy concerns regarding the study data. The respective model snapshots were gpt-3.5-turbo-0125, gpt-4o-2024-11-20, and o1-2024-12-17. The LLMs were called via the application programming interface, with the temperature parameter set to 1. We refrained from performing multiple classification runs as a previous study from the same research group had shown very consistent performance by LLMs for both classification and named-entity recognition tasks, as long as the temperature was kept at or below 1.50 [6].

Each model was evaluated in 4 different rounds. In round 1, the models were only provided with the conclusions of the abstract. In round 2, the models were provided with the methods and conclusions of the abstract. In round 3, the models were provided with the methods, results, and conclusions of the abstract. In round 4, the models were provided with the title and the full abstract.

The following system prompt (ie, the fixed instruction provided to the model to define its task) was used: “You will be provided with the {section} of a randomized controlled oncology trial. Your task will be to classify if the trial was positive, i.e. if it met its primary endpoint, or negative, i.e. if it did not meet its primary endpoint. Your response should be either the word POSITIVE (in all caps) or NEGATIVE (in all caps).”

The “{section}” part was replaced with either “conclusion,” “methods and conclusion,” “methods, results, and conclusion,” or “title and abstract.” The user prompt (ie, the specific input text) was the corresponding title, abstract, or sections of the abstract.

The prompts were designed to be as explicit as possible regarding the definition of a positive trial to minimize ambiguity and ensure consistent model behavior across conditions. However, we did not conduct a systematic comparison of different prompts.

XSL•FO
RenderX

Statistical Analysis

Interannotator agreement was calculated as the percentage of agreement divided by the total number of annotated trials.

The results were evaluated against the ground truth (ie, the human-annotated classification of whether the trial met its primary end point) by creating confusion matrices and computing several performance metrics to obtain a holistic picture of model performance. These included accuracy (the proportion of correctly classified trials among all trials), precision (the proportion of predicted positive trials that were truly positive; equivalent to positive predictive value), recall (the proportion of truly positive trials that were correctly predicted as positive; equivalent to sensitivity), and F_1-score (the harmonic mean of precision and recall). For completeness, specificity (true negative rate), and negative predictive value can also be derived from the confusion matrix but were not separately reported. The 95% CIs were estimated using normal approximation intervals. For the best-performing model, we further analyzed and categorized the trials that were incorrectly predicted as positive when provided with the conclusions but were correctly predicted as negative when provided with the title and abstract. For these trials, a single author (PW) reviewed the full conclusions and abstracts to categorize the patterns leading to incorrect classification (eg, omission of primary end point, emphasis on subgroup findings, or unclear distinction between end points). To contextualize these findings, we additionally selected 10 randomly chosen trials correctly classified as negative by GPT-o1 and performed the same qualitative assessment. All programming was performed in Python (Python Software Foundation; version 3.13.2) using, among others, the *pandas* (version 2.2.3) and *openai* (version 1.67.0) packages.

Ethical Considerations

This study used publicly available abstracts from published clinical trials. All data were deidentified and contained no patient-level information; therefore, ethics approval was not required.

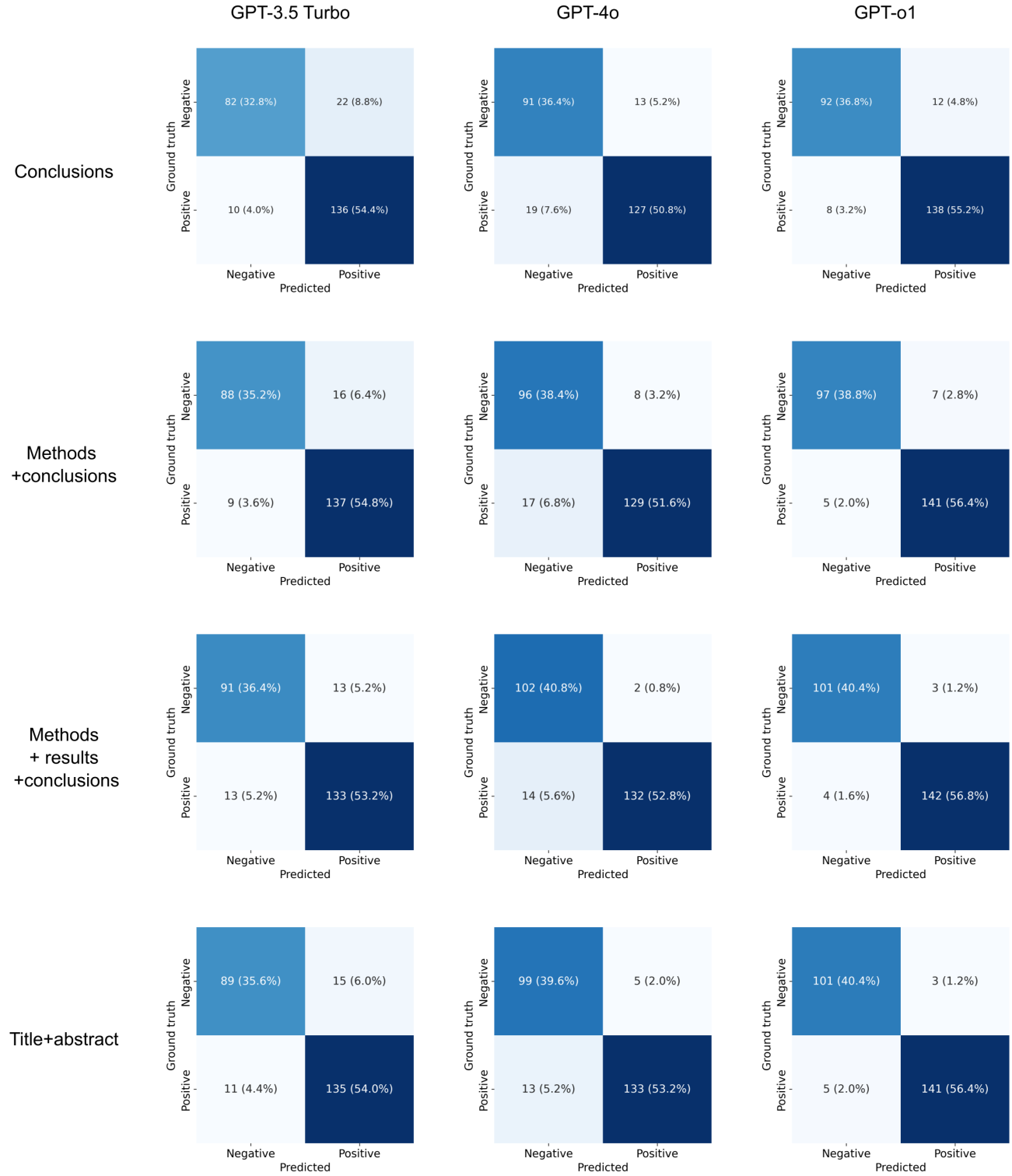
Results

Interannotator agreement was 97.2% (243/250). All of the disagreements were caused by simple mistakes and could be easily resolved during the discussion. Ultimately, 58.4% (146/250) of the trials were annotated as positive, and 41.6% (104/250) were annotated as negative.

The performances of the models when provided with different sections of the abstract are shown in Figure 1 and Table 1. GPT-o1 exhibited the best performance in each round, with F_1-scores of 0.932 (conclusions only), 0.96 (methods and conclusions), 0.98 (methods, results, and conclusions), and 0.97 (title and abstract). GPT-4o's F_1-scores across the 4 rounds were 0.89, 0.91, 0.94, and 0.94, respectively. GPT-3.5 Turbo exhibited F_1-scores of 0.89, 0.92, 0.91, and 0.91, respectively.

XSL•FO
RenderX

Figure 1. Confusion matrices. Classification performance of GPT-3.5 Turbo, GPT-4o, and GPT-o1 when predicting whether a trial was positive or negative based on different sections of the abstract.



XSL•FO
RenderX

Table . Classification performance. Accuracy, precision, recall, and F_1-score for GPT-3.5 Turbo, GPT-4o, and GPT-o1 when predicting whether a trial was positive based on different sections of the abstract.

	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
Conclusions only				
GPT-3.5 Turbo	0.87 (0.83 - 0.91)	0.86 (0.82 - 0.90)	0.93 (0.90 - 0.96)	0.89 (0.86 - 0.93)
GPT-4o	0.87 (0.83 - 0.91)	0.91 (0.87 - 0.94)	0.87 (0.83 - 0.91)	0.89 (0.85 - 0.93)
GPT-o1	0.92 (0.89 - 0.95)	0.92 (0.89 - 0.95)	0.95 (0.92 - 0.97)	0.93 (0.90 - 0.96)
Methods+conclusions				
GPT-3.5 Turbo	0.90 (0.86 - 0.94)	0.90 (0.86 - 0.93)	0.94 (0.91 - 0.97)	0.92 (0.88 - 0.95)
GPT-4o	0.90 (0.86 - 0.94)	0.94 (0.91 - 0.97)	0.88 (0.84 - 0.92)	0.91 (0.88 - 0.95)
GPT-o1	0.95 (0.93 - 0.98)	0.95 (0.93 - 0.98)	0.97 (0.94 - 0.99)	0.96 (0.93 - 0.98)
Methods+results+conclusions				
GPT-3.5 Turbo	0.90 (0.86 - 0.93)	0.91 (0.88 - 0.95)	0.91 (0.88 - 0.95)	0.91 (0.88 - 0.95)
GPT-4o	0.94 (0.91 - 0.97)	0.99 (0.97 - 1.00)	0.90 (0.87 - 0.94)	0.94 (0.91 - 0.97)
GPT-o1	0.97 (0.95 - 0.99)	0.98 (0.96 - 1.00)	0.97 (0.95 - 0.99)	0.98 (0.96 - 0.99)
Title+abstract				
GPT-3.5 Turbo	0.90 (0.86 - 0.93)	0.90 (0.86 - 0.94)	0.92 (0.89 - 0.96)	0.91 (0.88 - 0.95)
GPT-4o	0.93 (0.90 - 0.96)	0.96 (0.94 - 0.99)	0.91 (0.88 - 0.95)	0.94 (0.91 - 0.97)
GPT-o1	0.97 (0.95 - 0.99)	0.98 (0.96 - 1.00)	0.97 (0.94 - 0.99)	0.97 (0.95 - 0.99)

We further analyzed trials that were incorrectly predicted as positive by GPT-o1 when the model was only provided with the conclusions but predicted correctly when provided with the title and abstract. Of these 10 trials, 6 (60%) did not mention the primary end point in the conclusions [7-12]. One mentioned an improvement in the primary end point in a subgroup [13]. One mentioned both improved secondary end points and the unimproved primary end point without specifying which was which [14]. The remaining 2 trials mentioned that one arm was superior to the other one without specifying that it was the control arm that showed improved results [15,16]. The list of trials is provided in Table 2.

XSL•FO
RenderX

Table . Trials that were incorrectly predicted as positive by GPT-o1 when the model was only provided with the conclusions but predicted correctly when provided with the title and abstract.

Title	Conclusions reported on primary end point	Possible reason for incorrect prediction
"Total Body Irradiation or Chemotherapy Conditioning in Childhood ALL: A Multinational, Randomized, Noninferiority Phase III Study" [15]	Yes	Conclusions mentioned that TBI[a] plus etoposide showed improved overall survival. Therefore, the model likely thought that TBI plus etoposide was the intervention that was tested, whereas it was actually the control.
"Volasertib Versus Chemotherapy in Platinum-Resistant or -Refractory Ovarian Cancer: A Randomized Phase II Groupe des Investigateurs Nationaux pour l'Etude des Cancers de l'Ovaire Study" [7]	No	Primary end point was not discussed in the conclusions.
"High-Dose Therapy and Autologous Blood Stem-Cell Transplantation Compared With Conventional Treatment in Myeloma Patients Aged 55 to 65 Years: Long-Term Results of a Randomized Control Trial From the Group Myelome-Autogreffe" [14]	Yes	Conclusions mentioned both improved secondary end points and the unimproved primary end point without specifying which was which.
"Results of a Randomized Trial of Chlorambucil Versus Fludarabine for Patients With Untreated Waldenström Macroglobulinemia, Marginal Zone Lymphoma, or Lymphoplasmacytic Lymphoma" [8]	No	Primary end point was not mentioned in the conclusions.
"Bortezomib-Dexamethasone, Rituximab, and Cyclophosphamide as First-Line Treatment for Waldenström's Macroglobulinemia: A Prospectively Randomized Trial of the European Consortium for Waldenström's Macroglobulinemia" [9]	No	Primary end point was not mentioned in the conclusions.
"Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial" [12]	No	Primary end point was not mentioned in the conclusions.
"Addition of Bevacizumab to Bolus Fluorouracil and Leucovorin in First-Line Metastatic Colorectal Cancer: Results of a Randomized Phase II Trial" [10]	No	Primary end point was not mentioned in the conclusions.
"Oral ibandronic acid versus intravenous zoledronic acid in treatment of bone metastases from breast cancer: a randomised, open label, non-inferiority phase 3 trial" [16]	Yes	Conclusions mentioned the superiority of zoledronic acid. Therefore, the model likely thought that zoledronic acid was the intervention, whereas it was the comparator in this noninferiority study.
"Bcl-2 Antisense (oblimersen sodium) Plus Dacarbazine in Patients With Advanced Melanoma: The Oblimersen Melanoma Study Group" [13]	Yes	Improvement in the primary end point in a subgroup was mentioned.
"Efficacy and Safety of Trabectedin or Dacarbazine for Metastatic Liposarcoma or Leiomyosarcoma After Failure of Conventional Chemotherapy: Results of a Phase III Randomized Multicenter Clinical Trial" [11]	No	Primary end point was not mentioned in the conclusions.

[a]TBI: total body irradiation.

To confirm that those writing patterns were not equally frequent in trials correctly classified as negative, we also analyzed 10 random trials predicted correctly as negative by GPT-o1 and have provided the analysis in Multimedia Appendix 1. Of these trials, only 10% (1/10) did not mention the primary end point for the whole trial population in its conclusions but, instead, reported the results of the primary end point in a positive subgroup [17]. In total, 70% (7/10) of the trials explicitly mentioned that the primary end point failed to meet statistical significance or that the trial as a whole was negative or only mentioned the negative primary end point in their conclusions [18-24]. A total of 20% (2/10) of the trials mentioned both the primary end point and secondary end points or subgroups [25,26].

XSL•FO
RenderX

Discussion

Principal Findings

In this study, we evaluated the ability of 3 commercial LLMs to classify oncology RCTs as positive or negative based on different sections of trial abstracts. Our findings demonstrate that modern LLMs, particularly more advanced models, can achieve high classification accuracy even when provided with limited information. Our findings also support the hypothesis that trials that are correctly classified as negative by an LLM when provided with the title and abstract but incorrectly classified as positive when provided with only the conclusions are likely to contain patterns that may be interpreted as spin. While there is no ground truth of what constitutes spin, not mentioning the results for the primary end point at all in the conclusions, mentioning an improvement in the primary end point that only occurred in a subgroup, or mixed reporting of primary and secondary end points without clear distinction would be at least considered questionable by many readers [27]. Our findings also highlight that the LLM-based approach is not perfectly specific. In total, 20% (2/10) of the studies for which o1 was misled to believe they were positive when provided only with the conclusions had conclusions that clearly mentioned which arm had better outcomes. However, the LLM did not know which arm was the intervention and which arm was the control, so it assumed that the superior arm was the intervention arm. While this way of phrasing a conclusion may not be optimal for readability, it is certainly not an attempt at misleading the reader, who will still know which treatment yielded better results.

Therefore, our approach is likely not suitable as a fully automated solution. However, it demonstrated its potential to inform editors, reviewers, and authors alike of potential spin or unclear writing. The question of “Are the results for the primary endpoint clearly recognizable in the conclusion?” might serve as an alternative litmus test. Even though reviewers and journal editors are generally capable of recognizing questionable conclusions, we do believe that automated tools have value considering the ever-increasing list of items that have to be considered when conducting a careful review as they may, if implemented carefully, point toward parts of the manuscript that need increased attention. Another group of people who might benefit from a higher degree of automation are physicians who do not routinely read RCTs or have to do it in a situation in which they do not have time to fully digest all aspects of the research, such as in between patient consultations.

Comparison to Prior Work

While research on LLMs and spin is still in its infancy, Yun et al [28] evaluated 22 LLMs and found that they are actually more susceptible to spin than humans. As LLMs are being used increasingly for screening and synthesizing scientific literature, this highlights the importance of improved detection of spin, preferably at the prepublication stage. However, the approach demonstrated in this study could also be leveraged as part of a screening pipeline to detect spin when trying to systematically analyze the literature in an automated fashion.

Strengths and Limitations

This study has several strengths. The human annotation process was systematic, with independent dual review and consensus resolution, resulting in a reliable ground-truth dataset. Evaluating 3 LLMs of differing capability provided insights into how model complexity affects performance and sensitivity to unclear reporting. In addition, the structured comparison across 4 abstract conditions enabled us to isolate how specific sections of reporting contribute to misclassification.

This study has several limitations. First, the analysis was restricted to RCTs with 2 arms and a single primary end point. This constraint reduced complexity and helped ensure consistent interpretation but limits the applicability of our findings to trials with more complex designs, such as those involving multiple or co–primary end points. As noted in this paper, such designs introduce additional analytic considerations, for example, prespecified alpha splitting, that would have increased methodological heterogeneity and potentially confounded the evaluation [29]. Therefore, the restriction was deliberate, but it reduced generalizability.

Second, we did not include trials using analytical frameworks other than standard hypothesis testing, such as Bayesian designs [30]. Because these studies report results differently and may emphasize posterior probabilities rather than traditional statistical significance, the performance of LLMs in such contexts remains unknown. This limitation reflects the scope of the feasibility study rather than an inherent barrier of the method.

Third, it is uncertain whether the models had previously encountered some of the included abstracts during training. If so, prior exposure could have artificially increased performance, particularly when models were presented with only part of an abstract. Although this possibility cannot be fully eliminated for proprietary language models, our key analyses focused on discrepancies between conclusion-only and full-abstract predictions. These discrepancies are less susceptible to prior knowledge because recognizing internal inconsistencies requires examining the relationship between sections rather than retrieving memorized text. Nonetheless, this limitation may have influenced overall performance metrics.

Fourth, this study used a single, clearly defined prompt that specified what should be considered a positive or negative trial. While this approach ensured consistent instructions across models and conditions, it remains possible that different prompting strategies would yield different results. The choice of a single explicit prompt was intended to minimize variability, but it may limit insight into how models behave under alternative or less directive task formulations.

Future Directions

Future work could extend this approach to more complex trial designs, including studies with multiple or co–primary end points, adaptive designs, or Bayesian frameworks, to determine whether LLM-based assessments remain reliable under conditions in which end point interpretation is less straightforward. Evaluating models from different vendors and open-source architectures may also help clarify how

XSL•FO
RenderX

generalizable these findings are beyond the commercial systems examined in this study. In addition, refining prompting strategies or incorporating structured domain knowledge could improve model understanding of trial context, particularly in situations in which the distinction between intervention and control is not explicitly stated. Prospective integration of LLM-based screening tools into editorial workflows may help assess their practical utility in real-time manuscript evaluation. Finally, future studies may investigate whether LLMs can assist in promoting clearer reporting practices by providing automated feedback to authors during manuscript preparation.

Conclusions

In conclusion, this study demonstrates that LLMs can highlight potential spin in oncology trial reporting by identifying inconsistencies between conclusions and full abstracts. These findings suggest a possible role for LLMs as supportive tools that draw attention to areas in which reporting may be unclear or incomplete. While not a substitute for expert review, such tools may help promote clearer communication of trial results. Further evaluation in more complex trial settings will be needed to determine how broadly this approach can be applied.

Funding

No funding was received for this project.

Data Availability

The datasets generated or analyzed during this study are available in the GitHub repository [31].

Authors' Contributions

Conceptualization: CK, PW
Data curation: CK, PW
Formal analysis: CK, PW
Methodology: CK, PW
Project administration: DMA, DRZ
Supervision: DRZ
Writing – original draft: PW
Writing – review & editing: CK, FD, CS, DMA, RF, DRZ
All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Analysis of 10 random trials predicted correctly as negative by GPT-o1.
[PDF File, 64 KB - cancer_v12i1e78221_app1.pdf]

References

1. Booth CM, Cescon DW, Wang L, Tannock IF, Krzyzanowska MK. Evolution of the randomized controlled trial in oncology over three decades. J Clin Oncol 2008 Nov 20;26(33):5458-5464. [doi: 10.1200/JCO.2008.16.5456] [Medline: 18955452]
2. Mahtani KR. “Spin” in reports of clinical research. Evid Based Med 2016 Dec;21(6):201-202. [doi: 10.1136/ebmed-2016-110570] [Medline: 27737894]
3. Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. J Clin Oncol 2014 Dec 20;32(36):4120-4126. [doi: 10.1200/JCO.2014.56.7503] [Medline: 25403215]
4. Ito C, Hashimoto A, Uemura K, Oba K. Misleading reporting (spin) in noninferiority randomized clinical trials in oncology with statistically not significant results: a systematic review. JAMA Netw Open 2021 Dec 1;4(12):e2135765. [doi: 10.1001/jamanetworkopen.2021.35765] [Medline: 34874407]
5. Wayant C, Margalski D, Vaughn K, Vassar M. Evaluation of spin in oncology clinical trials. Crit Rev Oncol Hematol 2019 Dec;144:102821. [doi: 10.1016/j.critrevonc.2019.102821] [Medline: 31733444]
6. Windisch P, Dennstädt F, Koechli C, et al. The impact of temperature on extracting information from clinical trial publications using large language models. Cureus 2024 Dec 15;16(12):e75748. [doi: 10.7759/cureus.75748] [Medline: 39811231]
7. Pujade-Lauraine E, Selle F, Weber B, et al. Volasertib versus chemotherapy in platinum-resistant or -refractory ovarian cancer: a randomized phase II Groupe des Investigateurs Nationaux pour l’Etude des Cancers de l’Ovaire Study. J Clin Oncol 2016 Mar 1;34(7):706-713. [doi: 10.1200/JCO.2015.62.1474] [Medline: 26755507]

XSL•FO
RenderX

8. Leblond V, Johnson S, Chevret S, et al. Results of a randomized trial of chlorambucil versus fludarabine for patients with untreated Waldenström macroglobulinemia, marginal zone lymphoma, or lymphoplasmacytic lymphoma. J Clin Oncol 2013 Jan 20;31(3):301-307. [doi: 10.1200/JCO.2012.44.7920] [Medline: 23233721]
9. Buske C, Dimopoulos MA, Grunenberg A, et al. Bortezomib-dexamethasone, rituximab, and cyclophosphamide as first-line treatment for Waldenström's macroglobulinemia: a prospectively randomized trial of the European Consortium for Waldenström's macroglobulinemia. J Clin Oncol 2023 May 10;41(14):2607-2616. [doi: 10.1200/JCO.22.01805] [Medline: 36763945]
10. Kabbinavar FF, Schulz J, McCleod M, et al. Addition of bevacizumab to bolus fluorouracil and leucovorin in first-line metastatic colorectal cancer: results of a randomized phase II trial. J Clin Oncol 2005 Jun 1;23(16):3697-3705. [doi: 10.1200/JCO.2005.05.112] [Medline: 15738537]
11. Demetri GD, von Mehren M, Jones RL, et al. Efficacy and safety of trabectedin or dacarbazine for metastatic liposarcoma or leiomyosarcoma after failure of conventional chemotherapy: results of a phase III randomized multicenter clinical trial. J Clin Oncol 2016 Mar 10;34(8):786-793. [doi: 10.1200/JCO.2015.62.4734] [Medline: 26371143]
12. van de Velde CJ, Rea D, Seynaeve C, et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. Lancet 2011 Jan 22;377(9762):321-331. [doi: 10.1016/S0140-6736(10)62312-4] [Medline: 21247627]
13. Bedikian AY, Millward M, Pehamberger H, et al. Bcl-2 antisense (oblimersen sodium) plus dacarbazine in patients with advanced melanoma: the Oblimersen Melanoma Study Group. J Clin Oncol 2006 Oct 10;24(29):4738-4745. [doi: 10.1200/JCO.2006.06.0483] [Medline: 16966688]
14. Fermand JP, Katsahian S, Divine M, et al. High-dose therapy and autologous blood stem-cell transplantation compared with conventional treatment in myeloma patients aged 55 to 65 years: long-term results of a randomized control trial from the Group Myelome-Autogreffe. J Clin Oncol 2005 Dec 20;23(36):9227-9233. [doi: 10.1200/JCO.2005.03.0551] [Medline: 16275936]
15. Peters C, Dalle JH, Locatelli F, et al. Total body irradiation or chemotherapy conditioning in childhood ALL: a multinational, randomized, noninferiority phase III study. J Clin Oncol 2021 Feb 1;39(4):295-307. [doi: 10.1200/JCO.20.02529] [Medline: 33332189]
16. Barrett-Lee P, Casbard A, Abraham J, et al. Oral ibandronic acid versus intravenous zoledronic acid in treatment of bone metastases from breast cancer: a randomised, open label, non-inferiority phase 3 trial. Lancet Oncol 2014 Jan;15(1):114-122. [doi: 10.1016/S1470-2045(13)70539-4] [Medline: 24332514]
17. Hurwitz HI, Uppal N, Wagner SA, et al. Randomized, double-blind, phase II study of ruxolitinib or placebo in combination with capecitabine in patients with metastatic pancreatic cancer for whom therapy with gemcitabine has failed. J Clin Oncol 2015 Dec 1;33(34):4039-4047. [doi: 10.1200/JCO.2015.61.4578] [Medline: 26351344]
18. Gregorc V, Gaafar RM, Favaretto A, et al. NGR-hTNF in combination with best investigator choice in previously treated malignant pleural mesothelioma (NGR015): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet Oncol 2018 Jun;19(6):799-811. [doi: 10.1016/S1470-2045(18)30193-1] [Medline: 29753703]
19. Small EJ, Schellhammer PF, Higano CS, et al. Placebo-controlled phase III trial of immunologic therapy with sipuleucel-T (APC8015) in patients with metastatic, asymptomatic hormone refractory prostate cancer. J Clin Oncol 2006 Jul 1;24(19):3089-3094. [doi: 10.1200/JCO.2005.04.5252] [Medline: 16809734]
20. Cobleigh MA, Anderson SJ, Siziopikou KP, et al. Comparison of radiation with or without concurrent trastuzumab for HER2-positive ductal carcinoma in situ resected by lumpectomy: a phase III clinical trial. J Clin Oncol 2021 Jul 20;39(21):2367-2374. [doi: 10.1200/JCO.20.02824] [Medline: 33739848]
21. Yap YS, Kwok LL, Syn N, et al. Predictors of hand-foot syndrome and pyridoxine for prevention of capecitabine-induced hand-foot syndrome: a randomized clinical trial. JAMA Oncol 2017 Nov 1;3(11):1538-1545. [doi: 10.1001/jamaoncol.2017.1269] [Medline: 28715540]
22. Pisansky TM, Pugh SL, Greenberg RE, et al. Tadalafil for prevention of erectile dysfunction after radiotherapy for prostate cancer: the Radiation Therapy Oncology Group [0831] randomized clinical trial. JAMA 2014 Apr 2;311(13):1300-1307. [doi: 10.1001/jama.2014.2626] [Medline: 24691606]
23. Zhong LP, Zhang CP, Ren GX, et al. Randomized phase III trial of induction chemotherapy with docetaxel, cisplatin, and fluorouracil followed by surgery versus up-front surgery in locally advanced resectable oral squamous cell carcinoma. J Clin Oncol 2013 Feb 20;31(6):744-751. [doi: 10.1200/JCO.2012.43.8820] [Medline: 23129742]
24. Ray-Coquard I, Harter P, Lorusso D, et al. Effect of weekly paclitaxel with or without bevacizumab on progression-free rate among patients with relapsed ovarian sex cord-stromal tumors: the ALIENOR/ENGOT-ov7 randomized clinical trial. JAMA Oncol 2020 Dec 1;6(12):1923-1930. [doi: 10.1001/jamaoncol.2020.4574] [Medline: 33030515]
25. Butts C, Socinski MA, Mitchell PL, et al. Tecemotide (L-BLP25) versus placebo after chemoradiotherapy for stage III non-small-cell lung cancer (START): a randomised, double-blind, phase 3 trial. Lancet Oncol 2014 Jan;15(1):59-68. [doi: 10.1016/S1470-2045(13)70510-2] [Medline: 24331154]
26. Miller KD, Chap LI, Holmes FA, et al. Randomized phase III trial of capecitabine compared with bevacizumab plus capecitabine in patients with previously treated metastatic breast cancer. J Clin Oncol 2005 Feb 1;23(4):792-799. [doi: 10.1200/JCO.2005.05.098] [Medline: 15681523]

XSL•FO
RenderX

27. Reynolds-Vaughn V, Riddle J, Brown J, Schiesel M, Wayant C, Vassar M. Evaluation of spin in the abstracts of emergency medicine randomized controlled trials. Ann Emerg Med 2019 May 14:423-431. [doi: 10.1016/j.annemergmed.2019.03.011] [Medline: 31101371]
28. Yun HS, Zhang KY, Kouzy R, Marshall IJ, Li JJ, Wallace BC. Caught in the web of words: do LLMs fall for spin in medical literature? Proc Mach Learn Res 2025 Jun;287:458-479. [Medline: 41257216]
29. Multiple endpoints in clinical trials: guidance for industry. US Food and Drug Administration. 2022. URL: https://www.fda.gov/files/drugs/published/Multiple-Endpoints-in-Clinical-Trials-Guidance-for-Industry.pdf [accessed 2026-01-07]
30. van de Schoot R, Depaoli S, King R, et al. Bayesian statistics and modelling. Nat Rev Methods Primers 2021 Jan 14;1:1. [doi: 10.1038/s43586-020-00001-2]
31. Windisch-paul/positive_negative. GitHub. URL: https://github.com/windisch-paul/positive_negative [accessed 2025-05-25]

Abbreviations

LLM: large language model
RCT: randomized controlled trial

Edited by N Cahill; submitted 28.May.2025; peer-reviewed by D Hu, J Williams-Nguyen, S Matsuda; revised version received 04.Dec.2025; accepted 16.Dec.2025; published 21.Jan.2026.

Please cite as:
Koechli C, Dennstädt F, Schröder C, Aebersold DM, Förster R, Zwahlen DR, Windisch P
Large Language Models for Supporting Clear Writing and Detecting Spin in Randomized Controlled Trials in Oncology: Comparative Analysis of GPT Models and Prompts
JMIR Cancer 2026;12:e78221
URL: https://cancer.jmir.org/2026/1/e78221
doi:10.2196/78221

© Carole Koechli, Fabio Dennstädt, Christina Schröder, Daniel M Aebersold, Robert Förster, Daniel R Zwahlen, Paul Windisch. Originally published in JMIR Cancer (https://cancer.jmir.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Physical Activity Monitoring in Adolescents and Young Adults With Cancer: Observational Feasibility Study

Savannah V Rauschendorfer[1], MSc, PhD; Beatrice Cantoni[2], MSc; Michelle Audirac[3], MSc; Susan C Gilchrist[4], MD, MS; J Andrew Livingston[5], MD; Michael E Roth[6], MD; Corwin M Zigler[7], PhD; Eugenie S Kleinerman[6], MD

[1]Health, Human Performance, and Recreation, Baylor University, 1311 S 5th Street, Waco, TX, United States

[2]Department of Statistics and Data Sciences, The University of Texas at Austin, Austin, TX, United States

[3]Department of Biostatistics, Harvard University, Boston, MA, United States

[4]Department of Medicine, Division of Cardiology, University of North Carolina School of Medicine, Chapel Hill, NC, United States

[5]Department of Sarcoma Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[6]Department of Pediatrics, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[7]Department of Biostatistics, Brown University, Providence, RI, United States

Corresponding Author:
Savannah V Rauschendorfer, MSc, PhD
Health, Human Performance, and Recreation, Baylor University, 1311 S 5th Street, Waco, TX, United States

Abstract

Background: There is a growing interest in characterizing the relationship between long-term physical activity (PA) habits and cancer-related outcomes such as treatment-related toxicities, recurrence, and complications. Wearable devices can provide critical information to achieve this goal; however, inferences are significantly influenced by device wear adherence.

Objective: This study aimed to assess the feasibility of using wearable devices to monitor short- and long-term PA in adolescent and young adult (AYA) patients with sarcoma during and after chemotherapy in a free-living environment and evaluate the ability to accurately capture changes in PA over 3 years.

Methods: A total of 63 AYA patients with sarcoma were provided with a Fitbit Charge 3 to track daily steps, sedentary time, and heart rate for up to 3 years.

Results: On average, during the first 30 days of follow-up, 57.1% (36/63) of patients wore their device at least 10 hours per day, and only 23.8% (15/63) of patients wore their devices thereafter. Patients spent a mean of 80% (SD 11%) of their day in a sedentary state. Despite low adherence, daily step count trends increased over time.

Conclusions: This study highlights the adherence challenges met with longitudinal PA monitoring in AYA patients with sarcoma. Wearer discomfort, lost devices, and lack of data uploading compliance contributed to data missingness and attrition. Caution is warranted when relying on wearable activity trackers to inform program decisions, accurately assess PA outcomes, and monitor program adherence longitudinally without consideration of wearer bias. Alternative methods that would be more broadly accepted by AYA patients for effective long-term monitoring should be considered.

(JMIR Cancer 2026;12:e87591) doi:10.2196/87591

KEYWORDS

sarcoma; adolescents and young adults; physical activity tracking; wearable device; chemotherapy

Introduction

Adolescent and young adult (AYA) patients who receive high doses of anthracycline treatment such as doxorubicin are at high risk of long-term cardiac morbidity and mortality [1]. Physical activity (PA) has been previously demonstrated to protect against and mitigate doxorubicin-induced cardiotoxicity [2-4]. As a result, exercise is often proposed as a nonpharmacological early prevention intervention to combat cardiovascular side effects of cancer therapy. There is a growing interest in characterizing the relationship between long-term PA habits and cancer-related outcomes such as treatment-related toxicities, recurrence, and complications. Furthermore, there is interest in understanding how PA interventions may modify these crucial outcomes. Wearable devices are widely used for quantification of PA [5]. Among the demographic of AYA patients with sarcoma who receive high-dose anthracyclines, little is known regarding the feasibility of using wearable devices to monitor PA longitudinally for associations with late cancer-related outcomes.

In addition to wearable devices being capable of providing long-term activity data [6-8], it has been well established that PA monitoring can inform exercise prescription decisions and

XSL•FO
RenderX

quantify exercise intervention participation [9-11]. Information gathered from PA monitoring devices aids in designing tailored interventions for specific patient populations of interest. Exercise intervention programming and resultant prevention success hinge on the current behaviors, interest, and activity levels of a given cohort of individuals [12]. On the basis of the target population, exercise prescription in terms of frequency, intensity, time, and type can be highly variable. Among AYA patients with cancer, there is a lack of consensus on frequency, intensity, time, and type parameters to be used in patients receiving therapy [13]. Information gathered from regular activity levels may indicate optimal timing as well as intensity for a practical PA intervention. Therefore, recording PA in AYA patients undergoing chemotherapy is essential for the design of a feasible exercise regimen to improve adherence probability in a randomized trial.

The lack of participant willingness to wear and interact with PA tracking devices longitudinally can present a significant challenge for creating exercise interventions, monitoring program adherence, conducting unbiased data interpretation, and deducing critical scientific conclusions for determining the impact of exercise interventions on cancer-related toxicities. Researchers have called for investigations that evaluate the long-term adherence to and sustainability of activity-monitoring interventions using wearables [14-16]. Therefore, the purpose of this study was to evaluate the feasibility of assessing short- and long-term PA during and after chemotherapy using a wearable device in AYA patients with sarcoma in a free-living environment and determine whether such data can accurately describe the evolution of PA levels of AYA patients over the span of 3 years.

Methods

Patient Recruitment

PA was assessed as a component of a primary cardiotoxicity investigation. Patients between the ages of 15 and 39 years at the time of study entry were recruited at MD Anderson Cancer Center. The study follow-up period for the primary cardiotoxicity end points was 2 years after recruitment. Cardiotoxicity end points and sample size calculations were based on evidence that demonstrated significant cardiac remodeling 2 years after diagnosis [17]. However, patients were informed that the Fitbit study monitoring period would last up to 3 years from enrollment, wherein data would be collected passively without intervention to maximize the information gained from the wearable device and gaining new information from it.

Inclusion criteria were patients with a histologically confirmed diagnosis of sarcoma (including either soft tissue or bone sarcoma) and anticipated total doxorubicin dose of 300 mg/m^2 or higher or other anthracycline with doxorubicin equivalence of 300 mg/m^2 or higher as determined by the treating physician. Moreover, individuals included in the study must indicate that they were willing to wear and upload data from the PA wearable device throughout the study period. Patients were excluded if there was any condition that compromised the ability of the patient or their legally recognized representative to give written informed consent and/or comply with the study. Patients with a history of cardiac morbidity were also excluded from study participation.

Ethical Considerations

Informed consent was obtained from all participants included in the study. The study was approved by the MD Anderson Cancer Center institutional review board (PA18-0462). All study procedures were conducted in accordance with the ethical standards of the Declaration of Helsinki. The privacy and confidentiality of the research participants' data and identity were maintained. Patients were allowed to keep the Fitbit Charge 3 (Google) device after study conclusion or discontinuation. No other patient incentive was provided for participation.

PA Tracking

PA was assessed using an electronic wearable activity-monitoring device (Fitbit Charge 3 provided by the Department of Pediatrics of MD Anderson Cancer Center). Fitbit devices are widely used and well known for their commercial popularity, high wear compliance, and strong correlation with research-grade monitoring devices [15,18]. The Fitbit Charge 3 was selected based on the resources available for device purchase and implementation [19], as well as for its demonstrated significant correlation with research-grade PA monitoring devices such as the ActiGraph wGT3X-BT [18]. After consent, patients were given a Fitbit device and went through an educational introduction to the device whereby a study coordinator explained how to wear it and charge it and aided in downloading the Fitbit app to the participants' mobile devices or tablets for device data synchronization. Patients were asked to wear the device for a minimum of 6 months and were instructed to wear it at all times unless they were charging it. Participants were additionally informed that the study monitoring period would last a maximum of 3 years. No PA-specific instructions were given in order to to assess changes in PA without an intervention during and after chemotherapy. The Fitbit device was used to track aerobic PA, which included minute-level measures of steps; heart rate (HR); and a classification of activity as sedentary, light, moderate, or active. The device was linked to Fitabase (Small Steps Labs LLC), a database that stores activity data on a secure online website. Data storage on the Fitabase platform was participant dependent via app synchronization. Therefore, data were reviewed routinely throughout the study period. Study coordinators attempted to contact participants who did not upload data for 10 of the previous 14 days for participant retention. There were no device-integrated automatic or proactive reminders to participants to wear or synchronize their devices.

Device Wear Time Parameters

Whether the device was being worn in at a given minute was determined based on (1) whether the patient's device logged any data during the given day and (2) applying the algorithm developed by Choi et al [20] to consider the device worn during minutes for which there was an HR measurement or during minutes of spurious missing HR. PA was summarized between the hours of 8 AM and 8 PM, with daily summaries calculated

XSL•FO
RenderX

only for "valid days," defined as days during which the device was worn for at least 10 hours between 8 AM and 8 PM per best practice consensus for waking day sampling [19]. The 8 AM to 8 PM time window has been previously used and defined as a standard waking or daytime period used to monitor PA [21-23].

Statistical Analysis

Study device wear patterns were summarized by calculating the proportion of participants enrolled on a given follow-up day who wore the Fitbit device for at least 10 hours between 8 AM and 8 PM (a "valid day"). Daily PA summaries for each patient were calculated as the total number of steps, average daily HR, and proportion of time classified as sedentary activity among valid days of device wear. Means are reported with the SD, and medians are reported with the IQR. To evaluate changes in PA over time, a mixed-effects linear regression model with normally distributed random intercepts and slopes was fit to model trends in daily step count. This model specified patient-specific time trends, various patient characteristics (sex, tumor location, BMI, and age), and treatment status (during or after chemotherapy treatment) and adjusted for day of the week variability. The follow-up day was calculated as the day from primary cardiotoxicity study enrollment.

Results

Overview

A total of 75 patients were eligible, recruited, consented for participation, and enrolled in the study. In total, 84% (63/75) of these patients (33/63, 52.4% male and 30/63, 47.6% female) had valid Fitbit data that were included in the analysis (Table 1). The reasons for 16% (12/75) of the individuals not logging Fitbit data were variable; however, most related to inability or unwillingness to use the Fitbit device and/or its supporting app or study withdrawal before setting up the Fitbit device. The mean age and BMI of the cohort were 24 (SD 7.4) years and 27.7 (SD 10.0) kg/m^2, respectively. The most common types of sarcoma diagnosis were osteosarcoma (18/63, 28.6%) and synovial sarcoma (13/63, 20.6%), and the most common locations of the primary tumor were the trunk (33/62, 53.2%) and lower limb (21/62, 33.9%; Table 1). Patients were enrolled in the primary study for a median of 730 (IQR 622-730) days. The median number of days that a patient logged Fitbit data was 205 (IQR 58-352).

Over the study duration, 11.1% (7/63) of the wearable devices were lost, and 3.2% (2/63) were broken. All were replaced by the study team. Among the 63 patients, there were 5 (7.9%) reports of wearer discomfort due to skin sensitivity wherein patients subsequently stopped wearing the activity tracker. In several instances, wearers forgot to synchronize the device for data transfer, making consistent data collection a recurrent challenge for the study team.

XSL•FO
RenderX

Table . Patient characteristics (n=63).

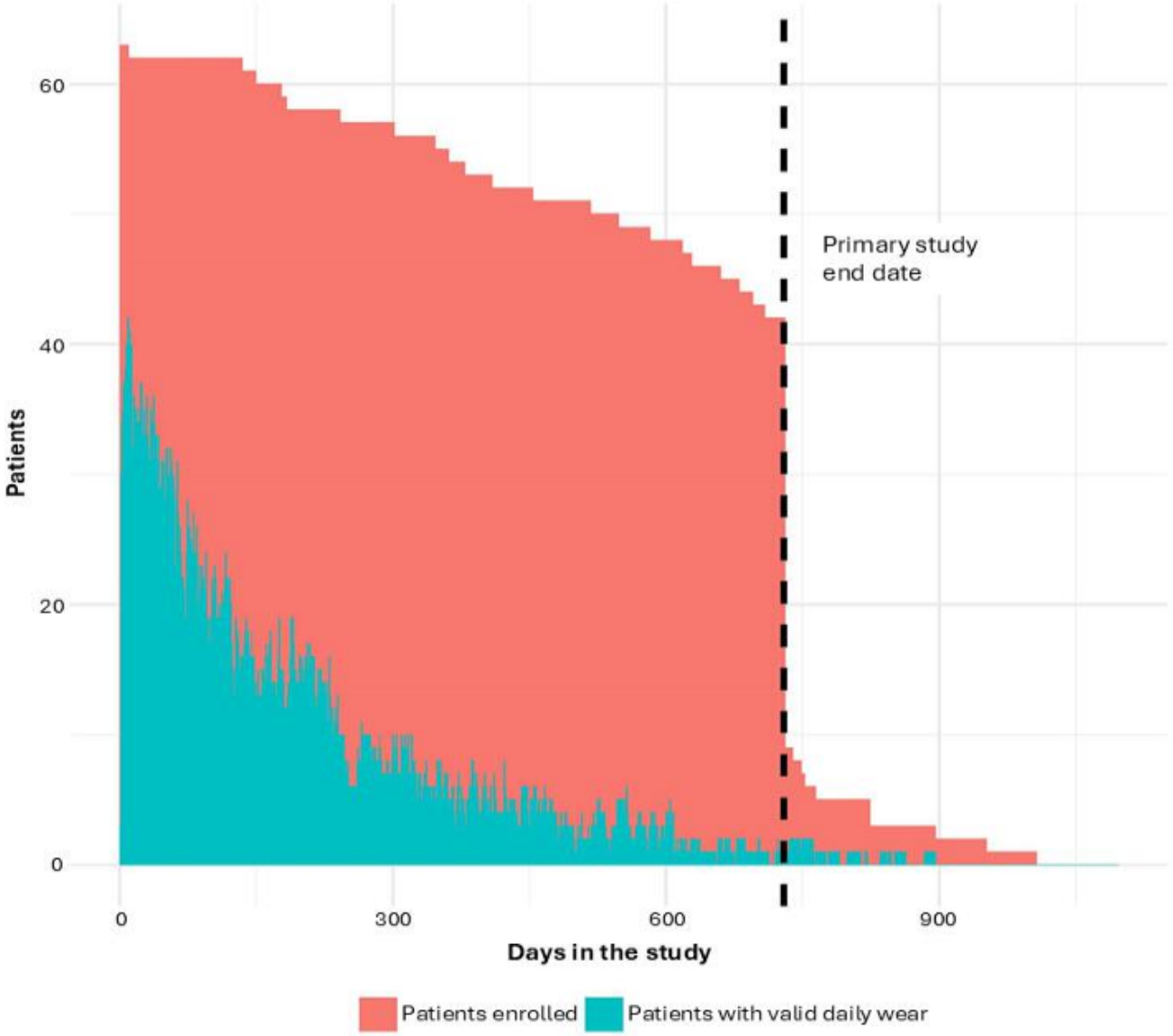
	Values
Duration of study enrollment (days), mean (SD)	630 (208)
Duration of study enrollment (days), median (IQR; range)	730 (622-730; 3-1006)
Duration of Fitbit participation (days), mean (SD)	245 (226)
Duration of Fitbit participation (days), median (IQR; range)	205 (58-352; 2-950)
Age (years), mean (SD)	24 (7.4)
Age (years), median (IQR; range)	23 (18-31; 15-39)
BMI (kg/m^2), mean (SD)	27.7 (10.0)
BMI (kg/m^2), median (IQR; range)	24.3 (21.0-31.0; 15.4-65.7)
Total doxorubicin dose (mg/m^2), mean (SD)	354.27 (125.84)
Total doxorubicin dose (mg/m^2), median (IQR; range)	379.80 (299.25-446.50; 25.00-633.37)
Radiation, n (%)	
Yes	26 (41.3)
No	37 (58.7)
Surgery, n (%)	
Yes	42 (66.7)
No	21 (33.3)
Sex, n (%)	
Female	30 (47.6)
Male	33 (52.4)
Sarcoma diagnosis, n (%)	
Angiosarcoma	2 (3.2)
Ewing sarcoma	9 (14.3)
Liposarcoma	5 (7.9)
Osteosarcoma	18 (28.6)
Rhabdomyosarcoma	4 (6.3)
Soft tissue sarcoma	3 (4.8)
Synovial sarcoma	13 (20.6)
Other	9 (14.3)
Tumor location (n=62), n (%)	
Trunk	33 (53.2)
Lower limb	21 (33.9)
Upper limb	3 (4.8)
Other	5 (8.1)

Device Wear Summary

The highest proportion of device wear occurred on day 9 of follow-up, when 66.7% (42/63) of study participants exhibited valid wear of at least 10 hours between 8 AM and 8 PM. All other study days saw fewer patients with valid device wear, which decreased as the study progressed. Allowing for a 7-day lag to set up the Fitbit app, the average proportion of patients exhibiting valid daily wear was 57% during follow-up days 8 to 30, with the proportion of valid wear averaging 24% during days 31 to 365 and 6% from day 366 to the primary cardiotoxicity study end point of 2 years. Device wear during the study period is summarized in Figure 1. Moreover, patients demonstrated intervariability for daily patterns of device wear. Figure 2 shows key examples of the heterogeneity of wear patterns across different hours of the day and days of the week for 3 anonymized patients.

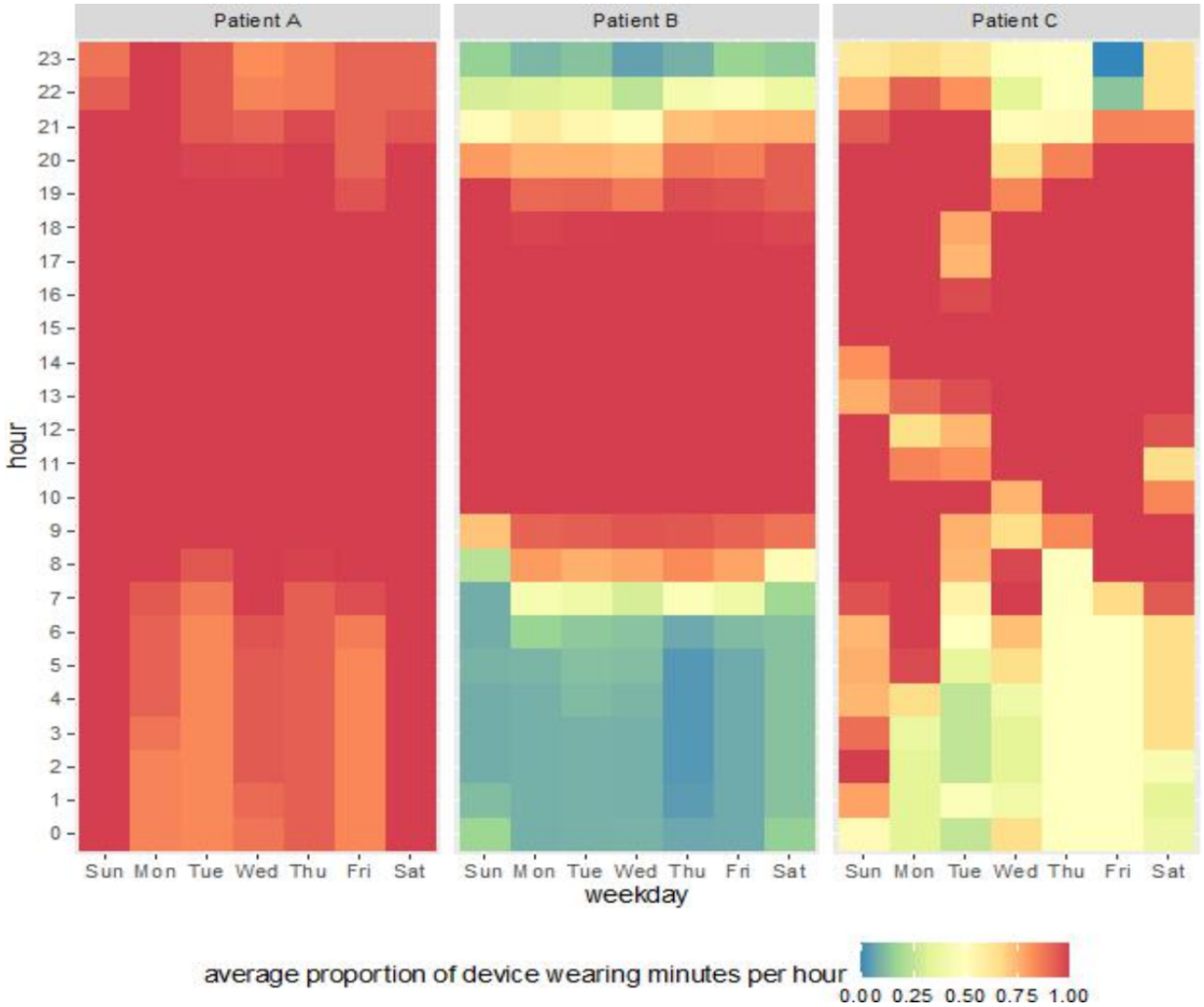
XSL•FO
RenderX

Figure 1. Number of patients enrolled (red) and number of patients exhibiting at least 10 hours of Fitbit wear time between 8 AM and 8 PM ("valid day"; blue) across their days in the study. Given that physical activity was assessed as a component of a primary cardiotoxicity investigation, the primary cardiotoxicity study end point has been indicated in the figure (black dotted line). Patients were informed that the Fitbit study monitoring period would last up to 3 years from enrollment, wherein data would be collected passively without intervention.



XSL•FO
RenderX

Figure 2. Proportion of device wear time during specific hours of days of the week across 3 anonymized patients.



PA During Device Wear

Table 2 summarizes the overall daily HR, step count, and proportion of time between 8 AM and 8 PM spent in the sedentary activity category among available valid days of device wear. The cohort was found to have a mean daily HR of 91 (SD 8.6) beats per minute and a mean daily step count of 2970 (SD 1902) steps. Moreover, patients spent a mean of 80% (SD 11%) of their day in a sedentary activity state. Figure 3A summarizes observed heterogeneity in PA across patients in terms of daily steps and HR during valid days for the entire study sample and for each patient individually. Table 3 describes Fitbit PA measures during patients' chemotherapy treatment period and after completion of chemotherapy. On average, and as expected, patients exhibited slightly less steps during chemotherapy (mean 2714, SD 1747 vs 3676, SD 2743 after), with comparable average HR and proportion of time spent sedentary. Importantly, we additionally found that 98.4% (62/63) of the patients had valid Fitbit wear during chemotherapy, whereas postchemotherapy summaries are based on only 68.3% (43/63) of the patients, who had valid Fitbit wear after completion of chemotherapy.

Figure 3B depicts the distribution of average daily step count across all participants exhibiting valid daily wear on each day of study follow-up, showing that the average daily step count among device wearers tended to increase over the course of the study. To further evaluate changes in daily step count over time, a mixed-effects regression model was fit with patient-specific random intercepts and slopes for follow-up time and adjustment for day of the week, patient sex, BMI, and tumor location, as well as whether the steps were recorded during or after chemotherapy treatment. Regression coefficient estimates from this model are reported in Table 4. This model indicated that, on average, across all patients, there was no significant trend in daily step counts as patients remained in the study for longer (P=.85), although there was substantial heterogeneity in patient-specific daily step trends—evidenced by a high variance of the patient-level random intercepts. The model also estimated that daily step count was 487 steps lower on average during chemotherapy treatment than after completion of chemotherapy (P<.001). With the exception of patient age, there was no evidence that any other patient characteristic or variable was significantly associated with daily step counts after accounting for chemotherapy, day of the week, and patient-specific time trends.

XSL•FO
RenderX

Table . Fitbit physical activity summary (n=63).

	Mean (SD; range)
Daily HR[a] (bpm[b])	91 (8.6; 69-118)
Daily steps	2970 (1902; 100-10,082)
Daily proportion of sedentary activity (%)	80 (0.11; 37-98)

[a]HR: heart rate.

[b]bpm: beats per minute.

Figure 3. (A) Summary of mean daily steps and heart rate: density of average daily heart rate (top left) and step count (top right) across all study participants and individual estimates and 95% CIs for mean daily heart rate (lower left) and daily step count (lower right) for individual study participants. (B) Daily steps by length of time in the study: box plots of steps per day for all participants exhibiting valid daily wear for each study follow-up day. The width of the box plots is proportional to the number of study patients logging Fitbit data on that follow-up day. bpm: beats per minute.

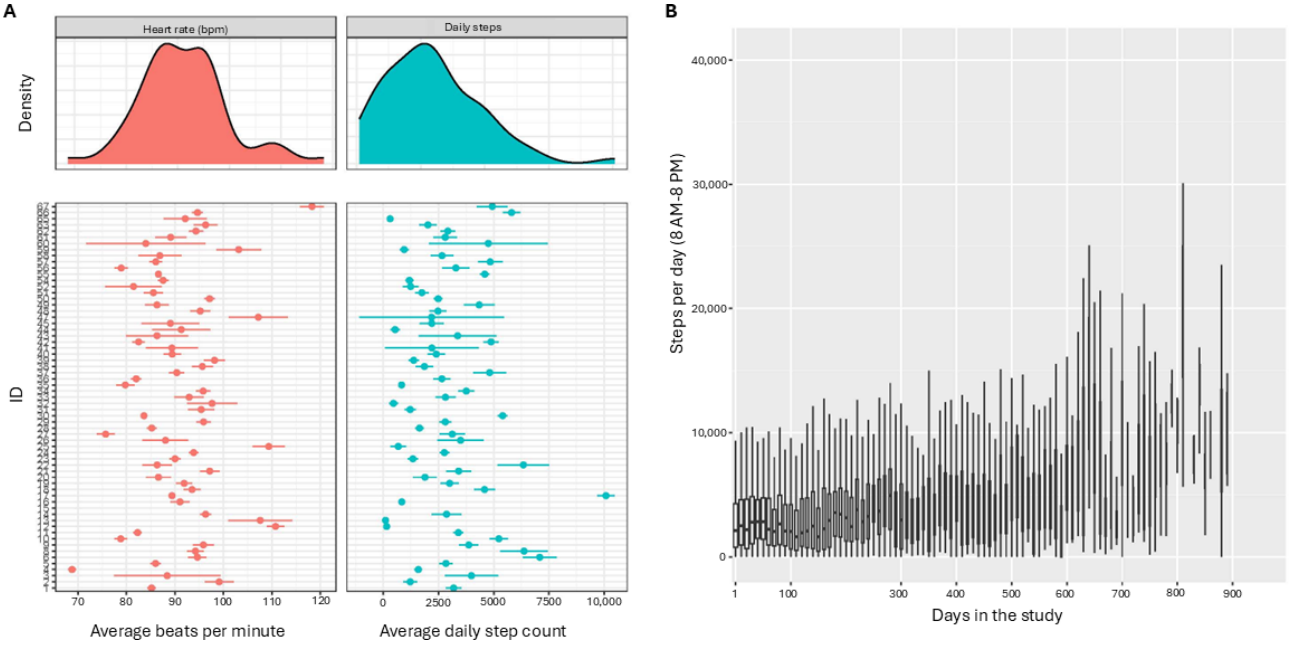


Table . Fitbit physical activity summary during and after chemotherapy.

	During chemotherapy (n=63)	After chemotherapy (n=43)
	Mean (SD; range)	Mean (SD; range)
Daily HR[a] (bpm[b])	91 (8.1; 70-114)	92 (11; 69-118)
Daily steps	2714 (1747; 82-9909)	3676 (2743; 105-12,533)
Daily proportion of sedentary activity	0.81 (0.11; 0.27-0.98)	0.79 (0.11; 0.56-0.98)

[a]HR: heart rate.

[b]bpm: beats per minute.

XSL•FO
RenderX

Table . Estimated regression coefficients from the linear mixed-effects model of daily steps during the study period[a].

Variable	Estimate (SE)	*P* value
Intercept	636.628 (1535.538)	.68
Days in the study	−0.345 (563.422)	.85
During chemotherapy	−487.326 (115.124)	<.001
Sex: male	625.104 (498.854)	.22
Age	77.911 (34.891)	.03
BMI	15.181 (25.771)	.56
Tumor location: lower limb	−184.588 (923.439)	.84
Tumor location: trunk	−286.307 (895.38)	.75
Tumor location: upper limb	1549.331 (1363.375)	.26

[a]Estimates represent the average difference in daily steps associated with a 1-unit increase in each individual feature controlling for all other variables presented in the model table, patient-specific trends in daily steps, and day of the week. The sex coefficient is the difference between male and female patients. Tumor location estimates are compared to "other" locations.

Discussion

Principal Findings

Given the popular desire by investigators to quantify longitudinal PA levels and changes in PA among AYA patients with cancer for the evaluation of the relationship between PA and cancer-related clinical outcomes, we investigated the practicality of assessing longitudinal PA using a wearable device in AYA patients with sarcoma who received high-dose doxorubicin in a free-living environment.

We found that, although participants were enrolled in the study for an average period of 630 (SD 208; median 730, IQR 622-730) days, PA wearable device data were logged for less than half of that time (mean 245, SD 226; median 205, IQR 58-352 days). Devices were worn and tolerated by an average of 57% of AYA patients for PA monitoring and data logging for at least 10 hours from 8 AM to 8 PM up to day 30 of study enrollment. Device wear declined drastically to 24% of patients on average after follow-up day 30. On the basis of AYA patient device wear patterns from this study, we identified that 30 days may be the upper limit for study duration in investigations that aim to use wearable devices such as Fitbit in similar cohorts, particularly without more active efforts to encourage device wear. We do not believe that these findings are exclusive to Fitbit device monitoring. Importantly, there is significant heterogeneity in the devices used in the field of PA monitoring [15]. We expect that our findings in AYA patients with sarcoma likely extend to the vast majority of devices. This study did not provide proactive reminders to participants to wear or synchronize the devices. Due to this passive monitoring modality, patients lacked routine motivational encouragement to wear their devices. Previous studies have demonstrated a significant positive impact of using SMS text message–, app-, and device-based reminders, notifications, and encouragement on PA-related behaviors, self-management, and adherence [24-26]. Therefore, motivational tactics should be considered for future investigations to modify behavioral patterns that augment device wear.

The practicality of using wearable devices in AYA patients with sarcoma appears to be greatly challenged by attrition and wearer difficulties. During the study period, qualitative reports of wearer and user challenges such as wearer discomfort, lost devices, and lack of data uploading compliance were identified as primary contributors to data missingness and participant attrition. Considering the high use of technology and mobile devices among AYA individuals [27,28], future studies that aim to monitor PA for longer periods should consider individual PA monitoring preferences and the use of digital health and mobile app alternatives rather than wearable devices to leverage current technology routinely used by this population [28]. Moreover, future studies should consider individual and treatment-related barriers such as side effects (eg, drug-induced skin sensitivity), psychological motivation, time availability, and technological challenges and preferences that AYA patients may face during monitoring periods that may largely impact adherence [14-16].

Our secondary longitudinal outcome evaluated what happened to the PA levels of AYA patients with sarcoma during and after exposure to high-dose chemotherapy. We found that, on average, across the subset of patients who persistently wore their devices both during and after chemotherapy, AYA patients performed 487 fewer steps during chemotherapy compared to after chemotherapy. These findings of reduced PA during chemotherapy are in agreement with those of several previous studies in the field [13,29,30]. After chemotherapy, the activity levels were observed to increase longitudinally over time (measured via step counts). However, such an inference is limited to the subset of patients who persistently wore their devices during and after chemotherapy. This key finding highlights the potential risk for self-selection or wearer bias and subsequent incorrect study conclusions. Individuals who are more physically active may be more likely to wear an activity-tracking device, which may result in inflation or overestimation of average PA. Future investigators should be wary of this inference limitation when collecting PA data using wearable devices.

XSL•FO
RenderX

Overall, individuals within the cohort demonstrated variable patterns of wear (length of time and days of the week) and heterogeneity in activity level. On average, the patients spent 80% (SD 11%) of their day sedentary. Interestingly, this cohort of AYA patients with sarcoma was found to have substantially lower PA levels (steps) than comparative reports. Previous studies have found that AYA patients with bone, soft tissue, and additional mixed cancer diagnoses undergoing variable therapies and procedures recorded an average of 5699 to 6434 steps per day and a notable range of patient variability [13]. In contrast, this cohort spent 80% (SD 11%) of their activity time between 8 AM and 8 PM sedentary and recorded an average of 2714 (SD 1747) steps during chemotherapy and 3676 (SD 2743) steps after chemotherapy completion. This difference in PA level may be explained by lifestyle behaviors, cancer location, surgery, radiation, and/or high anthracycline chemotherapy dosages. In total, 33.9% (21/62) of the patient population in our study had lower-extremity tumors, which may cause reduced lower body function and PA. Additionally, anthracycline chemotherapy has been observed to cause significant skeletal muscle weakness and fatigue [31]. Similar deleterious effects on skeletal muscle function and reduced PA are commonly reported after both radiation therapy and significantly reported after surgery.

Several limitations of this study should be considered, such as device selection, confounding factors, patient education and incentives, and potential biases. This study used the Fitbit Charge 3 device to quantify PA. While the Fitbit Charge 3 is known to be significantly correlated with research-grade devices such as the ActiGraph wGT3X-BT, it has been previously demonstrated that Fitbit may comparably overestimate PA [18]. The sedentary PA level among the cohort may be in part due to a combination of cancer location and fatigue. This study was not designed to determine the *impact* of either of these potentially confounding factors but, rather, report on PA tracker wear patterns and the resultant activity data and assess the usefulness of a PA monitoring device for future studies investigating the impact of PA on cancer-related clinical outcomes such as cardiac morbidity and fatigue. Moreover, this study was designed as an observational feasibility study. Our study was not designed with the intention to parse or investigate the cancer-related psychosocial or behavioral motivations that may impact the use of wearables. The psychological impact of cancer diagnoses and care should be considered in future wearable device research. Although device familiarization and patient education were provided for the use of wearable devices in this study, future studies should consider continuing education as a necessary component to ensure consistent use and study validity. This study did not offer patient incentives beyond allowing participants to keep the wearable device. Monetary or other incentives should be considered to increase adherence past 30 days and encourage PA throughout the study [32]. When interpreting patterns from wearable devices, it is essential to consider the possibility that device wear is related to PA, meaning that individuals who consistently adhere to wearing devices may differ systematically in their activity levels, potentially skewing data toward either higher or lower levels of PA. We restricted PA measurement to “valid days” of greater than or equal to 10 hours of wear during waking hours (8 AM-8 PM) to ensure that PA summaries actually represented daily behavior. Nonetheless, patients wore their devices for variable lengths of time throughout the study follow-up, implying that cross-sectional summaries at different study time points (eg, Table 3 and Figure 3B) cannot distinguish between behavior *changes* among persistent device wearers and behavioral *differences* between persistent wearers and patients who stopped wearing their devices. The mixed-effects regression model used in this study was designed to measure individual patient behavior changes over the time when the device was worn. The lack of significant trend in daily steps from this model suggests that individual patients did not tend to change their behavior over time. The apparent increase in daily steps among all patients with valid wear on a given follow-up day depicted in Figure 3B suggests that patients who wore their devices longer into the study were those who were more physically active.

Conclusions

This study demonstrated the need to identify alternative methods to a wearable device that would be more acceptable to AYA patients with cancer for monitoring PA longitudinally. On the basis of AYA patient device wear patterns from this study, we identified 30 days as a potential upper limit for study duration when using wearable devices such as Fitbit to accurately represent free-living environment PA without increased risk of biases. This study additionally confirmed poor PA levels among AYA patients with sarcoma. Importantly, future studies should consider AYA patient preferences for PA quantification and reporting, as well as study designs that implement continued patient education, regular participant notifications and reminders, and adherence incentives to ensure optimal longitudinal data collection and study success.

Funding

This work was supported in part by the Cancer Prevention and Research Institute of Texas (RP200381) and the Mary V. and John A. Reilly Distinguished Chair to ESK. SVR was partially supported by the Alan Ryden Endowment Fund and the Sanders Pediatric Sarcoma Research Fund.

Disclaimer

Generative artificial intelligence was not used for the preparation of this manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: SVR, SCG, AL, MER, ESK
Data curation: SVR, BC, MA, AL, MER, CMZ
Formal analysis: SVR, BC, MA, CMZ
Funding acquisition: ESK
Investigation: SVR, BC, SCG, AL, MER, CMZ, ESK
Methodology: SVR, BC, SCG, AL, MER, CMZ, ESK
Project administration: SCG, AL, MER, ESK
Resources: ESK
Supervision: AL, MER, ESK, CMZ
Visualization: SVR, BC, CMZ
Writing—original draft: SVR, BC
Writing—review and editing: SVR, BC, MA, SCG, AL, MER, CMZ, ESK

Conflicts of Interest

None declared.

References

1. Bhagat A, Kleinerman ES. Anthracycline-induced cardiotoxicity: causes, mechanisms, and prevention. Adv Exp Med Biol 2020;1257:181-192. [doi: 10.1007/978-3-030-43032-0_15] [Medline: 32483740]
2. Gaytan SL, Lawan A, Chang J, et al. The beneficial role of exercise in preventing doxorubicin-induced cardiotoxicity. Front Physiol 2023 Mar 9;14:1133423. [doi: 10.3389/fphys.2023.1133423] [Medline: 36969584]
3. Kang DW, Wilson RL, Christopher CN, et al. Exercise cardio-oncology: exercise as a potential therapeutic modality in the management of anthracycline-induced cardiotoxicity. Front Cardiovasc Med 2021;8:805735. [doi: 10.3389/fcvm.2021.805735] [Medline: 35097024]
4. Scott JM, Nilsen TS, Gupta D, Jones LW. Exercise therapy and cardiovascular toxicity in cancer. Circulation 2018 Mar 13;137(11):1176-1191. [doi: 10.1161/CIRCULATIONAHA.117.024671] [Medline: 29530893]
5. St Fleur RG, St George SM, Leite R, Kobayashi M, Agosto Y, Jake-Schoffman DE. Use of Fitbit devices in physical activity intervention studies across the life course: narrative review. JMIR Mhealth Uhealth 2021 May 28;9(5):e23411. [doi: 10.2196/23411] [Medline: 34047705]
6. Tang LM, Kay J. Harnessing long term physical activity data—how long-term trackers use data and how an adherence-based interface supports new insights. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017 Jun;1(2):1-28. [doi: 10.1145/3090091]
7. Speier W, Dzubur E, Zide M, et al. Evaluating utility and compliance in a patient-based eHealth study using continuous-time heart rate and activity trackers. J Am Med Inform Assoc 2018 Oct 1;25(10):1386-1391. [doi: 10.1093/jamia/ocy067] [Medline: 29850807]
8. Henriksen A, Sand AS, Deraas T, Grimsgaard S, Hartvigsen G, Hopstock L. Succeeding with prolonged usage of consumer-based activity trackers in clinical studies: a mixed methods approach. BMC Public Health 2020 Aug 27;20(1):1300. [doi: 10.1186/s12889-020-09406-w] [Medline: 32854671]
9. Sher J, Lewis CW, Lin C. Using wearable technologies to monitor physical activity and exercise in patients: a narrative review. Curr Sports Med Rep 2024 Aug 1;23(8):284-289. [doi: 10.1249/JSR.0000000000001187] [Medline: 39102583]
10. Kajzar M. Wearable devices for training and patient monitoring: a comprehensive review. Qual Sport 2024;29:55667. [doi: 10.12775/qs.2024.29.55667]
11. Butte NF, Ekelund U, Westerterp KR. Assessing physical activity using wearable monitors: measures of physical activity. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S5-S12. [doi: 10.1249/MSS.0b013e3182399c0e] [Medline: 22157774]
12. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. Int J Environ Res Public Health 2021 Feb 19;18(4):2023. [doi: 10.3390/ijerph18042023] [Medline: 33669679]
13. Munsie C, Ebert J, Joske D, Ackland T. The benefit of physical activity in adolescent and young adult cancer patients during and after treatment: a systematic review. J Adolesc Young Adult Oncol 2019 Oct;8(5):512-524. [doi: 10.1089/jayao.2019.0013] [Medline: 31090475]
14. Singh B, Zopf EM, Howden EJ. Effect and feasibility of wearable physical activity trackers and pedometers for increasing physical activity and improving health outcomes in cancer survivors: a systematic review and meta-analysis. J Sport Health Sci 2022 Mar;11(2):184-193. [doi: 10.1016/j.jshs.2021.07.008] [Medline: 34314878]
15. Keats MR, Yu X, Sweeney Magee M, et al. Use of wearable activity-monitoring technologies to promote physical activity in cancer survivors: challenges and opportunities for improved cancer care. Int J Environ Res Public Health 2023 Mar 8;20(6):4784. [doi: 10.3390/ijerph20064784] [Medline: 36981693]

XSL•FO
RenderX

16. Jermolenko GF, Semjonova G, Klavina A, et al. Measurements and digital technology solutions to monitor physical activity in patients with pediatric cancer: scoping review. JMIR Cancer 2026 Jan 29;12:e73889. [doi: 10.2196/73889] [Medline: 41611218]
17. Viguet C, Hughes A, Banchs J, et al. Abstract 4361008: longitudinal echocardiographic assessment of subacute cardiotoxicity in a prospective cohort of adolescents and young adults with sarcoma treated with high-dose doxorubicin. Circulation 2025 Nov;152(Suppl_3). [doi: 10.1161/circ.152.suppl_3.4361008]
18. Matlary RE, Holme PA, Glosli H, Rueegg CS, Grydeland M. Comparison of free-living physical activity measurements between ActiGraph GT3X-BT and Fitbit Charge 3 in young people with haemophilia. Haemophilia 2022 Nov;28(6):e172-e180. [doi: 10.1111/hae.14624] [Medline: 35830613]
19. Matthews CE, Hagströmer M, Pober DM, Bowles HR. Best practices for using physical activity monitors in population-based research. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S68-S76. [doi: 10.1249/MSS.0b013e3182399e5b] [Medline: 22157777]
20. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364. [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]
21. Yang Y, Chen L, Filippidis FT. Accelerometer-measured physical activity, frailty, and all-cause mortality and life expectancy among middle-aged and older adults: a UK Biobank longitudinal study. BMC Med 2025 Feb 27;23(1):125. [doi: 10.1186/s12916-025-03960-z] [Medline: 40016761]
22. Cui S, Lin Q, Gui Y, et al. CARE as a wearable derived feature linking circadian amplitude to human cognitive functions. NPJ Digit Med 2023 Jul 11;6(1):123. [doi: 10.1038/s41746-023-00865-0] [Medline: 37433859]
23. Narai E, Hagino H, Komatsu T, Togo F. Accelerometer-based monitoring of upper limb movement in older adults with acute and subacute stroke. J Geriatr Phys Ther 2016;39(4):171-177. [doi: 10.1519/JPT.0000000000000067] [Medline: 26428901]
24. Marin TS, Kourbelis C, Foote J, et al. Examining adherence to activity monitoring devices to improve physical activity in adults with cardiovascular disease: a systematic review. Eur J Prev Cardiol 2019 Mar;26(4):382-397. [doi: 10.1177/2047487318805580] [Medline: 30322268]
25. Mutsuddi A, Connelly K. Text messages for encouraging physical activity are they effective after the novelty effect wears off? Presented at: 6th International Conference on Pervasive Computing Technologies for Healthcare; May 21-24, 2012. [doi: 10.4108/icst.pervasivehealth.2012.248715]
26. Heizmann AN, Chapelle C, Laporte S, Roche F, Hupin D, Le Hello C. Impact of wearable device-based interventions with feedback for increasing daily walking activity and physical capacities in cardiovascular patients: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2023 Jul 11;13(7):e069966. [doi: 10.1136/bmjopen-2022-069966] [Medline: 37433730]
27. Dicker AP, Jim HS. Intersection of digital health and oncology. JCO Clin Cancer Inform 2018 Dec;2:1-4. [doi: 10.1200/CCI.18.00070] [Medline: 30652577]
28. Devine KA, Viola AS, Coups EJ, Wu YP. Digital health interventions for adolescent and young adult cancer survivors. JCO Clin Cancer Inform 2018 Dec;2:1-15. [doi: 10.1200/CCI.17.00138] [Medline: 30652583]
29. Murnane A, Gough K, Thompson K, Holland L, Conyers R. Adolescents and young adult cancer survivors: exercise habits, quality of life and physical activity preferences. Support Care Cancer 2015 Feb;23(2):501-510. [doi: 10.1007/s00520-014-2446-2] [Medline: 25294658]
30. Winter C, Müller C, Brandes M, et al. Level of activity in children undergoing cancer treatment. Pediatr Blood Cancer 2009 Sep;53(3):438-443. [doi: 10.1002/pbc.22055] [Medline: 19415742]
31. Gilliam LA, St. Clair DK. Chemotherapy-induced weakness and fatigue in skeletal muscle: the role of oxidative stress. Antioxid Redox Signal 2011 Nov;15(9):2543-2563. [doi: 10.1089/ars.2011.3965] [Medline: 21457105]
32. Fukaya E, Welden S, Bukari A, Khan Z, Leeper N, Mohler E. Incentivizing physical activity through activity monitoring interventions in PAD - a pilot study. Vasa 2021 Feb;50(2):145-150. [doi: 10.1024/0301-1526/a000924] [Medline: 33150850]

Abbreviations

AYA: adolescent and young adult
HR: heart rate
PA: physical activity

XSL•FO
RenderX

Edited by M Balcarras; submitted 11.Nov.2025; peer-reviewed by M Yamakita, R Taylor; revised version received 05.Mar.2026; accepted 05.Mar.2026; published 13.Apr.2026.

Please cite as:
Rauschendorfer SV, Cantoni B, Audirac M, Gilchrist SC, Livingston JA, Roth ME, Zigler CM, Kleinerman ES
Physical Activity Monitoring in Adolescents and Young Adults With Cancer: Observational Feasibility Study
JMIR Cancer 2026;12:e87591
URL: https://cancer.jmir.org/2026/1/e87591
doi:10.2196/87591

© Savannah V Rauschendorfer, Beatrice Cantoni, Michelle Audirac, Susan C Gilchrist, J Andrew Livingston, Michael E Roth, Corwin M Zigler, Eugenie S Kleinerman. Originally published in JMIR Cancer (https://cancer.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Efficacy of Telehealth-Based Coaching to Improve Physical Activity and Overall Experience for Cancer Survivors: Secondary, Mixed Methods Analysis of a Randomized Controlled Trial

Saif Khairat[1,2,3,4], PhD, MPH; Erin E Kent[3,5], PhD, MS; John Geracitano[2], MS; Kaushalya Mendis[1], MD; Zhaoqiang Zhou[2], MS; Carly Bailey[3], MA; William A Wood[3,6], MD, MPH

[1]School of Nursing, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]Carolina Health Informatics Program, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[3]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[4]Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[5]Department of Health Policy and Management, Gillings School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[6]Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
William A Wood, MD, MPH
Lineberger Comprehensive Cancer Center
University of North Carolina at Chapel Hill
321 S Columbia Street
Chapel Hill, NC, 27599
United States
Phone: 1 9199628331
Email: william_wood@med.unc.edu

Abstract

Background: Cancer survivors face significant challenges in maintaining adequate physical activity levels, which are essential for overall health and quality of life. Telehealth-based interventions offer promising opportunities to provide accessible support and promote healthier lifestyles throughout the cancer survivorship continuum. HealthScore is a telehealth coaching program designed to optimize the health of cancer survivors.

Objective: This study assessed the effectiveness of HealthScore in improving physical activity metrics among cancer survivors compared to controls. We also evaluated participants' qualitative experiences with the program to understand its impact on motivation, accountability, and overall health-related quality of life.

Methods: We performed a secondary analysis of a randomized controlled study of cancer survivors who participated in a comprehensive health coaching intervention called HealthScore. Participants in control and intervention groups received a Fitbit activity tracker that collected heart rate, step counts, active minutes, and calories burned. These metrics were analyzed using statistical methods to compare overall averages and temporal trends between intervention and control groups. Eleven exit interviews were conducted with intervention arm participants to ascertain their experiences with HealthScore. Inductive thematic analysis was performed to identify emerging themes. Data were collected between May 2020 and March 2022.

Results: Of the 32 participants enrolled, 20 (62%) were in the intervention group. Compared to the control group, intervention participants had significantly higher average daily steps (mean 3660, SD 3344; 95% CI 3557-3764 vs mean 3408, SD 3288; 95% CI 3299-3518; P=.001) and more moving average daily steps (mean 4813, SD 1723; 95% CI 4680-4946 vs mean 4581, SD 1224; 95% CI 4494-4669; P=.003). Moving average daily step counts in the intervention arm showed an increasing trend, which was significantly higher than that of the control group (regression slope=5.89 vs 2.80; P<.001). Compared to the control group, the intervention participants had significantly higher average daily walking distance (mean 2.6, SD 2.5; 95% CI 2.5-2.7 vs mean 2.4, SD 2.3; 95% CI 2.3-2.5; P<.001) and more moving average daily walking distance (mean 3.5, SD 1.3; 95% CI 3.4-3.6 vs mean 3.2, SD 0.8; 95% CI 3.1-3.3; P<.001). Moving average daily walking distances among intervention participants increased, which was also significantly higher than that of the control group (regression slope=0.0046 vs 0.0017; P<.001). Participants in the intervention group reported a growing sense of accountability and motivation. One barrier was completing weekly monitoring

XSL•FO
RenderX

of patient-reported outcome surveys, which focused on symptoms and physical function and did not always align with participants' goals.

Conclusions: The HealthScore telehealth coaching program improved physical activity levels among cancer survivors and enhanced motivation and accountability. These findings support the integration of telehealth-based health coaching into posttreatment care, promoting healthier lifestyles and improved quality of life for cancer survivors.

Trial Registration: ClinicalTrials.gov NCT04923997; https://clinicaltrials.gov/study/NCT04923997

(JMIR Cancer 2026;12:e78968) doi:10.2196/78968

KEYWORDS

physical activity; quality-of-life; cancer survivor; telehealth; health coaching; digital health; Fitbit; wearable; step count; walking distance

Introduction

Maintaining adequate physical activity levels poses a significant challenge for cancer survivors [1]. The toxicity associated with cancer treatments often results in fatigue, pain, and decreased physical function, which can discourage individuals from participating in regular exercise [2,3]. Additionally, psychological symptoms, such as anxiety and depression, further impede motivation and adherence to physical activity regimens [4]. The lack of tailored programs for cancer survivors increases these challenges, making it harder for individuals to incorporate physical activity consistently into their routines [5].

Low physical activity levels among cancer survivors can have detrimental effects on their overall well-being and health care outcomes [1,6]. Inactivity has been associated with reduced health-related quality of life (HRQOL), increased fatigue, and higher anxiety and depression scores [7]. Additionally, a sedentary lifestyle is linked to higher risks of comorbidities, complicating the health status of cancer survivors [8]. Conversely, regular physical activity can alleviate treatment-related adverse effects, reduce cancer-specific and overall mortality, and improve HRQOL [9]. Furthermore, studies have demonstrated that telehealth exercise-based interventions can significantly enhance cardiorespiratory fitness, quality of life (QOL), and physical activity levels in cancer survivors [10,11]. Home-based physical activity interventions have been shown to be safe, with very low adverse event rates [12]. Finally, exercise has been recommended to cancer survivors since 2019 [13]. Therefore, promoting and maintaining adequate physical activity is essential for improving overall health among cancer survivors.

Research has shown that health coaching can be an effective method to promote physical activity among cancer survivors [14-16]. By definition, health coaching involves participant-led personalized support and guidance to help individuals set and achieve health-related goals [7]. In general, health coaching can lead to significant improvements in the frequency and intensity of physical activity, thereby improving overall physical activity levels [17,18]. However, to date, many health coaching programs have not shown improvements in physical activity among those with cancer, largely due to lower intervention intensity or low quality of coaching [19,20].

Telehealth-based health coaching has emerged as a promising solution to address these challenges by providing ongoing, personalized support with remote convenience. Telehealth allows for higher accessibility for individuals who may not have access to in-person services [21]. Telehealth coaching combined with remote symptom monitoring can effectively increase patients' physical activity levels, improve biomarkers associated with diabetes, and reduce body weight [4,22]. Health coaching is distinct among behavioral interventions for cancer survivors for a number of reasons: (1) its emphasis on improving patient activation and motivation, (2) its ability to be delivered by trained and supervised lay health coaches, and (3) the extent to which participants drive the direction. Coaching is less scripted than other behavioral interventions in general, given the focus on encouraging participants to take ownership of their health. HealthScore places emphasis on improving motivation, mood, mindfulness, and movement (the 4M model) [23]. However, the effect of telehealth coaching on the physical activity levels of cancer survivors is unknown.

HealthScore is a telehealth coaching program that focuses on optimizing the overall health of cancer survivors (ie, from diagnosis through end-of-life) through (1) weekly, structured coaching sessions and (2) physiologically based patient-generated health data (PGHD) capture to enable comprehensive support to promote health. Preliminary work has demonstrated high levels of acceptability and feasibility of the HealthScore intervention [24]. The purpose of this study was to assess HealthScore's potential for improving physical activity and report on the experiences of intervention participants.

Methods

Overview

We conducted a secondary analysis of a randomized controlled pilot study [24] of cancer survivors who participated in a comprehensive health coaching intervention called HealthScore, using an explanatory sequential mixed methods approach [25]. This study was a registered clinical trial (NCT04923997). The pilot study assessed the feasibility and acceptability of the HealthScore program by measuring changes in physical function from baseline, evaluations conducted at 3 and 6 months, and patient-reported outcomes (PROs) as secondary outcomes. This secondary analysis assesses HealthScore's potential for improving physical activity by examining daily step count and walking distance, coupled with thematic analysis of participant exit interviews. For this study, we adopted the National Cancer

XSL•FO
RenderX

Institute (NCI) definition of cancer survivor: "An individual is considered a cancer survivor from the time of diagnosis through the balance of life" [26]. Data were collected between May 2020 and March 2022.

HealthScore's key components consist of weekly meetings with coaches to develop and clarify goals, the collection of PGHD from Fitbit activity tracking, and weekly participant-completed surveys centered on physical function, HRQOL, and symptom burden. Participants in the intervention group completed weekly sessions with a trained health coach who guided them in creating individualized SMART (Specific, Measurable, Achievable, Relevant, and Time-Bound) goals, reviewed their survey responses, including a graphic display of their self-reported physical function metrics over time (labeled their Health Score), and, when needed, connected them to additional supportive care resources [27]. Health coaches were a combination of 2 full-time cancer center staff and 12 volunteer health coaches. All coaches received training in motivational interviewing strategies, common symptoms affecting cancer survivors, and goal setting to facilitate focused, goal-oriented sessions. Coaches had participated in a series of didactic presentations surrounding the transtheoretical model of behavior change, coaching foundations, cancer center resources, and other health-related topics, including exercise, nutrition, and sleep. Training videos and manuals were developed by a board-certified health coach and maintained for reference [24].

The team was formally trained in conducting interviews and comprised of medical professionals and 12 interdisciplinary volunteer coaches. The majority of coaches were recruited through the National Board of Health and Wellness Coaches job board, had already completed an accredited program, and were planning to pursue the National Board examination for health and wellness coaching. Coaches were asked to complete a HealthScore-specific training, which included human participants protections, 5 hour-long HealthScore coaching trainings and associated assessments, and practice coaching sessions with the same board-certified staff coach. Having a singular coach trainer provided consistency across the multiple volunteer health coaches. Coaches were not permitted to work with patients who participated in HealthScore until the study team and the volunteer coach were confident that the volunteer had sufficient coaching skills and understanding of cancer center resources. Coaches were required to attend monthly team meetings to ensure they were up to date with study progress and to have a shared space where they could brainstorm participant-specific coaching challenges together. All volunteer coaches were supervised by a board-certified health coach.

The intervention group completed semistructured exit interviews after the intervention ended at 6 months. Exit interviews lasted 30-45 minutes and elicited participants' perspectives about the impact of the HealthScore intervention on their health goals and outcomes. The objective of the current analysis was to report findings from both the objective Fitbit physical activity data and intervention participant exit interview themes and to provide insight into HealthScore's mechanisms of action and refine the intervention for a future fully powered efficacy trial.

Study Design

To promote the autonomy and self-efficacy of participants and improve their long-term QOL, the HealthScore program is grounded on self-determination theory [28]. Additionally, the principles and the transtheoretical model of behavior change were supplemented by motivational interview techniques [29,30].

Patients were either referred to our program by clinical oncology teams, identified through screening of clinical visits, or self-referred to our program in response to the University of North Carolina (UNC) Research For Me listing. Patients across a variety of cancer types with advanced cancer staging were recruited. Once patients were identified, they were screened to ensure they met the inclusion criteria. Information, such as cancer type and treatment status, was collected; however, this information was used for evaluation of study eligibility and not used as exclusion criteria based on type and treatment status. If eligible, a study team member contacted them via phone to gauge interest, provide more information about the study. Once consented, patients completed the PROMIS (Patient-Reported Outcomes Measurement Information System) Physical Function 8b, [31] which evaluates limitations in physical activities (eg, mobility and extremity function). Responses were stratified into high or low physical function and then randomized into either the intervention or a waitlist control arm to study the impact of the intervention on the participants' physical function, HRQOL, and physical activities as measured by weekly averages of daily step counts. Participants were randomized by the research team and stratified at the median of baseline PROMIS physical function scores of (45+ and <45). The final number of participants included 20 in the intervention group and 12 in the control group, totaling 32 participants for the entire cohort. While there were 46 total participants in the parent trial, 14 participants were excluded from this secondary analysis due to missing or abnormal data. The TIDieR (Template for Intervention Description and Replication) and the CONSORT (Consolidated Standards of Reporting Trials) checklists were referenced to describe this study (Multimedia Appendices 1 and 2, respectively). Qualitative methods were guided and reported in compliance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) reporting guideline [32].

Measures

Quantitative Data Collection

Participants in control and intervention groups received a Fitbit activity tracker that collected metrics, such as heart rate, step counts, active minutes, and calories burned. The study team provided training and troubleshooting for participants, if needed. Data from each device were automatically sent via an intermediary application programming interface to the UNC Connected Health for Applications & Interventions (CHAI) Core, a UNC-developed secure data collection system used by the study team. CHAI Core created a HealthScore platform for coaches and participants to provide participant monitoring, generate reports, visualize data capture (including the PROMIS Physical Function measure, referred to as the HealthScore), and serve as a platform for collaborative goal setting. Survey

XSL•FO
RenderX

responses were collected and aggregated through another internal platform called PRO-Core.

Qualitative Data Collection

A semistructured interview guide (Multimedia Appendix 3) designed to synthesize themes of participants' perspectives of the HealthScore program was collaboratively developed by the study team and pilot tested during earlier phases of HealthScore. This guide structured questions around patients' perceived facilitators and barriers to participation, their interactions with their health coach, thoughts on the program's metrics, and caregiver involvement. Intervention group participants were asked about their perceptions after completing the one-time 6-month HealthScore program and its components, such as the Fitbit tracker, short- and long-term benefits, and recommendations for changes to HealthScore.

Study team members conducted interviews using web-based videoconferencing platforms or telephonically. The 30-minute interviews were audio recorded and transcribed, and representative quotes were identified. Field notes were captured during the sessions to amplify recordings. Transcripts were not shown or returned to participants for comment or correction.

Outcomes

This paper's analytical outcomes were (1) physical activity (daily step count and daily walking distance over the week) and (2) perceptions of participants in the intervention group elicited from exit interviews with intervention participants.

Data Analysis

Quantitative Analysis

Frequencies and means describing participant characteristics in both arms were calculated. A pragmatic methodological approach was undertaken for this mixed methods study [25].

For assessment of physical activity, we used the records of each participant's daily step count captured from participant-worn Fitbits. We calculated moving averages of daily step count for each participant using a window size of 7 days, as we inferred that the data collected from the participants tended to change over the course of the week, and this interval matched the frequency of weekly coach calls. The missing data and abnormal data (negative step counts) were removed (18%). The moving average of daily step count between participants in the intervention and control arms was compared, and *t* test was applied to detect the difference between them. Then, with the time change as the independent variable and the moving averages per participant as the dependent variable, linear regression models were fit to further analyze the trend of step counts per participant during this experimental phase, helping to understand how participants' adherence and motivation for physical activity changed over time. The same procedures were also applied to the analysis of walking distance. A *P* value less than .05 was considered significant.

Qualitative Analysis

For the participant perspectives of the intervention, an inductive qualitative research approach was used for analysis, with each unique participant serving as the unit of analysis [33]. The qualitative sample was limited to those participants for whom full interviews were available, totaling 11 participants overall. The interview sessions were automatically transcribed and then analyzed using Dedoose, a qualitative management web-based tool [34]. The research team followed a phased approach to ensure rigor and reproducibility in our analysis. Two coders (JG and KM) developed a codebook of 25 unique codes based on the initial readings of transcripts reviewed and standardized to code the interview transcripts [35]. Codes were focused on facilitators, barriers, and suggestions for improvement. We then reviewed codes to identify broader themes and refined them to ensure no key aspects of the data were overlooked. Finally, we captured direct quotes to provide meaningful insights that are aligned with our study's aim. Participants did not provide feedback on our findings.

Mixed Methods Integration

Following the explanatory sequential mixed methods design, quantitative and qualitative findings were integrated during the interpretation phase to provide a comprehensive understanding of HealthScore's intervention effectiveness. Integration involved comparing quantitative physical activity outcomes (eg, step counts) with qualitative themes. This integration aimed to explain how and why the intervention achieved its quantitative outcomes by examining participants' perspectives on program components, such as accountability and measurement tools. This integrated analysis facilitated a richer interpretation of the findings and leveraged the GRAMMS (Good Reporting of a Mixed Methods Study) checklist (Multimedia Appendix 4).

Ethical Considerations

The IRB (Institutional Review Board) of the UNC at Chapel Hill approved the study and its use of electronic consent (IRB number 20-0051). Consents were emailed and signed via UNC REDCap (Research Electronic Data Capture; Vanderbilt University). All study data complied with institutional guidelines and were deidentified. Participants provided written informed consent and received a US $20 gift card as compensation for returning questionnaires at each study milestone.

Results

Overview

There were 163 patients approached to participate in the HealthScore program (Figure 1). Of those, 53 consented, and 46 ultimately enrolled. Participants were randomized and stratified at the median of baseline PROMIS physical function scores of (45+ and <45). The final number of participants included 20 participants in the intervention group and 12 participants in the control group for a total of 32 participants (Figure 1). No adverse events related to exercise were noted.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Criteria for Reporting Qualitative Research) flow diagram.

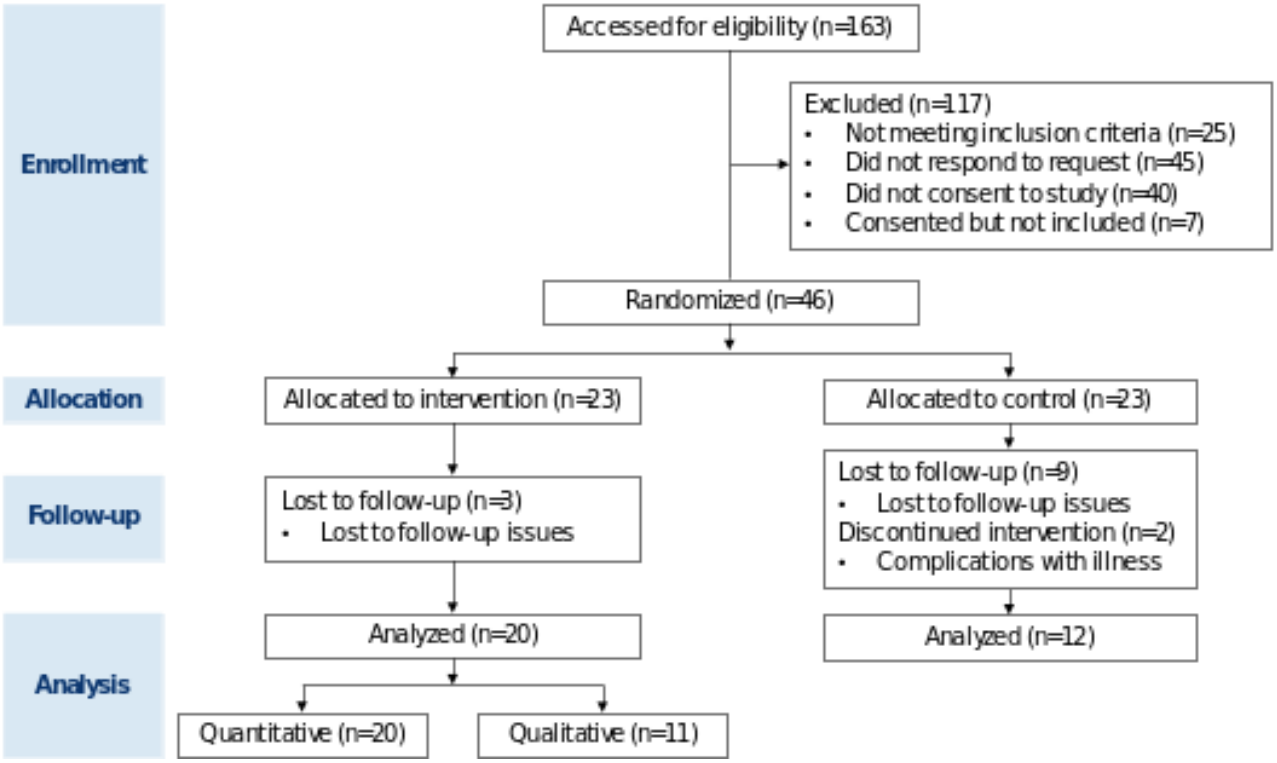


Table 1 provides participant demographics (N=32). Of the 20 intervention participants, 13 (65%) were female, 17 (85%) were Non-Hispanic White, and 11 (55%) were between 60 and 79 years old. Of 12 participants in the control group, 9 (75%) were female, 6 (50%) were White, 4 (33%) were between 40 and 59 years old, and 4 (33%) were between 60 and 79 years old. Baseline demographic data did not significantly differ between intervention and control participants (Table 1). Furthermore, within the intervention group, the participant demographics of the 11 interviewees were not significantly different from those of the other 9 intervention group participants. Multimedia Appendix 5 includes a demographics table specifically for the 11 interviewees.

XSL•FO
RenderX

Table 1. Demographics of the finally included participants.

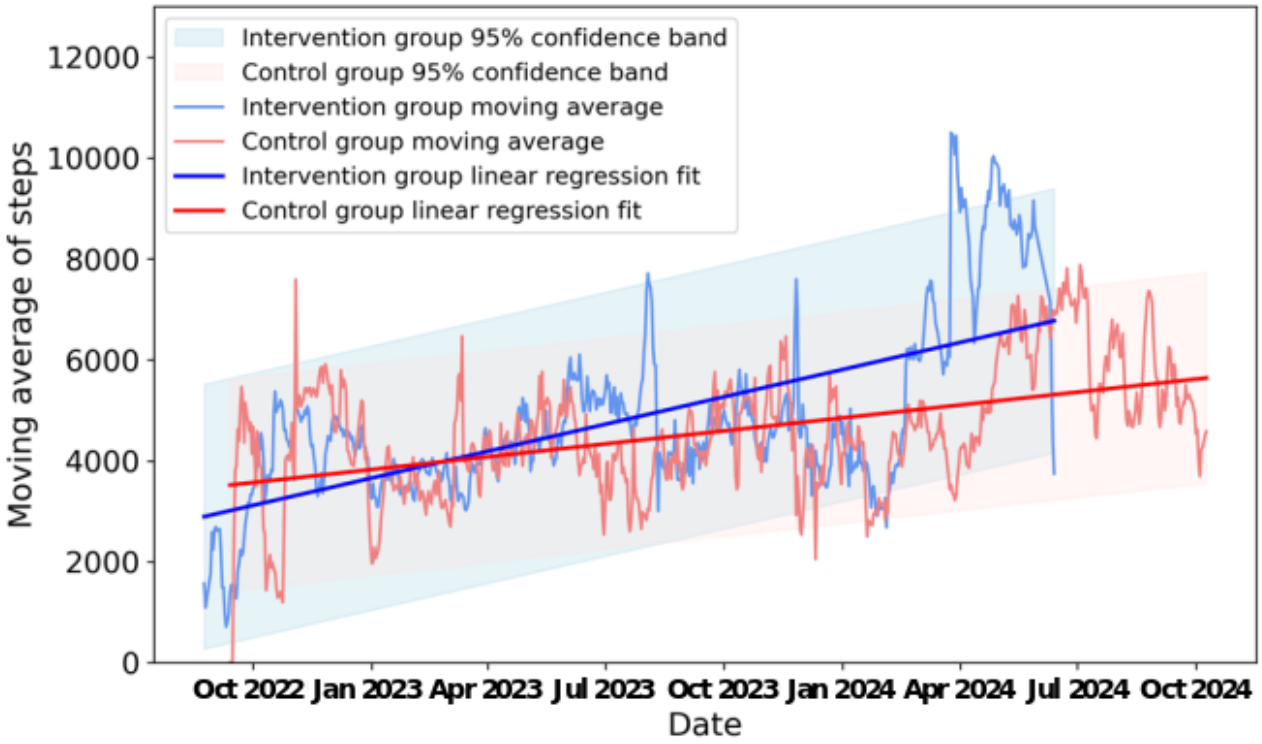
Characteristic	Participants (N=32)	Intervention group (n=20)	Control group (n=12)	*P* value
Age (years)				.50
Mean (SD)	58.7 (12.5)	59.2 (10.8)	57.7 (16.3)	
Range	23-76	29-75	23-76	
Age group (years), n (%)				.35
20-39	2 (6)	1 (5)	1 (8)	
40-59	11 (34)	7 (35)	4 (33)	
60-79	15 (47)	11 (55)	4 (33)	
Not specified	4 (13)	1 (5)	3 (25)	
Sex , n (%)				.26
Female	22 (69)	13 (65)	9 (75)	
Male	9 (28)	7 (35)	2 (17)	
Not specified	1 (3)	0 (0)	1 (8)	
Race , n (%)				.16
Non-Hispanic White	24 (75)	17 (85)	6 (50)	
Non-Hispanic Black	4 (13)	2 (10)	3 (25)	
More than one race	2 (6)	0 (0)	2 (17)	
Not specified	2 (6)	1 (5)	1 (8)	
Ethnicity , n (%)				≥.99
Hispanic	2 (6)	1 (1)	1 (8)	
Non-Hispanic	30 (93)	19 (95)	11 (92)	
Education , n (%)				.51
9th-12th grade (no diploma)	1 (3)	0 (0)	1 (9)	
High school graduate or equivalent	2 (6)	1 (5)	1 (9)	
Some college (no degree)	5 (16)	4 (20)	1 (9)	
Vocational or associate’s degree	7 (22)	4 (20)	3 (27)	
Bachelor’s degree	10 (31)	8 (40)	2 (18)	
Higher than a bachelor’s degree	6 (19)	3 (15)	3 (27)	
Employment, n (%)				.62
Employed full-time	4 (12)	2 (10)	2 (18)	
Unemployed, because of illness	5 (16)	3 (15)	2 (18)	
On disability	7 (23)	6 (30)	1 (9)	
Retired	14 (45)	8 (40)	6 (55)	
Other	1 (3)	1 (5)	0 (0)	
Marital status , n (%)				.71
Single, never married	2 (7)	1 (5)	1 (9)	
Married or partnered	20 (67)	12 (63)	8 (73)	
Separated	3 (10）	2 (11)	1 (9)	
Divorced	3 (10)	3 (16)	0 (0)	
Widowed	2 (7)	1 (5)	1 (9)	

XSL•FO
RenderX

Quantitative Results: Physical Activity

Participants in the intervention group had an overall average number of 3660 (SD 3344; 95% CI 3557-3764) daily steps, which was significantly more than participants in the control group, who had an overall average number of 3408 (SD 3288; 95% CI 3299-3518; t_{7476}=3.3; *P*=.001) daily steps. The comparison of weekly moving averages showed similar results. The weekly moving average step count of participants in the intervention group was 4813 (SD 1723; 95% CI 4680-4946), which was significantly more than that of the control group 4581 (SD 1224; 95% CI 4494-4669; t_{1400}=2.9; *P*=.003). Weekly moving average step counts in the intervention arm showed an increasing trend as the study progressed, with a regression slope parameter of 5.89 (*P*<.001). In contrast, the control weekly average step counts increased with a regression slope parameter of 2.80 (*P*<.001) but at a lower rate than the intervention arm. The slope of the line of best fit of the intervention group is steeper than that of the control group (Figure 2).

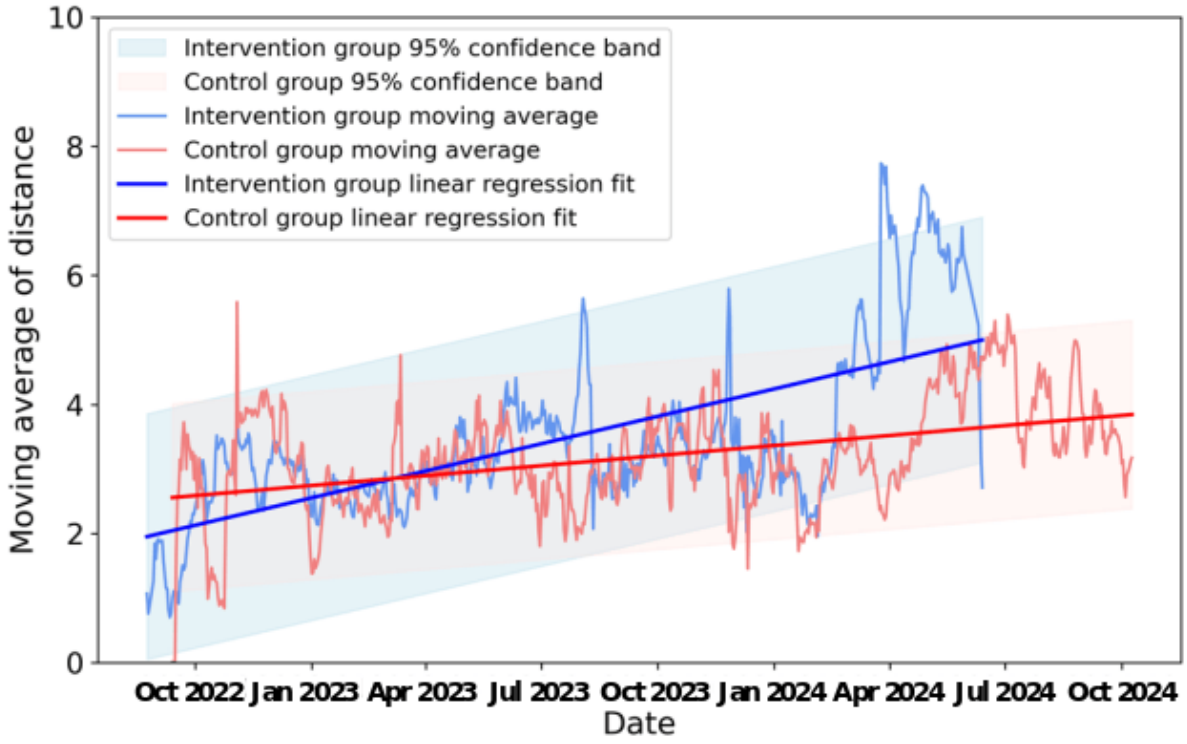
Figure 2. Linear regression of the weekly moving average steps over date per participant.



For the daily walking distance, intervention participants had an overall average daily walking distance of 2.6 (SD 2.5*;* 95% CI 2.5-2.7) miles, which was significantly more than control participants, who had a weekly average of daily walking distance of 2.4 (SD 2.3; 95% CI 2.3-2.5; t_{7476}=3.8; *P*<.001) miles. The comparison of the weekly moving average showed similar results. The weekly moving average walking distance of intervention participants was 3.5 (SD 1.3; 95% CI 3.4-3.6) miles, which was significantly more than that of control participants, 3.2 (SD 0.8; 95% CI 3.1-3.3; t_{1400}=4.6; *P*<.001) miles. Weekly moving average walking distances among intervention participants increased, with a regression slope parameter equal to 0.0046 (*P*<.001). In contrast, though the weekly average of daily walking distances also increased for controls, the regression slope parameter was smaller than that of the intervention group, less than half of that, at 0.0017 (*P*<.001). The slope of the line of best fit of the intervention group is steeper than that of the control group (Figure 3). Thus, walking distances of the intervention participants showed greater increases than those of the controls.

XSL•FO
RenderX

Figure 3. Linear regression of the weekly moving average of walking distance over date per participant.



Qualitative Results

Participant Perspectives of the HealthScore Intervention

The inductive coding and analysis of 11 intervention participant interviews revealed 3 major thematic areas, including facilitators and barriers to participant engagement and program improvement. Each theme and subtheme was supported by multiple participant quotes, and thematic saturation was achieved after coding 6 interviews, as no new codes emerged in subsequent interviews (Table 2).

XSL•FO
RenderX

Table 2. Qualitative excerpts about the facilitators, barriers, and areas for improvement of the HealthScore from participants in the intervention group.

Major theme and subtheme	Representative quote
Facilitators	
Relationship with coach	• "[The health coach] and I talked a lot about my goals each week: what went well, what didn't go as well. And she would help me reflect on how I was doing which was really helpful. She helped me move forward." (F[a], 50-64)
Accountability via weekly calls	• "It's hard to feel motivated when you are sick, so having someone to talk to every week was really beneficial." (F, 50-64)
Increased physical function	• "It really helped me increase it. It was an inspiration to move, to get to the gym, to exercise. It definitely helped." (F, 65+)
Improved self-efficacy or QoL[b]	• "The more exercise you get, the more positive your attitude is, the more positive your attitude is the more chance you have to live longer." (F, 50-64) • "The cancer might get me, but at least I will die healthy!" (F, 65+)
Barriers	
Ineffective survey tool	• "Questions are always the same on the survey, which might not be great, and they are vague. 'In the last 7 days..' was hard to remember. Run 3 miles question is never going to happen." (M[c], 65+)
Lack of caregiver involvement	• "For them to learn to meet you where you are and to understand that more. Having caregivers learn how to release their expectations of what you need to do and not to push their own desires onto your experience." (F, 65+)
HealthScore metrics are not comprehensive	• "My symptoms were not addressed by the program- I talked to Bri (Health Coach) about my symptoms." (F, 50-64)
More attention on topics outside of physical activity	• "Diet- especially for diet with a compromised immune system. Being coached on nutrition- what is/isn't the best foods to eat, what helps your system regain immunity and strength." (F, 65+)
Areas for improvement	
Survey improvement	• "Some of the questions were redundant. The redundancy made me question my answers. Some of those questions could be removed." (F, 50-64) • "Let us tell you how we're feeling and what would help us, especially if we don't have a counselor. More of a mental component to the program might be helpful for some people who are struggling more with their cancer." (F, 50-64)
Caregiver support	• ".. husband could have used some mental support. Some caregivers might need physical and mental support. They will have different needs. Meet the caregiver where they are at and be able to answer questions for the caregiver because they are going to have a lot of questions..." (F, 50-64)
HealthScore metrics	• "Disappointed that I used the bike a lot and it didn't pick up a lot of steps." (M, 65+)

[a]F: female.

[b]QoL: quality of life.

[c]M: Male.

Facilitators

Patients reported that maintaining a strong positive relationship with the coaches and the weekly coaching calls were key facilitators that enabled them to establish accountability. Many patients expressed their appreciation for the encouragement and empathy they received from the coaches, which facilitated the patient-coach connections and openness. The rapport established enabled the coaches to guide participants to achieve success by reframing setbacks, shifting mindsets, and achieving SMART goals. Personal connection and trust enhanced patients' engagement more during the weekly check-ins, fostering accountability and contributing to overall success.

Patients voiced that Fitbit tracking and goal setting facilitated consistency of participation and adherence to health goals. Patients reported increased levels of physical activity, often making progress toward pretreatment levels and improved QOL. Many patients expressed eagerness to continue goal setting, exercise, and FitBit tracking beyond the conclusion of the study

XSL•FO
RenderX

period. This sustained interest reflects an improvement in their confidence to manage their own health and well-being without their coaches, indicating an improvement in self-efficacy.

Additionally, patients reported improved communication with their care teams, which enabled them to be more proactive in managing their symptoms. They felt more confident in expressing their concerns clearly and seeking help early.

Barriers

Some patients expressed that the survey tool, which primarily focused on symptoms and physical function, was ineffective in capturing the progress they observed through their Fitbits. Many participants reported frustration at having to complete weekly surveys, which they felt were tedious and repetitive. Some patients felt that specific survey questions provided led to the setting of unrealistic goals and feelings of discouragement. Completing the survey at the end of each week also posed a challenge, as some struggled to recall their weekly progress accurately. Since the surveys directly impacted the HealthScore metrics, patients felt that regularly displaying the HealthScore (physical function measure) during the study could have negatively affected their motivation. Additionally, participants expressed a preference for coaching with a focus on their personal needs, such as mental health, nutrition, and weight management.

One idea that some patients noted was that involving a family or friend caregiver in the coaching sessions could be an additional component to the program, with additional implications. Some noted that the program could be adapted to allow for shared goal-setting across patients and caregivers, which may benefit both members. Others expressed concern about placing an additional burden on their caregivers.

Areas for Improvement

Patients identified scope for improvement in the weekly surveys, HealthScore metrics, caregiver involvement, and customized coaching topics. They felt that the surveys needed to be revisited and adapted to better capture the progress they were making. They suggested the incorporation of the Fitbit data (step count, heart rate, and sleeping pattern) into the HealthScore Metric to provide a more accurate and comprehensive reflection of their progress.

Most patients perceived that there is a significant potential for this program to improve caregiver support by offering tailored resources, information, and services that address their caregiver responsibilities as well as their own physical and emotional well-being. Some patients believe that by providing the resources, tools, and services that alleviate caregiver burden, the program has the potential to improve caregiver HRQOL and ultimately the support they can provide to patients. Additionally, participants also perceived that caregiver support can be improved through the program by supporting the physical and mental well-being of their caregivers.

Some patients also identified ideas for personalizing the program by including additional health coaching topics, such as nutrition, sleep, weight management, and mental health. The program could also be improved by enhancing the referral process to other community-based resources and services.

Integrated Mixed Methods Findings

The integration of quantitative physical activity data with qualitative interview themes revealed explanatory insights into the findings of the intervention group. Quantitatively, the findings demonstrated that intervention participants achieved significantly higher average daily step counts and daily walking distances compared to the control group, with steeper increases over time. Our qualitative analysis explained these improvements through specific program facilitators (eg, weekly coaching calls or growth in self-efficacy). However, this analysis also revealed a disparity in quantitative outcomes and participant experiences. While physical activity metrics improved overall, participants expressed frustration with the weekly PRO surveys, as they felt they were not aligned with their Fitbit-captured progress. Moreover, participants desired additional coaching in areas outside physical activity (eg, nutrition and sleep), topics not captured quantitatively but perceived as integral to their overall recovery and well-being.

Discussion

Principal Findings

We conducted a secondary analysis of a randomized controlled pilot study of cancer survivors who participated in a novel telehealth-based coaching program called HealthScore [24]. We found that participants in the intervention group had significantly higher physical activity than participants in the control group. Average daily step counts and average daily walking distance in the intervention participants were both significantly more than those of control participants and increased over the course of the 6-month intervention more than twice that of control participants. The integration of quantitative and qualitative findings revealed that these improvements were driven by enhanced accountability through weekly coaching relationships and increased self-efficacy, although the measurement tools themselves presented barriers that may have affected some participants' motivation. Our findings underscore the potential effectiveness of the HealthScore coaching program in enhancing physical activity levels among cancer survivors.

The combination of personalized coaching and physiologically based PGHD positions HealthScore as a robust approach to overcoming barriers associated with physical inactivity among cancer survivors. Our findings demonstrate patterns resonant with those in similarly designed health coaching interventions tested in cancer survivors [19,20], align with existing literature on the benefits of structured health coaching, and affirm that tailored interventions can significantly impact exercise adherence [36,37]. However, the sustained engagement via weekly sessions allowed participants to set realistic SMART goals and receive ongoing support, fostering an environment conducive to behavior change, which was praised by participants as an effective design for the coaching program. Moreover, the inclusion of motivational interviewing techniques proved beneficial in empowering participants, fostering autonomy, and addressing intrinsic barriers, such as anxiety and depression.

XSL•FO
RenderX

Qualitatively, participants indicated a growing sense of accountability and motivation throughout the program duration. Many expressed that the regular weekly check-ins with their health coaches provided essential encouragement and instilled a sense of connection, which is critical for cancer survivors who often experience feelings of isolation [38,39]. Participants recognized improvements in their physical capabilities and reported enhanced emotional well-being, highlighting the multifaceted benefits of the HealthScore program.

Key barriers identified by participants included aspects of the weekly monitoring PRO surveys, which were focused on symptoms and physical function. Participants felt that the surveys did not accurately capture their progress, which led to frustration with the tedious and repetitive nature of these assessments. This finding is congruent with prior literature showing survey fatigue as a common pain point in cancer health services research [1,40]. Additionally, challenges in recalling weekly progress and the potential negative impact of the displayed HealthScore on motivation were noted. Prior research reported similar findings that using self-monitoring tools can negatively impact participants' motivation [41,42]. Potential solutions include motivational messages accompanied by positive statistics showing improvement over time or comparing the participant's activity to the overall performance of participants.

The mixed methods integration illuminated meaningful insights. For example, participants attributed their increased activity to enhanced accountability through weekly coaching calls and the relationships they had built. Additionally, participants experienced growth in self-efficacy as they internalized goal-setting practices and saw tangible progress via their Fitbit devices. However, identified barriers, such as survey fatigue, may help explain the variability in quantitative outcomes. The participants' frustration with repetitive surveys that failed to capture their perceived progress suggests that while objective metrics improved overall, the measurement tools themselves may have dampened motivation for some. Therefore, it is important to align measurement approaches that optimize participant engagement throughout a sustained trial, such as HealthScore. Our convergence of quantitative improvements with qualitative insights strengthens the findings and can inform recommendations for refinement of similar programs.

Future iterations of HealthScore, or other similar programs, will integrate feedback from participants to further streamline the program for improved user experience. As identified by participants, the number and length of monitoring surveys should be reduced to avoid survey fatigue by considering biweekly or monthly surveys. In addition, incorporating Fitbit data into HealthScore metrics and providing personalized, evidence-based guidelines around nutrition, sleep, and mental health could improve the intervention. Emerging technologies, including artificial intelligence, have the power to collect, analyze, and synthesize PGHD and to provide real-time information from trusted sources.

Comparison to Prior Work

Prior research underscores the importance of high-quality, standardized coaching programs, as previous studies have shown that intervention intensity and participant engagement are critical for effective behavior change [43-45]. These limitations can be mitigated by implementing user-centered design principles through design thinking methods, expanding and incorporating objective measures alongside self-reports, and ensuring consistent training for all health coaches, all of which could enhance the effectiveness of future health interventions for cancer survivors.

This study contributes to prior literature by exploring the effectiveness of a telehealth-based coaching program, HealthScore, specifically designed to enhance physical activity among cancer survivors. It integrates PGHD from Fitbit devices for objective assessments, uses a mixed methods approach to combine quantitative and qualitative insights, and emphasizes the importance of structured, personalized coaching sessions aligned with SMART goals. The rigorous training of health coaches in motivational interviewing further distinguishes this study, highlighting the importance of interpersonal support in behavioral change, thereby offering a comprehensive and modern perspective on improving health outcomes in cancer survivors.

Limitations and Future Directions

Some limitations are worth noting. First, the sample size, while adequate for preliminary analysis, may not fully represent the diversity of the patients with cancer and cancer survivors. Future studies should consider larger, more diverse cohorts to better understand the generalizability of HealthScore's findings across the cancer continuum. Second, longer follow-up periods could help ascertain the sustainability of physical activity levels beyond the intervention, as well as the long-term effects on overall health outcomes. Third, there was a significant attrition imbalance between the control and intervention groups. While this is likely attributed to the fact that the intervention group entailed more time demand, there may also have been reluctance to the increased accountability of having a health coach. Importantly, the imbalance may have influenced both the qualitative and quantitative findings. Fourth, there are some limitations inherent in secondary data analysis. The original trial was not powered to detect the specific outcomes examined here, and the relatively small, demographically homogeneous sample, coupled with reliance on volunteer coaches, constrains both generalizability and scalability. Fifth, we acknowledge the potential for selection bias given that participation relied on self-referral and that a substantial number of individuals were lost between screening and analysis. As a result, those who remained in the study may represent a more motivated or otherwise distinct subset of the broader population, which may limit the generalizability of our findings. Future research should examine the long-term effects and applicability across diverse cancer populations and equally sized groups.

Conclusion

The HealthScore telehealth coaching program shows promise in enhancing physical activity and QOL for cancer survivors. Personalized coaching and PGHD foster sustained engagement in exercise, overcoming common barriers. Participants reported increased motivation and adherence to regimens, highlighting the importance of tailored support for psychological and physical

XSL•FO
RenderX

challenges. Qualitative feedback indicated that the program's holistic approach improved participants' confidence in managing their health after treatment. Integrating telehealth coaching into standard care for cancer survivors might be possible once these findings are replicated in larger and more diverse cohorts to promote active lifestyles and better health outcomes.

Acknowledgments

Authors (SK, JG, KM, and ZZ) were supported by the National Center for Advancing Translational Sciences of the National Institutes of Health (R21TR002088 and RC2TR004380). The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, nor does it mention of department or agency names imply endorsement by the US government. All authors had access to the data and a role in writing the manuscript. No generative artificial intelligence (AI) was used in the development of this study or manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to patient privacy issues, but are available from the corresponding author on reasonable request.

Funding

KRT received funding from the National Cancer Institute (T32CA116339). This study was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health (NIH) under award number RC2TR004380.

Authors' Contributions

Conceptualization: SK, EK, CB, WW
Data curation: JG, KM, ZZ, CB
Formal analysis: JG, KM, ZZ
Funding acquisition: WW
Investigation: SK, EK, CB, WW
Methodology: SK, EK, WW
Project administration: CB
Supervision: SK, EK, CB, WW
Validation: JG, KM, ZZ
Writing–original draft: SK, EK, JG, KM, ZZ, CB, WW
Writing–review & editing: SK, EK, JG, KM, ZZ, CB, WW

Conflicts of Interest

None declared

Multimedia Appendix 1
TIDieR Checklist.
[DOCX File , 37 KB - cancer_v12i1e78968_app1.docx]

Multimedia Appendix 2
CONSORT extension Pilot and Feasibility Trials Checklist.
[DOC File , 240 KB - cancer_v12i1e78968_app2.doc]

Multimedia Appendix 3
Participant exit interview questions.
[DOCX File , 26 KB - cancer_v12i1e78968_app3.docx]

Multimedia Appendix 4
GRAMMS Checklist.
[DOCX File , 46 KB - cancer_v12i1e78968_app4.docx]

Multimedia Appendix 5
Intervention Group Interviewee Characteristics.
[DOCX File , 20 KB - cancer_v12i1e78968_app5.docx]

XSL•FO
RenderX

References

1. Frikkel J, Götte M, Beckmann M, Kasper S, Hense J, Teufel M, et al. Fatigue, barriers to physical activity and predictors for motivation to exercise in advanced cancer patients. BMC Palliat Care 2020;19(1):43 [FREE Full text] [doi: 10.1186/s12904-020-00542-z] [Medline: 32234027]
2. Yang L, Morielli AR, Heer E, Kirkham AA, Cheung WY, Usmani N, et al. Effects of exercise on cancer treatment efficacy: a systematic review of preclinical and clinical studies. Cancer Res 2021;81(19):4889-4895 [FREE Full text] [doi: 10.1158/0008-5472.CAN-21-1258] [Medline: 34215623]
3. Takemura N, Chan SL, Smith R, Cheung DST, Lin C. The effects of physical activity on overall survival among advanced cancer patients: a systematic review and meta-analysis. BMC Cancer 2021;21(1):242 [FREE Full text] [doi: 10.1186/s12885-021-07988-1] [Medline: 33678180]
4. Getie A, Ayalneh M, Bimerew M. Global prevalence and determinant factors of pain, depression, and anxiety among cancer patients: an umbrella review of systematic reviews and meta-analyses. BMC Psychiatry 2025;25(1):156 [FREE Full text] [doi: 10.1186/s12888-025-06599-5] [Medline: 39972435]
5. Wong JN, McAuley E, Trinh L. Physical activity programming and counseling preferences among cancer survivors: a systematic review. Int J Behav Nutr Phys Act 2018;15(1):48 [FREE Full text] [doi: 10.1186/s12966-018-0680-6] [Medline: 29879993]
6. Qiu L, Ye M, Tong Y, Jin Y. Promoting physical activity among cancer survivors: an umbrella review of systematic reviews. Support Care Cancer 2023;31(5):301 [FREE Full text] [doi: 10.1007/s00520-023-07760-0] [Medline: 37097500]
7. Berzins NJ, Mackenzie M, Galantino ML, Pickles N, Hebbel S, Leonard T, et al. Preliminary effectiveness study of a community-based wellness coaching for cancer survivors program. Am J Lifestyle Med 2024;18(4):465-474. [doi: 10.1177/15598276221076040] [Medline: 39262887]
8. Gilchrist SC, Howard VJ, Akinyemiju T, Judd SE, Cushman M, Hooker SP, et al. Association of sedentary behavior with cancer mortality in middle-aged and older US adults. JAMA Oncol 2020;6(8):1210-1217 [FREE Full text] [doi: 10.1001/jamaoncol.2020.2045] [Medline: 32556069]
9. Cormie P, Zopf EM, Zhang X, Schmitz KH. The impact of exercise on cancer mortality, recurrence, and treatment-related adverse effects. Epidemiol Rev 2017;39(1):71-92. [doi: 10.1093/epirev/mxx007] [Medline: 28453622]
10. Batalik L, Chamradova K, Winnige P, Dosbaba F, Batalikova K, Vlazna D, et al. Effect of exercise-based cancer rehabilitation via telehealth: a systematic review and meta-analysis. BMC Cancer 2024;24(1):600 [FREE Full text] [doi: 10.1186/s12885-024-12348-w] [Medline: 38760805]
11. Li J, Liu Y, Jiang J, Peng X, Hu X. Effect of telehealth interventions on quality of life in cancer survivors: a systematic review and meta-analysis of randomized controlled trials. Int J Nurs Stud 2021;122:103970. [doi: 10.1016/j.ijnurstu.2021.103970] [Medline: 34303269]
12. Su JJ, Winnige P, Chamradova K, Dosbaba F, Batalikova K, Lin R, et al. Feasibility, safety, and adherence of home-based exercise interventions in people diagnosed with cancer: a systematic review. J Cancer Surviv 2025:1-22. [doi: 10.1007/s11764-025-01778-5] [Medline: 40119984]
13. Campbell KL, Winters-Stone KM, Wiskemann J, May AM, Schwartz AL, Courneya KS, et al. Exercise guidelines for cancer survivors: consensus statement from International Multidisciplinary Roundtable. Med Sci Sports Exerc 2019;51(11):2375-2390 [FREE Full text] [doi: 10.1249/MSS.0000000000002116] [Medline: 31626055]
14. Brown JC, Ma C, Shi Q, Saltz LB, Shields AF, Meyerhardt JA. The association of physical activity with survival in colon cancer versus a matched general population: data from Cancer and Leukemia Group B 89803 and 80702 (Alliance). Cancer 2025;131(5):e35727. [doi: 10.1002/cncr.35727] [Medline: 39989023]
15. Pinto BM, Stein K, Dunsiger S. Peers promoting physical activity among breast cancer survivors: a randomized controlled trial. Health Psychol 2015;34(5):463-472 [FREE Full text] [doi: 10.1037/hea0000120] [Medline: 25110844]
16. Barakat S, Boehmer K, Abdelrahim M, Ahn S, Al-Khateeb AA, Villalobos, et al. Does health coaching grow capacity in Cancer survivors? A systematic review. Popul Health Manag 2018;21(1):63-81. [doi: 10.1089/pop.2017.0040] [Medline: 28636526]
17. Barrett S, Begg S, O'Halloran P, Kingsley M. A physical activity coaching intervention can improve and maintain physical activity and health-related outcomes in adult ambulatory hospital patients: the Healthy4U-2 randomised controlled trial. Int J Behav Nutr Phys Act 2020;17(1):156 [FREE Full text] [doi: 10.1186/s12966-020-01063-x] [Medline: 33256753]
18. Nerek A, Wesołowska-Górniak K, Czarkowska-Pączek B. Enhancing feedback by health coaching: the effectiveness of mixed methods approach to long-term physical activity changes in nurses. An intervention study. BMC Nurs 2024;23(1):196 [FREE Full text] [doi: 10.1186/s12912-024-01815-1] [Medline: 38519943]
19. Yun YH, Kim YA, Lee MK, Sim JA, Nam B, Kim S, et al. A randomized controlled trial of physical activity, dietary habit, and distress management with the Leadership and Coaching for Health (LEACH) program for disease-free Cancer survivors. BMC Cancer 2017;17(1):298 [FREE Full text] [doi: 10.1186/s12885-017-3290-9] [Medline: 28464804]
20. Weemaes ATR, Beelen M, Weijenberg MP, van Kuijk SMJ, Lenssen AF. Effects of remote coaching following supervised exercise oncology rehabilitation on physical activity levels, physical fitness, and patient-reported outcomes: a randomised controlled trial. Int J Behav Nutr Phys Act 2024;21(1):8 [FREE Full text] [doi: 10.1186/s12966-024-01561-2] [Medline: 38273361]

XSL•FO
RenderX

21. Dennett A, Harding KE, Reimert J, Morris R, Parente P, Taylor NF. Telerehabilitation's safety, feasibility, and exercise uptake in cancer survivors: process evaluation. JMIR Cancer 2021;7(4):e33130 [FREE Full text] [doi: 10.2196/33130] [Medline: 34854817]
22. Bollyky JB, Bravata D, Yang J, Williamson M, Schneider J. Remote lifestyle coaching plus a connected glucose meter with certified diabetes educator support improves glucose and weight loss for people with type 2 diabetes. J Diabetes Res 2018;2018:3961730 [FREE Full text] [doi: 10.1155/2018/3961730] [Medline: 29888288]
23. Nair B, Otaki F. Promoting university students' mental health: a systematic literature review introducing the 4M-Model of individual-level interventions. Front Public Health 2021;9:699030 [FREE Full text] [doi: 10.3389/fpubh.2021.699030] [Medline: 34249852]
24. Wood WA, Bailey C, Castrogivanni B, Mehedint D, Bryant AL, Lavin K, et al. Piloting HealthScore: feasibility and acceptability of a clinically integrated health coaching program for people living with cancer. Cancer Med 2023;12(7):8804-8814 [FREE Full text] [doi: 10.1002/cam4.5625] [Medline: 36647557]
25. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Los Angeles: SAGE Publications; 2022:320.
26. Survivorship terms. National Cancer Institute (NCI). URL: https://cancercontrol.cancer.gov/ocs/definitions [accessed 2025-03-27]
27. White ND, Bautista V, Lenz T, Cosimano A. Using the SMART-EST goals in lifestyle medicine prescription. Am J Lifestyle Med 2020;14(3):271-273 [FREE Full text] [doi: 10.1177/1559827620905775] [Medline: 32477026]
28. Deci EL, Ryan LRM. Handbook of Theories of Social Psychology. In: Van Lange PAM, Kruglanski AW, Higgins ET, editors. Self-determination theory. Thousand Oaks, CA: Sage Publications Ltd; 2012:416-436.
29. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
30. Simmons LA, Wolever RQ. Integrative health coaching and motivational interviewing: synergistic approaches to behavior change in healthcare. Glob Adv Health Med 2013;2(4):28-35 [FREE Full text] [doi: 10.7453/gahmj.2013.037] [Medline: 24416683]
31. Jensen RE, Potosky AL, Reeve BB, Hahn E, Cella D, Fries J, et al. Validation of the PROMIS physical function measures in a diverse US population-based cohort of cancer patients. Qual Life Res 2015;24(10):2333-2344 [FREE Full text] [doi: 10.1007/s11136-015-0992-9] [Medline: 25935353]
32. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
33. Chapman AL, Hadfield M, Chapman CJ. Qualitative research in healthcare: an introduction to grounded theory using thematic analysis. J R Coll Physicians Edinb 2015;45(3):201-205 [FREE Full text] [doi: 10.4997/JRCPE.2015.305] [Medline: 26517098]
34. Dedoose. 2025. URL: https://www.dedoose.com./ [accessed 2021-03-21]
35. Attride-Stirling J. Thematic networks: an analytic tool for qualitative research. Qual Res 2001;1(3):385-405. [doi: 10.1177/146879410100100307]
36. Bullard T, Ji M, An R, Trinh L, Mackenzie M, Mullen SP. A systematic review and meta-analysis of adherence to physical activity interventions among three chronic conditions: cancer, cardiovascular disease, and diabetes. BMC Public Health 2019;19(1):636 [FREE Full text] [doi: 10.1186/s12889-019-6877-z] [Medline: 31126260]
37. Yang Z, Huang H, Xu L, Gao Y, Zhang C, Wang A. From need to action: decoding the role of motivation in adherence to home-based cardiac rehabilitation exercise in chronic heart failure patients. BMC Nurs 2024;23(1):913 [FREE Full text] [doi: 10.1186/s12912-024-02617-1] [Medline: 39695568]
38. Zhao J, Reese JB, Han X, Yabroff KR. Loneliness and mortality risk among cancer survivors in the United States: a retrospective, longitudinal study. J Natl Compr Canc Netw 2024;22(4):244-248. [doi: 10.6004/jnccn.2023.7114] [Medline: 38663443]
39. Kittel JA, Seplaki CL, van Wijngaarden E, Richman J, Magnuson A, Conwell Y. Fatigue, impaired physical function and mental health in Cancer survivors: the role of social isolation. Support Care Cancer 2024;33(1):16. [doi: 10.1007/s00520-024-09075-0] [Medline: 39661200]
40. Xu J, Li Q, Gao Z, Ji P, Ji Q, Song M, et al. Impact of cancer-related fatigue on quality of life in patients with cancer: multiple mediating roles of psychological coherence and stigma. BMC Cancer 2025;25(1):64 [FREE Full text] [doi: 10.1186/s12885-025-13468-7] [Medline: 39794768]
41. Zhang S, Rehman S, Zhao Y, Rehman E, Yaqoob B. Exploring the interplay of academic stress, motivation, emotional intelligence, and mindfulness in higher education: a longitudinal cross-lagged panel model approach. BMC Psychol 2024;12(1):732 [FREE Full text] [doi: 10.1186/s40359-024-02284-6] [Medline: 39696514]
42. Michaelsen MM, Esch T. Understanding health behavior change by motivation and reward mechanisms: a review of the literature. Front Behav Neurosci 2023;17:1151918 [FREE Full text] [doi: 10.3389/fnbeh.2023.1151918] [Medline: 37405131]
43. Matthews L, Simpson SA. Evaluation of Behavior Change Interventions. In: Hagger MS, Cameron LD, Hamilton K, Hankonen N, Lintunen T, editors. The Handbook of Behavior Change. Cambridge: Cambridge University Press; 2020:318-332.

XSL•FO
RenderX

44. Passarelli AM, Moore S, Van Oosten EB. How leaders and their coaches describe outcomes of coaching for intentional change. Consult Psychol J 2022;74(4):329-346. [doi: 10.1037/cpb0000240]
45. Michie S, West R, Sheals K, Godinho CA. Evaluating the effectiveness of behavior change techniques in health-related behavior: a scoping review of methods used. Transl Behav Med 2018;8(2):212-224 [FREE Full text] [doi: 10.1093/tbm/ibx019] [Medline: 29381786]

Abbreviations

CHAI: Connected Health for Applications and Interventions
CONSORT: Consolidated Standards of Reporting Trials
COREQ: Consolidated Criteria for Reporting Qualitative Research
GRAMMS: Good Reporting of a Mixed Methods Study
HRQOL: health-related quality of life
IRB: Institutional Review Board
NCI: National Cancer Institute
PGHD: patient-generated health data
PRO: patient-reported outcome
PROMIS: Patient-Reported Outcomes Measurement Information System
QOL: quality of life
REDCap: Research Electronic Data Capture
SMART: Specific, Measurable, Achievable, Relevant, and Time-Bound
TIDieR: Template for Intervention Description and Replication
UNC: University of North Carolina

Edited by N Cahill; submitted 12.Jun.2025; peer-reviewed by B Lai, A Dennett, L Batalik, N Lukkahatai; comments to author 11.Sep.2025; revised version received 13.Nov.2025; accepted 25.Nov.2025; published 15.Jan.2026.

Please cite as:
Khairat S, Kent EE, Geracitano J, Mendis K, Zhou Z, Bailey C, Wood WA
Efficacy of Telehealth-Based Coaching to Improve Physical Activity and Overall Experience for Cancer Survivors: Secondary, Mixed Methods Analysis of a Randomized Controlled Trial
JMIR Cancer 2026;12:e78968
URL: https://cancer.jmir.org/2026/1/e78968
doi:10.2196/78968
PMID:

©Saif Khairat, Erin E Kent, John Geracitano, Kaushalya Mendis, Zhaoqiang Zhou, Carly Bailey, William A Wood. Originally published in JMIR Cancer (https://cancer.jmir.org), 15.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Feasibility and Acceptability of a Remotely Delivered, Combined Exercise Intervention on Cognitive Function in Patients With Breast Cancer Following Chemotherapy: Randomized Controlled Trial

Linda Trinh[1], PhD; Natalie Cuda[1], MSc; Catherine M Sabiston[1], PhD; Ellen Warner[2], MD, MSc; Jennifer D Ryan[3], PhD; Arthur F Kramer[4], PhD; Michelle Ha[1], BSc; Hui Xiao[1], PhD; Sarah O'Rourke[1], MSc; Edward McAuley[4,5], PhD

[1]Faculty of Kinesiology and Physical Education, University of Toronto, Toronto, ON, Canada

[2]Odette Cancer Centre, Sunnybrook Health Science Centre, Toronto, ON, Canada

[3]Baycrest Academy for Research and Education, Rotman Research Institute, Toronto, ON, Canada

[4]Beckman Institute, University of Illinois, Urbana, IL, United States

[5]Department of Kinesiology and Community Health, University of Illinois Urbana-Champaign, Urbana, MD, United States

Corresponding Author:
Linda Trinh, PhD
Faculty of Kinesiology and Physical Education
University of Toronto
55 Harbord Street
Toronto, ON, M5S2W6
Canada
Phone: 1 4169463624
Email: linda.trinh@utoronto.ca

Abstract

Background: Cognitive impairments, a prevalent quality-of-life concern in breast cancer (BC), are particularly pronounced in women having undergone adjuvant chemotherapy. These impairments—affecting executive function, attention, and processing speed—are often underdiagnosed, with no established treatments. Exercise is a potential intervention to mitigate cancer-related cognitive impairment (CRCI). Since virtual care delivery is feasible, remotely delivered exercise interventions for CRCI management in patients with BC may be explored.

Objective: We examined the feasibility of an 8-week remotely delivered combined exercise program (aerobic+resistance training) compared to a stretching and toning active control in postchemotherapy patients with BC.

Methods: Patients with BC who completed adjuvant chemotherapy within 48 months were recruited across Canada from February to July 2023. The combined exercise group engaged in unsupervised aerobic exercise (30 minutes, thrice per week), supervised group-based resistance training (30 minutes, twice per week via Zoom [Zoom Communications, Inc]), and one recorded class weekly, supplemented with 4 biweekly behavioral counseling sessions. The active control group participated in low-intensity balance and flexibility classes (30 minutes, twice per week live, once per week recorded). Feasibility was assessed via enrollment, adherence, attrition, measurement completion, adverse events, and participant satisfaction; cognitive function, using the National Institutes of Health Toolbox Cognition Battery Remote Administration (V2) at baseline and post intervention.

Results: Twenty-one participants (mean age 51.6, SD 7.2 years; 11.8, SD 12.9 months since treatment) were randomized to the combined exercise (n=10) or the active control (n=11) group. Final analyses included 18 participants (mean age 51.9, SD 7.4 years; mean months since treatment 12.6, SD 13.5) with 9 participants in each group (51.2% enrollment rate; 14.3% attrition; no adverse events). Measurement completion rates were 85.7% and participants reported high satisfaction with the intervention, indicating minimal burden. Adherence rates for exercise classes were 70.8% in both groups. Adherence to behavioral counseling sessions was 77.5% (combined exercise group only). There were no significant differences in objectively measured cognitive function, but small-to-medium effect size improvements were observed in objectively measured episodic memory (mean difference 5.33, 95% CI –12.5 to 23.2; ηp^2=0.03), working memory (mean difference 8.17, 95% CI –4.2 to 20.6; ηp^2=0.12), executive function updating (mean difference –394.35, 95% CI –1035.67 to 246.96 ms; ηp^2=0.07), and immediate memory and verbal learning (mean difference +3.22, 95% CI –2.0 to 8.5; ηp^2=0.12), trending toward the multicomponent exercise group vs the active

XSL•FO
RenderX

control group. In contrast, small-to-medium effect size improvements were observed in the Oral Reading Recognition Test (mean difference –9.65, 95% CI –22.9 to 3.5; *P*=.14; ηp^2=0.14) and Picture Vocabulary Test (mean difference –2.48, 95% CI –5.2 to 0.3, *P*=.07; ηp^2=0.20), trending toward the active control group.

Conclusions: A remotely delivered combined exercise intervention is feasible and may improve CRCI in patients with BC. Larger randomized controlled trials are warranted to confirm its efficacy in enhancing cognitive function and quality of life in this population.

Trial Registration: ClinicalTrials.gov NCT05704855; https://clinicaltrials.gov/study/NCT05704855

(JMIR Cancer 2026;12:e73393) doi:10.2196/73393

KEYWORDS

cancer-related cognitive impairment; breast cancer; quality of life; exercise, feasibility study

Introduction

An increasing number of patients with breast cancer (BC) are experiencing chronic symptoms due to rising cancer incidence and improving survival rates [1]. Treatments, such as chemotherapy, radiation therapy, and hormone therapy, while effective at reducing mortality, are associated with many adverse side effects, including cognitive decline known as cancer-related cognitive impairment (CRCI) [2]. CRCI has been reported in up to 85% of patients with BC, lasting for months, years, and not uncommonly indefinitely following treatment completion [3]. CRCI manifests as problems with attention, processing speed, memory, and executive function and is most commonly reported following chemotherapy [4-6]. Chemotherapeutic drugs may have neurotoxic effects, leading to structural and functional changes in the brain [4,7-9]. Other chemotherapy-related side effects, such as fatigue, anxiety, depression, stress, and sleep dysfunction, may further degrade cognition in patients with BC [10,11]. These cognitive problems can significantly impair work performance, social relationships, and daily functioning, ultimately reducing quality of life (QoL) [12,13]. Despite the prevalence and significance of this problem, no established treatment has yet been identified to mitigate the adverse effects of CRCI on patients with BC [14].

Although the majority of evidence for CRCI in patients with BC is attributed to chemotherapy, one confounding aspect is the impact of aging and changes in menopausal status that result from chemotherapy-induced menopause and lowered levels of estradiol [15-17]. Case studies in patients with BC show that CRCI may vary in those who received the same course of chemotherapy, potentially due to differences in menopausal status [18]. Schilder et al [19] demonstrated that after one year of adjuvant therapy, tamoxifen use was associated with statistically significant lower functioning in verbal memory and executive functioning, whereas exemestane use was not associated with statistically significant lower cognitive functioning in postmenopausal patients with BC. Klemp et al [20] conducted a longitudinal study examining both subjective and objective measures of cognitive function and QoL in pre- and perimenopausal patients with BC receiving chemotherapy. Assessments were conducted prior to chemotherapy, after cycle 3, within 2-3 weeks of completing adjuvant chemotherapy, and more than 8+ years later. Neither age nor estradiol levels were associated with cognitive complaints. Vega et al [21] compared a group of primarily postmenopausal women with persistent CRCI to 2 groups of postmenopausal women: women who endorse menopause - associated subjective cognitive decline (maSCD+) and women who do not (maSCD−) to explore the potential role of menopause in CRCI. Women with persistent CRCI reported more severe subjective symptoms of cognitive decline, along with measurable differences in objective performance, compared to women who experienced natural menopause. In addition, women with persistent CRCI reported greater menopausal symptoms compared with the maSCD- group, but not the maSCD+ group. These findings were unrelated to menopausal status prior to chemotherapy or current endocrine therapy use. Although menopausal symptoms may contribute to some CRCI experienced by patients with BC, they do not fully account for CRCI. Future research is needed to determine the specific predictors and causal mechanisms for CRCI.

Implementing combined (aerobic+resistance training) exercise interventions could be a promising strategy to improve cognitive outcomes and reduce the impact of CRCI. Studies with older adults have demonstrated that combined exercise improves memory more than aerobic exercise alone by producing changes in brain structure, brain function, and brain connectivity [22,23]. Recent reviews indicate that exercise interventions, such as aerobic and combined aerobic-resistance training, may help improve cognitive function in patients with BC [3,24]. Across interventions, programs delivered 2-5 times per week for 10-60 minutes, or aerobic training combined with whole-body resistance, most consistently improved short-term self-reported cognition [3,24]. Another scoping review of 97 studies on physical activity (PA) and cognition among cancer survivors, primarily patients with BC, found that 32% of studies reported positive associations, 66% had inconclusive results, and only 2% showed negative associations, with cognition assessed using both objective and subjective measures [25]. While the review encompassed various study designs, including observational and interventional research, most interventions were delivered after, rather than during, chemotherapy. Complementing these findings, the recent ACTIVATE (Aerobic exercise and CogniTIVe functioning in women with breAsT cancEr) trial specifically examined the effects of aerobic exercise during chemotherapy, demonstrating significant improvements in self-reported cognitive function and QoL in women with BC, despite no significant effects on objective cognitive outcomes [26].

XSL•FO
RenderX

Despite growing interest in this area, findings remain preliminary with limited evidence derived from high-quality, adequately powered randomized controlled trials (RCTs) in patients with BC [3,24]. For example, a 3-arm RCT compared 16 weeks of concurrent aerobic-resistance exercise, continuous moderate-intensity aerobic exercise, and high-intensity interval training to usual care in 206 patients with BC following chemotherapy [27]. Immediately post intervention, the aerobic-resistance training group had no changes in self-reported cognitive function [27]. However, lower levels of cognitive cancer-related fatigue were noted at the 2-year follow-up in this group [28]. Nevertheless, the lack of data on the specific type and frequency of exercise during the follow-up period makes it challenging to attribute these improvements solely to the initial aerobic-resistance exercise intervention.

Meanwhile, many trials are limited to subjective measures of cognitive function and do not prioritize objective cognitive function as a primary outcome. While self-reported CRCI is correlated with mood and fatigue, the gold standard for measuring cognitive function, as developed by the International Cognition and Cancer Task Force (ICCTF), is objective tests of processing speed, memory, and executive function [6]. In practice, objective batteries are often relegated to secondary end points, and when included, they tend to show smaller or inconsistent effects compared with perceived cognition. For example, the recent ACTIVATE trial found improvements in self-reported cognitive function and QoL without parallel benefits on objective tests [26]. Similarly, Galiano-Castillo et al [29] conducted an RCT to test the effects of an 8-week remotely delivered combined exercise program on cognitive function for patients with BC. Following this program, improvements were found in short-term memory, measured by the recall of auditory consonant trigrams, but not mental flexibility, measured by Trail Making Test performance. While this study highlights the potential benefits of a remotely delivered combined exercise program on cognitive function, its primary focus was on physical functional capacity, with cognitive function evaluated as a secondary outcome [29]. Koevoets et al [30] examined the effects of a 6-month supervised combined exercise intervention on cognitive function for patients with BC exposed to chemotherapy. Beneficial intervention effects were found for secondary outcomes of self-reported cognitive function, fatigue, QoL, and depression, but not for the primary outcome of total recall measured by the Hopkins Verbal Learning Test-Revised, or other objective measures of cognitive function measured by the Amsterdam Cognition Scan [30].

Overall, key gaps were noted in the literature regarding exercise interventions, including limited understanding of the exercise dose needed to elicit cognitive benefits, insufficient attention to behavior change support required for maintaining long-term exercise engagement and reducing long-term CRCI [3,25,31,32], and a lack of remotely delivered exercise interventions. Moreover, inconsistent inclusion of patients with BC with self-reported CRCI following chemotherapy, absence of active comparators, and considerable heterogeneity in cognitive function measures are also notable gaps in the literature. Attention-matched or active comparators (eg, stretching and toning) are needed to isolate exercise effects and minimize bias [33-35]. Active control groups (receiving intervention during or after the intervention period) have lower contamination (ie, increasing exercise) and maintains participant engagement regardless of group assignment. This noncompliance may lead to decreased power to detect a significant intervention effect [34]. As such, the available evidence is insufficient for expert consensus to guide exercise recommendations for CRCI management and additional research is warranted.

Previous studies were conducted in person under supervised conditions, and little is known about remotely delivered combined exercise interventions. Patients with BC commonly report issues with transportation to fitness centers, weather constraints, and lack of accessibility as barriers to in person exercise [36]. Remotely delivered combined exercise programs have the potential to reduce these barriers and enhance reach and accessibility to patients with BC [29,37]. Preliminary evidence suggests that 10- to 16-week [28,38] supervised or 8-week [29] home-based combined exercise can improve self-reported cognitive function in patients with BC. Systematic reviews have also shown the effectiveness of remotely delivered exercise interventions in improving physical outcomes [39,40], with high adherence and satisfaction rates among patients with BC [41]. Therefore, there is a clear need for more rigorous neuropsychological testing to examine the effects of combined exercise on cognitive function in patients with BC, especially given the promising preliminary evidence supporting its benefits and the limited knowledge of remotely delivered interventions [28,29,38].

The primary purpose of this study was to assess the feasibility of an 8-week, remotely delivered, combined exercise intervention (aerobic+resistance training) on cognitive function in patients with BC following chemotherapy. Feasibility indicators included enrollment rates, adherence (defined as maintaining an exercise rate of ≥70%), attrition rates (<30%), adverse events, participant satisfaction, and feedback. It was hypothesized that the trial would meet the following a priori feasibility indicators, including an adherence rate 70%, an attrition rate <30% [42-44], and that no adverse events or major injuries would occur as a result of the combined exercise intervention. We also predicted that the combined exercise group would result in significantly greater improvements in objective measures of attention, executive function, and memory (secondary outcomes) compared to the active control group at post intervention (8 weeks), with sustained exercise.

Methods

Study Design

This 2-armed, parallel-design pilot RCT evaluated the effects of an 8-week, remotely delivered combined (aerobic + resistance training) exercise program compared to an active control group. Participants were randomized in a 1:1 allocation ratio. Reporting of the trial outcomes followed the CONSORT-eHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth; Multimedia Appendix 1) 2011 guidelines [45,46].

XSL•FO
RenderX

Participants

Participants were recruited from cancer care organizations, support groups, university listserves, an existing database of research individuals who participated in prior research studies in the laboratory, and social media advertisements (ie, Facebook [Meta Platforms, Inc], Instagram [Meta Platforms, Inc], and X [X Corp]) across Canada between February and July 2023. Eligibility criteria included (1) 40-65 years of age, (2) a history of stage I-III BC (nonmetastatic), (3) a history of receiving chemotherapy within 48 months prior to enrollment, (4) mild cognitive impairment as determined by the Modified Telephone Interview of Cognitive Status (TICS-M) [scores between 21 and 24 to separate individuals with mild cognitive impairment and normal cognition (>24)] [47], (5) self-reported low-activity, defined as <3 days of exercise (<20 min/d) per week in the previous 6 months [48], (6) physician clearance to participate in exercise if required, (7) no previous invasive cancer, (8) no neurological or musculoskeletal comorbidity inhibiting exercise, (9) access to a webcam and internet for videoconferencing, and (10) fluency in English. Participants who were interested in the study contacted the study team to schedule a screening call to further assess their eligibility. Additionally, participants were asked to provide their tentative availability for attending 2 live, remotely delivered classes each week.

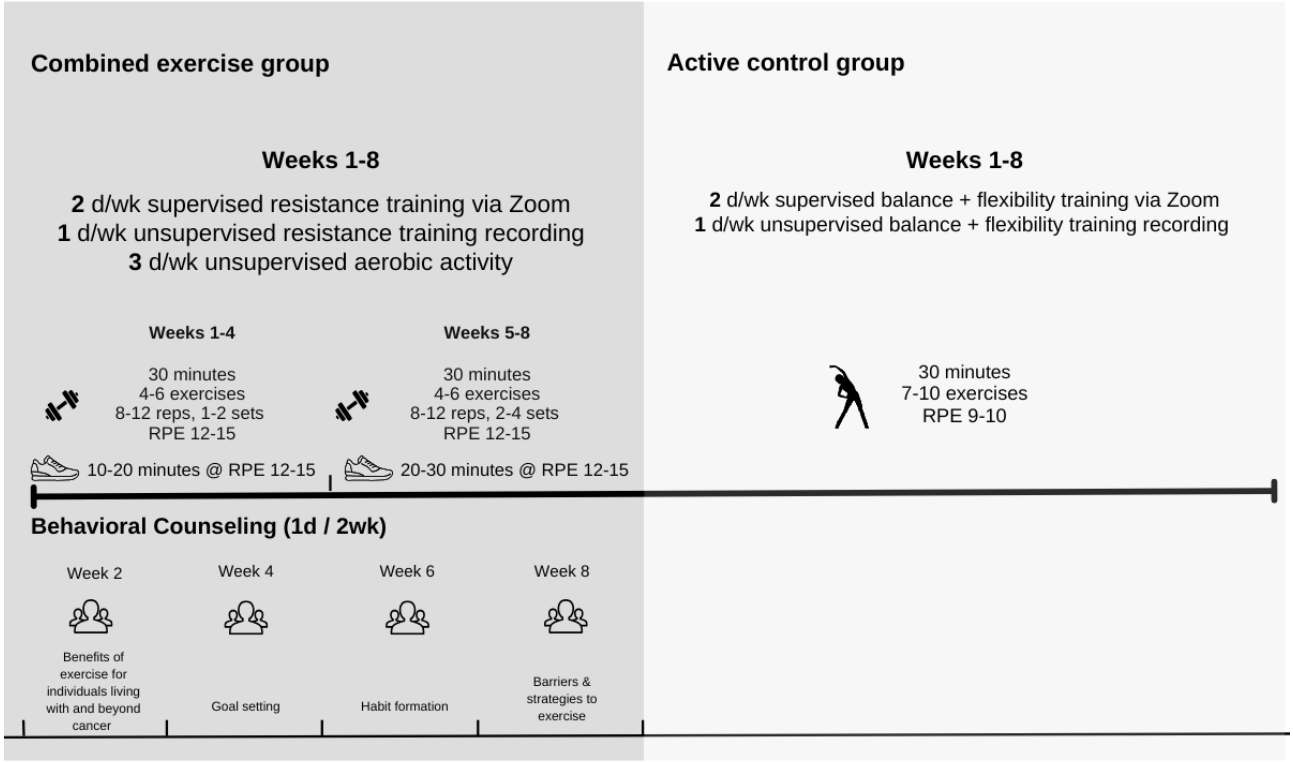
Randomization and Blinding

Eligible participants were randomized in a 1:1 ratio to either the combined exercise group or the active control group using REDCap (Research Electronic Data Capture; Vanderbilt University) [49,50] by a graduate research assistant (NC). Randomization was conducted upon completion of the informed consent form, all baseline measures (ie, PA) assessment, and a battery of questionnaires. Participants were blinded to the study hypotheses and were only informed of their assigned group. Outcome assessors were kept blinded to the allocation.

Intervention

Participants in both groups participated in an 8-week, remotely delivered, supervised exercise program conducted via videoconferencing (Zoom [Zoom Communications, Inc]). The program consisted of two 30-minute live sessions and one 30-minute recorded session per week (Figure 1). Recorded sessions aligned with live classes to maintain consistency and allowed for individualized adjustments. All live sessions were led by Qualified Exercise Professionals (QEPs) certified as Registered Kinesiologists or through the American College of Sports Medicine (ACSM) or Canadian Society of Exercise Physiologists (CSEP). Participants were encouraged to keep their cameras on during live sessions to facilitate safety monitoring and provide real-time exercise modifications.

Figure 1. Combined exercise group.



XSL•FO
RenderX

Participants received resistance bands, a Fitbit Inspire 2 heart rate (HR) monitor, and an exercise log. The program included 30 minutes of unsupervised aerobic exercise (eg, walking) 3 times per week, gradually increasing intensity and duration from 40% to 59% of the maximum HR reserve initially to 60%-70% by the end of the program. The program also included 30-minute supervised resistance training sessions 2 times per week. Resistance training focused on major muscle groups using 2-3 sets of 8-12 repetitions at a Rating of Perceived Exertion (RPE) of 12-15, progressively adjusted with stronger resistance bands. A 30-60 second rest was provided between sets and 60 seconds between exercises, during which the QEP demonstrated the next exercise and offered modifications. Participants recorded their HR and RPE in exercise logs after each session. Participants also attended 4 remotely delivered, biweekly, 30-minute behavioral counseling sessions based on the multi-process action control (M-PAC) framework, emphasizing goal-setting, habit formation, and behavioral control to support long-term PA (Table 1 [51-54]).

Table 1. Overview of biweekly behavioral counseling sessions in the combined exercise group.

Week	Topic	Description
2	Benefits of exercise for individuals living with and beyond cancer	Participants learned about the exercise guidelines for general health benefits and the cancer-specific exercise guidelines. Participants were invited to share their meaning of the term "exercise" and some of the exercises they have engaged in or would like to engage in.
4	Goal setting	Participants learned about the importance of goal setting, the difference between short- and long-term goals, SMART[a] goals, and intrinsic and extrinsic rewards. Participants created and discussed their short- and long-term SMART goals and their prospective intrinsic and extrinsic rewards. Participants were invited to share their responses with the group.
6	Habit formation	Participants learned about habits and how to develop PA[b] habits, including "tagging" activities to existing environmental cues and creating "cues to action." Participants brainstormed ideas for creating PA habits using environmental cues that exist within their work and nonworkdays. Participants were invited to share their responses with the group.
8	Barriers and strategies to exercise	Participants learned about cancer-specific barriers and facilitators to exercise. In addition, participants were provided with solutions that may help them overcome some of the barriers to PA. Participants were reminded of past behavioral counseling sessions (ie, revisiting SMART goals and habit formation). Participants applied these concepts into practice by creating a new SMART goal and reward, identifying an environmental cue that will trigger the behavior, and a strategy to overcome any barriers that may arise (or arose) that may get in the way of completing this goal. Participants were invited to share their responses with the group.

[a]SMART: specific, measurable, achievable, relevant, and time-bound.

[b]PA: physical activity.

Active Control Group

The active control group participated in a low-intensity, whole-body stretching and toning program designed to improve balance and flexibility. Exercises targeted major muscle groups potentially affected by cancer treatments, such as steroid use, radiation therapy, or surgery [55]. Flexibility exercises, incorporating both static and dynamic movements, were introduced weekly to maintain engagement and progress using resistance bands. Intensity was maintained at a light level (RPE 9-10) [56]. Participants recorded their HR (using Fitbit devices) and RPE in exercise logs after each session.

Measures

Primary Outcome: Feasibility

Feasibility was assessed by evaluating the enrollment, adherence (ie, attendance and exercise prescription adherence), attrition rates, measurement completion rates, adverse events, program satisfaction, and therapeutic alliance with the QEP. The enrollment rate was determined by the percentage of participants assessed for eligibility who subsequently enrolled. Attendance was expressed as a percentage of exercise sessions attended. Exercise prescription adherence was determined by assessing self-reported RPE and Fitbit Inspire 2-measured HR during each exercise session. Attrition was measured as the percentage of participants who did not complete the intervention. Participants completed a patient satisfaction and therapeutic alliance questionnaire following the 8-week intervention. The Working Alliance Inventory Short Revised (WAI-SR) [57,58] was used to assess therapeutic alliance, with higher scores representing better therapeutic alliance. Closed-ended questions with response options on a 7-point Likert scale ranging from 1 (not at all) to 7 (very much) were used.

Secondary Outcomes

Objectively Measured Cognitive Outcomes

Cognitive function was assessed at baseline and post intervention using the remotely administered National Institutes of Health (NIH) Toolbox Cognition Battery (v2), a validated tool with minimal practice effects comparable to other widely used gold-standard cognitive measures [59,60]. The NIH Toolbox Cognition Battery included 4 tests: the Picture Vocabulary Test, which assessed language and vocabulary knowledge; the Picture Sequence Memory Test, which measured

XSL•FO
RenderX

episodic memory; the List Sort Working Memory Test, which evaluated working memory and executive function updating; and the Auditory Verbal Learning Test, which assessed immediate memory and verbal learning [59,60]. A supplementary measure of executive function shifting—a domain frequently impaired in BC survivors treated with chemotherapy [14,61]—was measured using the PsyToolkit software platform (Gijsbert Stoet) to administer Task Switching cognitive test [62,63].

All cognitive assessments were administered remotely via Zoom at both time points through a shared iPad (9.7" iPad 2; Apple) screen in a quiet, distraction-free room, with each session lasting about 45 minutes. The examiner completed the remote administration training protocol [64] under the guidance of researchers with expertise in cognitive psychology (AFK and JDR). To ensure standardized administration across participants, a single examiner conducted all assessments through screen-sharing to monitor participants and their responses in real-time.

Device-based exercise minutes were measured at baseline and post intervention with ActiGraph GTX3+ accelerometers. Participants wore the accelerometer on their nondominant hip during waking hours for 7 consecutive days. Data were analyzed if there were no extreme counts (> 20,000) and if participants had at least 10 valid hours of wear time on 4 or more days. Data were downloaded in 60-second epochs, processed, and converted to mean counts per minute in the ActiLife software package (v6.13.5; ActiGraph). These counts were used to estimate daily minutes of light (101-1951 counts min^{-1}), moderate (1952-5724 counts min^{-1}), vigorous (≥5725 counts min^{-1}), and total moderate-to-vigorous intensity exercise (≥1952 counts min^{-1}) based on established cut points [65]. Further, weekly minutes spent in each exercise intensity category were calculated and presented.

Demographic and Clinical Information

Demographic and clinical information were collected at baseline using a self-reported standardized health history questionnaire used in previous studies [44,66,67]. Demographic variables included age, marital status, education, employment, ethnicity, smoking, and alcohol use. Clinical information included months since diagnosis, months since treatment, type of cancer treatment received, current cancer status, and general health status. Additional data on comorbidities, such as high blood pressure, high cholesterol, diabetes, and arthritis, were also obtained.

Data Analyses

Data analyses were performed using IBM SPSS Statistics (version 29). As this was a feasibility study, an a priori power calculation was not performed [68,69]. Descriptive statistics (ie, means and SDs) characterized the sample for demographic, clinical, and feasibility outcomes. z-scores were used to identify any outliers in the field. Analyses of covariance (ANCOVA) were used to assess both primary and secondary outcomes of the study. The dependent variable was the mean difference of each intervention outcome (ie, the difference between post- and preintervention scores), while the independent variable was the group (ie, combined exercise or active control). All analyses were conducted while controlling for the baseline values of each outcome. All analyses were conducted on an intention-to-treat basis. Given that the purpose of this feasibility trial was to inform a larger RCT, outcomes were interpreted for potential clinical significance based on the direction and magnitude of numerical differences. Partial eta squared (ηp^2) values were reported to describe the observed effect sizes.

Ethical Considerations

The trial protocol was approved by the Research Ethics Board at the University of Toronto (number 43675), and all participants provided written informed consent. Participant privacy and confidentiality were strictly maintained throughout the study; all data were deidentified and stored on secure, password-protected servers accessible only to authorized research personnel, in accordance with institutional and ethical guidelines. Participants received a total of $20 CAD compensation, plus a Fitbit Inspire 2 for their time and participation, consistent with the approved ethics protocol.

Results

Sample Characteristics

Demographic and clinical characteristics of the participants are summarized in Table 2. Overall, all participants were women (21/21, 100%), predominantly White (16/21, 76%), with an average age of 51.9 (SD 7.2) years. The majority were married (17/21, 81%), some had completed university or college (10/21, 48%), and many were employed full-time (13/21, 62%). All participants had chemotherapy and were, on average, 22.9 (SD 13.9) months post diagnosis and 11.8 (SD 12.9) months post treatment. Regarding weekly exercise minutes, participants engaged in 149.4 (SD 96.1) moderate-to-vigorous intensity exercise minutes per week.

XSL•FO
RenderX

Table 2. Demographic and clinical characteristics of patients with breast cancer at baseline (N=21).

Variable	Value
Age (years), mean (SD)	51.9 (7.2)
Sex, n (%)	
Female	21 (100)
Marital status, n (%)	
Married	17 (81)
Divorced	1 (4.8)
Common-law	3 (14.3)
Education, n (%)	
Completed high school	2 (9.5)
Some university or college	3 (14.3)
Completed university or college	10 (47.6)
Completed graduate school	6 (28.6)
Employment, n (%)	
Disability	4 (19)
Retired	3 (14.3)
Part-time	1 (4.8)
Full-time	13 (61.9)
Race, n (%)	
White	16 (76.2)
Chinese	2 (9.5)
South Asian (East Indian, Pakistani)	2 (9.5)
Other (did not specify)	1 (4.8)
Months (n=19), mean (SD)	
Since treatment	11.8 (12.9)
Since diagnosis	22.9 (13.8)
Treatment type, n (%)	
Chemotherapy	21 (100)
Current cancer status, n (%)	
Currently receiving maintenance therapy	10 (47.6)
Not currently receiving any treatment	9 (42.9)
Currently receiving primary cancer treatment	2 (9.5)
General health status, n (%)	
Good	10 (47.6)
Fair	6 (28.6)
Very good	4 (19)
Excellent	1 (4.8)
Other comorbidities, n (%)	
High blood pressure	3 (14.3)
High cholesterol	3 (14.3)
Diabetes	1 (4.8)
Arthritis	2 (9.5)
Smoking status, n (%)	

XSL•FO
RenderX

Variable	Value
Never smoked	12 (57.1)
Ex-smoker	8 (38.1)
Occasional	1 (4.8)
Alcohol consumption, n (%)	
Never	7 (33.3)
Less than once a month	11 (52.4)
2-3 times per month	2 (9.5)
2-3 times per week	1 (4.8)
Exercise minutes (min/wk), mean (SD)	
Sedentary	5212. 3 (1033.8)
Light intensity	2093.6 (487.1)
Moderate intensity	139.7 (82.9)
Vigorous intensity	8.4 (26.3)
Total moderate-to-vigorous intensity	149.4 (96.1)

Feasibility Outcomes

Participant flowchart is presented in Figure 2. Feasibility outcomes are summarized in Table 3. Of the 41 women who responded to the invitation, 22 were eligible, and 21 consented to be randomized into either the combined exercise group (aerobic+resistance training; n=10) or the active control group (balance and flexibility; n=11), representing a 51.2% enrollment rate. Participants were divided into two cohorts: (1) the first cohort (n=11) began in May 2023, and (2) the second cohort (n=10) began in July 2023. The sample size aligns with previous research on feasibility and group cohesion for counseling programs for patients with BC [70,71]. Overall, 18 participants completed the 8-week intervention. Three participants withdrew after randomization, but before starting the intervention for reasons, including a change in BC treatment plan (n=1), an issue with the timing of the programming and their schedule (n=1), and a change of interest in participating (n=1). An intention-to-treat analysis was conducted, and therefore adherence was calculated based on attendance and included the dropouts (n=3), since they were enrolled and randomized at the start of the intervention. Consequently, the attrition rate was 14.3%. For the ANCOVA analyses, a complete case analysis was applied (n=18).

XSL•FO
RenderX

Figure 2. Participant flowchart through the study.

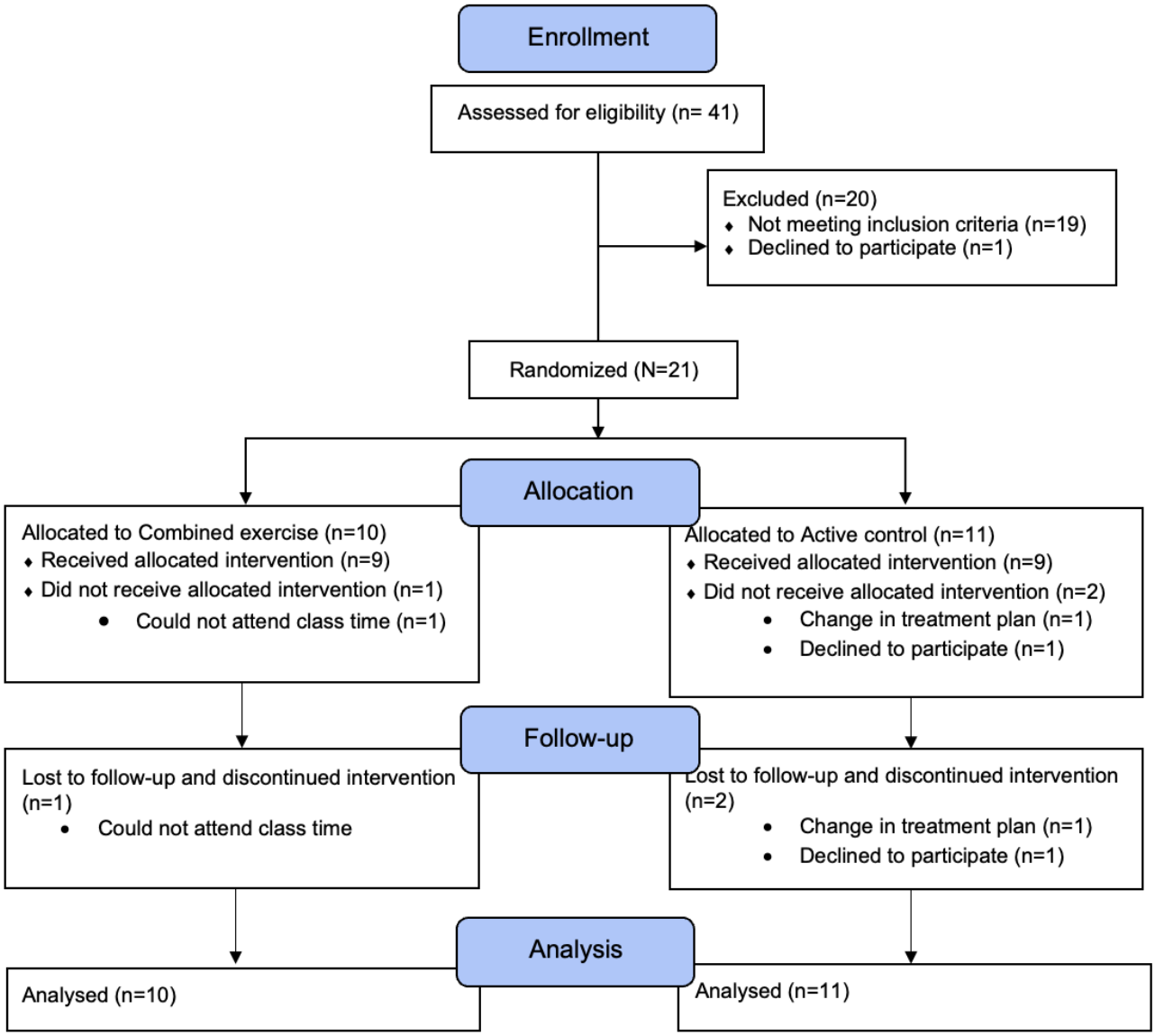


Table 3. Feasibility and acceptability outcomes.

Feasibility variables	Definition	Adherence and completion, n/N (%)
Enrollment rate	Consented and randomized divided those invited	21/41 (51.2)
Attendance adherence (exercise sessions)	Attendance rate in exercise sessions	34/48 (70.8)
Attendance adherence (counseling sessions)[a]	Attendance rate in counseling sessions	31/40 (77.5)
Adverse events	Number of adverse events reported	0/0 (0)
Participant satisfaction questionnaire	Completion rate	18/21 (85.7)
Working Alliance Inventory	Completion rate	17/21 (81)
Remotely delivered NIH[b] Toolbox Cognition Battery	Completion rate	18/21 (85.7)
PsyToolkit executive function test	Completion rate	17/21 (81)
Accelerometer wear compliance	Met wear-time criteria rate	14/21 (66.7)
Attrition	Dropouts after randomization	3/21 (14.3)

[a]Counseling sessions based on the multi-process action control were delivered only to participants assigned to the combined exercise group.

[a]NIH: National Institutes of Health.

XSL•FO
RenderX

Attendance adherence (Table 3) rate for the remotely delivered, combined exercise classes was 70.8% (17/24) and 77.5% (3/4) for behavioral counseling sessions based on the M-PAC that was delivered only to those in the combined exercise group (n=10). Attendance adherence for the active control group was 70.8% (17/24). Participants in the combined exercise group who completed the intervention and postintervention measures (n=9) reported an average RPE of 12.8 and a HR of 111.8 during the resistance training classes. This corresponds to a moderate-intensity RPE and an HR within the light-moderate target HR zone. In addition, participants in the combined exercise intervention (n=9) reported a mean RPE of 12.2 and a mean HR of 119.7 during the unsupervised aerobic (ie, walking) sessions, which corresponds to moderate-intensity RPE and a HR within the moderate-intensity target HR zone. Participants in the active control group who completed the intervention and postintervention measures (ie, including the exercise log; n=8) reported a mean RPE of 11.1 and a mean HR of 85.5 during their balance and flexibility program, which corresponds to a light-moderate intensity RPE and light-intensity target HR zone. There were no adverse events during the intervention in either group. The measurement completion rates for the study were as follows: 85.7% (18/21) for the participant satisfaction questionnaire and 81% (17/21) for the Working Alliance Inventory. For objectively measured cognitive function, completion rates were 85.7% (18/21) for the memory test and 81% (17/21) for the executive function test. Additionally, accelerometer wear compliance was achieved by 66.7% (14/21) of participants.

Participant responses regarding program satisfaction are summarized in Tables 4 and 5. Overall, participants reported high satisfaction with the intervention. The majority of participants (17/18, 94.4%) agreed or strongly agreed that the program was rewarding and appreciated the virtual delivery format. Most participants (15/18, 83.3%) agreed or strongly agreed that the intervention helped them increase their PA levels, and 77.7% (14/18) indicated that they would recommend the program to other patients with BC. Most participants found the intervention beneficial, and participants 62.5%) agreed or strongly agreed that the behavioral counseling sessions were personally useful. Similarly, 61.1% of participants agreed or strongly agreed that the intervention was not burdensome.

Table 4. Satisfaction scores from the participant satisfaction questionnaire in patients with breast cancer at post intervention (ie, 8 weeks; n=18).

Variable	Value, mean (SD)	Strongly agree and agree (%)
Overall intervention participation		
I found the exercise intervention rewarding.	4.6 (0.6)	94.4
I enjoyed the virtual delivery of the exercise intervention.	4.6 (0.6)	94.4
The exercise intervention helped me to increase my physical activity levels.	4.0 (1.1)	83.3
The exercise intervention helped me to increase my overall fitness.	3.9 (1.1)	72.2
The behavioral counseling sessions were useful for me personally.	3.1 (1.5)	62.5
I am more confident in my goal-setting and coping abilities after participating in the intervention.	3.7 (1.1)	61.1
The research study was useful for me personally.	3.7 (1.0)	83.3
I would recommend the exercise intervention to other patients with breast cancer.	4.2 (0.7)	77.7

Table 5. Burden scores from the participant satisfaction questionnaire in patients with breast cancer at post intervention (ie, 8 weeks; n=18).

Burden	Value, mean (SD)	Strongly disagree and disagree (%)
The exercise intervention was a waste of time.	1.8 (1.3)	83.3
The behavioral counseling sessions were a waste of time.	2.3 (1.5)	75.0
The questionnaires were a burden.	2.3 (1.2)	61.1
The neurological assessments (iPad cognition tests) were a burden.	2.1 (1.1)	77.8
The 3-times weekly exercise class was a burden.	2.1 (0.9)	77.8
It was difficult to participate in the exercise classes because of cancer or chemotherapy-related side effects.	2.5 (1.3)	55.6

Participant responses on the WAI-SR are summarized in Table 6. Overall, participants reported positive perceptions of their alliance with the QEP. Most (16/18, 88.2%) indicated that they very often or always felt mutual respect, and a substantial majority (14/18, 76.5%) felt appreciated and cared for.

XSL•FO
RenderX

Table 6. Working alliance inventory-short revised scores in patients with breast cancer post intervention (ie, 8 weeks; n=17)[a].

Variable	Value, mean (SD)	Very often or always (%)
As a result of these sessions, I am clearer as to how I might be able to change.	2.9 (1.2)	35.3
What I am doing in therapy gives me new ways of looking at my problem.	3.2 (1.3)	41.2
I believe [INSTRUCTOR] likes me.	4.1 (0.8)	70.6
[INSTRUCTOR] and I collaborate on setting goals for my therapy.	2.9 (1.4)	35.3
[INSTRUCTOR] and I respect each other.	4.5 (0.7)	88.2
[INSTRUCTOR] and I are working towards mutually agreed upon goals.	3.2 (1.2)	47.1
I feel that [INSTRUCTOR] appreciates me.	4.2 (0.9)	76.5
[INSTRUCTOR] and I agree on what is important for me to work on.	3.2 (1.2)	35.3
I believe the way we are working with my problem is correct.	3.7 (1.3)	64.7
I feel that the things I do in therapy will help me to accomplish the changes that I want.	3.6 (1.3)	64.7
[INSTRUCTOR] and I have established a good understanding of the kind of changes that would be good for me.	3.1 (1.5)	41.2
I feel [INSTRUCTOR] cares about me even when I do things that he/she does not approve of.	4.2 (0.9)	76.5

[a]Participants rated items on a 5-point Likert scale, where 1=seldom, 2=sometimes, 3=fairly often, 4=very often, and 5=always. Participants were told to replace the term “therapy” with “exercise program” when thinking about their responses to each item.

Objectively Measured Cognitive Outcomes

Changes in objectively measured cognitive function scores from baseline to post intervention for the combined exercise group vs the active control group are presented in Table 7. There were no statistical differences in objectively measured cognitive function in the combined exercise group compared to the active control group. However, adjusted group mean differences trending toward the combined exercise group were observed in the scores for the Picture Sequence Memory Task (mean difference +5.33, 95% CI –12.5 to 23.2; ηp^2=0.03), List Sorting Working Memory Test (mean difference +8.17, 95% CI –4.2 to 20.6; ηp^2=0.12), and Auditory Verbal Learning (mean difference +3.22, 95% CI –2.0 to 8.5; ηp^2=0.12). Table 7 also shows the 4 cognition outcomes from the PsyToolkit: repeat trial time (ms), switch trial time (ms), repeat trials correct (%), and switch trials correct (%). No significant differences were found between the combined exercise group compared to the active control group. However, a reduction in repeat trial time trended toward the combined exercise group (mean difference –146.14, 95% CI 420.87 to128.59 ms; ηp^2=0.04). Likewise, a greater decrease in the switch trial time trended toward the combined exercise group (mean difference –394.35, 95% CI –1035.67 to 246.96 ms; ηp^2=0.07) compared to the control group (mean difference –124.39 95% CI –522.06 to 273.27 ms; ηp^2=0.07). For accuracy, improvements in repeat trials correct (mean difference +4.86%, 95% CI –0.61 to 10.36; ηp^2=0.02) and switch trials correct (mean difference +19.70%, –0.30 to 39.70; ηp^2=0.02) also trended towards the combined exercise group, while the control group exhibited decreases in these measures.

XSL•FO
RenderX

Table 7. Effects of a supervised, remotely delivered, combined exercise or active control group on cognitive function at post intervention (n=18).

Variable	Pre, mean (SD)	Post, mean (SD)	Mean difference (95% CI)	*P* value	Partial η^2
Picture sequence memory			5.33 (–12.5 to 23.2)	.53	0.03
Combined exercise (n=9)	107.1 (12.1)	112.2 (16.6)			
Active control (n=9)	108.4 (18.8)	107.6 (22.0)			
List sorting memory			8.17 (–4.2 to 20.6)	.18	0.12
Combined exercise (n=9)	106.4 (16.1)	111.4 (13.3)			
Active control (n=9)	100.7 (17.7)	102.0 (11.2)			
Auditory verbal learning test			3.22 (–2.0 to 8.5)	.21	0.12
Combined exercise (n=9)	28.5 (5.9)	32.0 (6.7)			
Active control (n=9)	26.1 (7.0)	26.6 (8.2)			
Oral reading recognition			–9.65 (–22.9 to 3.5)	.14	0.14
Combined exercise (n=9)	129.8 (22.4)	127.7 (20.4)			
Active control (n=9)	128.0 (15.01)	136.4 (9.0)			
Picture vocabulary			–2.48 (–5.2 to 0.3)	.07	0.20
Combined exercise (n=9)	103.7 (14.5)	102.0 (15.0)			
Active control (n=9)	106.0 (12.0)	106.7 (11.6)			
Repeat trials reaction time (ms)				.42	0.04
Combined exercise (n=9)	1266.23 (86.29)	1120.09 (88.92)	–146.14 (–420.87 to128.59)		
Active control (n=8)	1119.22 (68.91)	1179.75 (91.80)	60.53 (–132.71 to 253.78)		
Switch trials reaction time (ms)				.34	0.07
Combined exercise (n=9)	2268.51 (205.67)	1874.16 (123.64)	–394.35 (–1035.67 to 246.96)		
Active control (n=8)	2227.06 (114.37)	2102.67 (187.89)	–124.39 (–522.06 to273.27)		
Percentage of correct repeat trials (%)				.62	0.02
Combined exercise (n=9)	90.59 (1.61)	95.46 (1.67)	4.86 (–0.61 to 10.36)		
Active control (n=8)	96.68 (0.92)	92.86 (2.66)	–3.83 (–10.62 to 2.96)		
Percentage of correct switch trials (%)				.63	0.02
Combined exercise (n=9)	69.70 (6.86)	89.39 (2.83)	19.70 (–0.30 to39.70)		
Active control (n=8)	92.05 (2.06)	81.25 (5.26)	–10.80 (–24.26 to 2.67)		

Table 8 presents the changes in device-based exercise levels among patients with BC. There were no statistical differences between the two groups. However, an increase in total moderate-to-vigorous intensity exercise minutes was observed from pre to post intervention, trending toward participants in the active control group compared to those in the combined exercise group (mean difference in change 58.8, 95% CI 105.8.8-223.6 minutes; ηp^2=0.06).

XSL•FO
RenderX

Table 8. Effects of supervised, remotely delivered, combined exercise or active control program on total moderate-to-vigorous exercise in patients with breast cancer at pre and post intervention (ie, 8 weeks; n=14).

Variable	Pre intervention, mean (SD)	Post intervention, mean (SD)	Adjusted mean difference between groups[a]		
			Mean difference (95% CI)	*P* value	Partial n^2
Total moderate-to-vigorous intensity exercise			58.8 (105.8.8-223.6)	.44	0.06
Active control (n=6)	179.3 (110.3)	272.0 (230.6)			
Combined exercise (n=8)	127.0 (96.0)	148.5 (162.9)			
Light-intensity exercise			74.5 (–625.0 to 774.1)	.82	0.15
Active control (n=6)	1940.3 (431.9)	1991.6 (319.9)			
Combined exercise (n=8)	2140.4 (388.8)	2168.3 (555.5)			
Moderate-intensity exercise			26.6 (–89.8 to 143.0)	.62	0.02
Active control (n=6)	158.6 (89.8)	211.6 (153.3)			
Combined exercise (n=8)	123.2 (88.1)	144.1 (154.5)			
Vigorous-intensity exercise			48.2 (–6.6 to 103.0)	.08	0.28
Active control (n=6)	17.3 (41.9)	55.2 (71.1)			
Combined exercise (n=8)	3.75 (10.6)	4.25 (9.01)			

[a]Mean differences were adjusted for the baseline value of the outcome and total minutes of wear time.

Discussion

Principal Findings

To our knowledge, this is the first study to pilot a remotely delivered, supervised combined exercise program for patients with BC following chemotherapy, exhibiting mild cognitive impairment. The findings of this study will be used to inform larger, definitive RCTs by examining which aspects of the trial were feasible, as well as further refinements that need to be made. Specifically, there are few exercise training studies that target BC treatment-related cognitive impairment as a primary outcome. Recent reviews examined evidence for the use of exercise as an intervention for CRCI [25,31,72]. However, these studies did not consistently measure cognitive function as a primary outcome, and less than half assessed cognitive function with well-normed objective measures. The varied results within neuropsychological testing and the recent emphasis on individualized care all highlight the critical need to identify patients with BC who are especially at risk for cognitive decline. Our pilot is noteworthy, as it addressed some of these conceptual and methodological issues that may advance progress in the field. These include having an active control condition to match contact time, examining different types of exercise, such as combined exercise, including computer-based objectively measured cognitive function, as well as theory-based, behavioral counseling for exercise uptake and adherence.

This study was feasible and exceeded the a priori feasibility indicators of adherence rate ≥70%, attrition rate <30%, and no adverse events, and was well received by patients with BC. Adherence rates in our remotely delivered trial were consistent (63%-100% [28,30,73]) or lower (88% [72]; 87.9% [26]; 80% [74]; 99.9% [75]) than previous exercise trials targeting CRCI. However, it is important to note that these exercise trials were supervised (in person) sessions that included different exercise modalities, such as high-intensity interval training [28,73], combined exercise with Nordic or power walking [30], home-based walking [71], and aerobic exercise [26,72,75]. Nevertheless, our adherence rates are encouraging when compared to other virtually supervised exercise interventions delivered via Zoom in a group format, with various combinations of aerobic and resistance exercises, where adherence ranged from 78% to 100% [76]. Our trial also had low attrition rates, which were consistent with prior research with remotely delivered exercise interventions (0%-29%; [76]).

Recruitment rates for virtually supervised exercise interventions vary widely, with reported rates between 8% and 60% [76] likely due to the heterogeneity of recruitment methods. While our recruitment rate (51.2%) was consistent with previous research, our study relied mainly on self-referrals, which tended to attract highly motivated cancer survivors who were well-educated, White, and in better overall health and physical functioning. Homogeneous samples (eg, White, young, and highly educated) are a common limitation across distance-based exercise oncology trials [77]. Recruitment strategies and partnerships with organizations that serve racial and ethnic populations of cancer survivors should be strengthened at the onset of the intervention, which may assist with recruitment efforts [78].

In our secondary outcomes, no statistical differences were observed in objectively measured cognitive function. Small to medium effect size improvements were observed in episodic memory (Picture Sequence Memory Task), working memory, and executive function updating (List Sorting Working Memory Test), and immediate memory and verbal learning (Auditory Verbal Learning), trending toward the combined exercise group. On the other hand, large effect size improvements were observed in crystallized abilities (Oral Reading Recognition Test) plus language and vocabulary (Picture Vocabulary), slightly trending

XSL•FO
RenderX

toward the active control group. For executive function shifting, greater accuracy and reductions in reaction time (small to medium effect sizes) were observed for the switch trials in the combined exercise group. To date, few studies have tested the effects of exercise on objectively measured cognitive function outcomes in patients with BC following chemotherapy, with mixed results [40,72,75]. Preliminary results from these trials support cognitive changes following an exercise intervention, which is consistent with a study that examined the effects of a 3-month PA intervention compared to a waitlist control arm on cognitive functioning in patients with BC [43]. The intervention improved processing speed only among patients with BC diagnosed within the past 2 years. One proof-of-concept aerobic exercise intervention examined cognitive function using objective neuropsychological testing and functional magnetic resonance imaging in patients with BC. This study found improvements in cognitive symptoms with notable patterns of neural activation. However, the changes either did not reach statistical significance or had small effect sizes [72]. Finally, one of the first large-scale studies to examine the effects of aerobic exercise on neurocognitive function tested the effects of 6 months of moderate-intensity aerobic exercise on neurocognitive function in patients with BC receiving endocrine therapy. The trial focused on women who initiate exercise within 2 years of completing primary therapy (surgery +/− chemotherapy) [75]. Findings show improved performance in processing speed and verbal memory in the exercise group and no change in the usual care control group. It is important to note that these studies were either home-based aerobic programs or in person supervised aerobic activity. They also used a wide range of cognitive measures that were collected in person, whereas our study used a cognitive battery that was assessed remotely.

Recent reviews on the use of exercise as a potential intervention for cognitive deficits associated with cancer and its treatment demonstrated notable improvements in the cognitive domains of working and visual memory, attention and concentration, cognitive flexibility, inhibitory control, and verbal fluency [26,79], consistent with our study. Null effects were observed in other RCTs with a walking intervention [74], exercise intervention with supervised aerobic and strength training, as well as Nordic or power walking [30], and interval and continuous aerobic exercise in patients with BC following chemotherapy [73]. While findings from previous trials have been inconclusive, the positive improvements in cognitive function reported in our study could be attributed to recruiting patients with BC who exhibit mild cognitive impairment at the onset of the intervention. This may have provided greater room for improvement post intervention. Also, a range of cognitive assessments measuring cognitive domains that are most impaired by chemotherapy (eg, memory, attention, and executive function) increased the sensitivity to detect impairments post intervention. Finally, patients with BC in the combined exercise intervention were adhering to the prescribed exercise dose (ie, moderate-to-vigorous intensity physical activity [MVPA]), which is critical in achieving greater cognitive benefits and is consistent with reviews demonstrating that sustained MVPA elicits changes in cognitive function in patients with BC [3,25,80]. One of the most important mediators of cognition in patients with BC may be cardiorespiratory fitness [81,82]. Exercise-induced changes in cognitive function can be explained by alterations in brain morphology and functional dynamics (ie, prefrontal cortex and hippocampus) and may be partially mediated by improvements in cardiorespiratory fitness [82]. This underscores the pivotal role of cardiorespiratory fitness in amplifying the neuroprotective effects of MVPA. Given the link between cognition and cardiorespiratory fitness and the established effects of exercise on cardiorespiratory fitness, exercise might also remediate CRCI in patients with BC.

Notable improvements were also observed in the active control group performing light-intensity exercise, but these improvements were not in the cognitive domains most impaired by chemotherapy. The active control group engaged in a light-intensity exercise, which is insufficient to elicit improvements in cardiorespiratory fitness. Previous reviews demonstrated that both aerobic and resistance training that were moderate-to-vigorous intensity may be required to elicit cognitive benefits [3]. Nevertheless, improvements in the Oral Reading Recognition Test and Picture Vocabulary Test trended toward the active control group (although exploratory). These domains are related to language and/or crystallized abilities, and it is possible that increases in cognitive performance likely reflect expectancy or practice effects from repeated exposure to the assessments. Therefore, the remote administration of the NIH Toolbox Cognition Battery may not fully capture subtle CRCI-related changes.

In terms of changes in exercise minutes, improvements in MVPA from pre to post intervention trended toward the active control group (although exploratory). This finding is inconsistent with previous studies, which demonstrated improvements in device-based exercise minutes in patients with BC [30,43]. Given that we had an active control group, both groups were motivated to engage in exercise. Accelerometers do not accurately measure resistance training, which may not have been captured in patients with BC, especially in the combined exercise group, who continued engaging in resistance training. Patients with BC in the combined exercise group may have engaged in less structured exercise outside of the planned classes because they were receiving a moderate-intensity multicomponent exercise program compared to the active control group. Therefore, the reported effect sizes between our groups are possibly smaller given the active comparator group. Also, patients with BC self-reported not meeting exercise guidelines at screening but subsequently showed higher levels of activity with device-based measures at baseline. Future research should consider using device-based measures to screen patients with BC not meeting exercise guidelines to ensure that the inclusion criteria are met.

Lessons Learned and Future Research Directions

Overall, patients with BC were highly satisfied with the intervention. Most patients with BC found the study rewarding, enjoyed the virtual delivery format, would recommend the study to other patients with BC, and did not find the study assessments burdensome. Patients with BC also felt a mutual respect with their QEP and felt appreciated by their QEP, further contributing to their positive experience. This suggests that few adjustments

XSL•FO
RenderX

need to be made to the intervention or the study design before implementing a future RCT. In our study, an additional QEP attended the classes to monitor participant safety and to assist the lead QEP in managing the virtual environment. This was beneficial to both the lead QEP and patients with BC, and future remotely delivered interventions should consider a second QEP. Additionally, the QEP's personality, approachability, and teaching style played an important role in the enjoyment and feasibility for patients with BC. Future studies should consider QEPs with cancer-specific training to offer exercise modifications to patients with BC with limitations or injuries. Finally, our classes included 15 minutes of optional time before or after each class for patients with BC to interact for overall class enjoyment and accountability and to help foster a sense of community. This was beneficial to the overall satisfaction of the trial and should be implemented in future studies.

Strengths and Limitations

Our trial should be interpreted within the context of important strengths and limitations. The remotely delivered component of the trial was a strength, as it removed preexisting barriers to exercise (eg, time constraints, access, and hesitancy toward public space). The combined exercise and active control group sessions were delivered by knowledgeable QEPs who were engaging and provided modifications that allowed all patients with BC to participate in exercise, despite any physical limitations. The addition of a recorded session reduced participant burden by providing flexibility and autonomy. Our study was strengthened by the inclusion of behavioral counseling sessions, which were guided by the M-PAC framework to encourage exercise adoption and adherence [51]. The inclusion of a rigorous comparison group was an additional strength since both groups received supervised, remotely delivered exercise programs at differing intensities. Finally, the use of a standardized, well-validated battery of cognitive tests via the NIH Toolbox was another strength and in line with the ICCTF's recommendations [6].

Although this study had a small sample size and a heterogeneous population, which limits the generalizability of the findings, this feasibility study is crucial for calculating sample sizes and refining methodologies for larger, future trials. In addition, the sample was predominantly White, highly educated, middle- to older-aged, and within the early cancer survivorship period. The focus on early-stage patients with BC within 48 months of completing treatment may not be generalizable to other cancer populations, such as individuals with metastatic disease or patients who are pretreatment, undergoing treatment, or immediately post treatment. Furthermore, the heterogeneity in the time since treatment in our study may preclude identifying the optimal exercise timing to prevent or ameliorate CRCI, which requires future investigation. It is important to note that patients with BC may be more likely to seek out care for cancer rehabilitation and view themselves as research participants compared to patients with BC who are more culturally and socially diverse populations. Future research should consider the recommendations for increasing racial and ethnic diversity in cancer clinical trials by the American Society of Clinical Oncology and Association of Community Cancer Centers Joint Research Statement [78]. Another limitation of this study was not accounting for menopausal status, which could be a confounding variable that may impact cognitive function in patients with BC. Future studies should consider chemotherapy-induced menopause as a possible mechanism for CRCI. Furthermore, our study did not include self-reported cognitive function, which is another important indicator of CRCI, as it taps into day-to-day cognitive performance in relation to cancer and its treatments [6,83]. Therefore, including both self-report and objective cognitive function assessments in future research is warranted.

Conclusions

In conclusion, this feasibility study provides preliminary evidence that a remotely delivered, combined exercise program is feasible and may improve objectively measured cognitive function. There are a few exercise studies that have rigorously evaluated CRCI as a primary outcome in patients with BC. The limited number of studies conducted makes it difficult to draw conclusions about the efficacy of exercise for improving cognitive function in patients with BC. There is a need for rigorous and well-designed RCTs. While the use of shorter exercise protocols like our study provides preliminary evidence for cognitive benefits, long-term exercise is more efficacious in addressing age-related cognitive decline and improving cognition [84]. Future trials should include active control conditions, test multicomponent exercises (ie, aerobic and resistance training), and examine the mechanisms underlying the relationship between exercise and cognitive function. Given that CRCI is an emerging cancer survivorship issue, the next step for this work is to conduct a larger RCT that is adequately powered to examine the effects of a remotely delivered, combined exercise intervention on cognitive function in patients with BC following chemotherapy. If replicated, these findings have the potential to identify effective and modifiable management strategies to enhance cognitive health in patients with BC.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to the sensitive information of patients but are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: LT, NC, HX, CMS, EW, JDR, AFK, EM
Formal analysis: LT, NC, HX
Investigation: LT, NC, HX
Methodology: LT, NC, HX, CMS, EW, JDR, AFK, EM
Project administration: LT, NC, HX
Writing – original draft: LT, NC, HX
Writing – review & editing: CMS, EW, JDR, AFK, EM, MH, SO

Conflicts of Interest

None declared

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2938 KB - cancer_v12i1e73393_app1.pdf]

References

1. Cathcart-Rake EJ, Tevaarwerk AJ, Haddad TC, D'Andre SD, Ruddy KJ. Advances in the care of breast cancer survivors. BMJ 2023;382:e071565. [doi: 10.1136/bmj-2022-071565] [Medline: 37722731]
2. Lange M, Joly F, Vardy J, Ahles T, Dubois M, Tron L, et al. Cancer-related cognitive impairment: an update on state of the art, detection, and management strategies in cancer survivors. Ann Oncol 2019;30(12):1925-1940 [FREE Full text] [doi: 10.1093/annonc/mdz410] [Medline: 31617564]
3. Campbell KL, Zadravec K, Bland KA, Chesley E, Wolf F, Janelsins MC. The effect of exercise on cancer-related cognitive impairment and applications for physical therapy: systematic review of randomized controlled trials. Phys Ther 2020;100(3):523-542 [FREE Full text] [doi: 10.1093/ptj/pzz090] [Medline: 32065236]
4. Ahles TA, Root JC, Ryan EL. Cancer- and cancer treatment-associated cognitive change: an update on the state of the science. J Clin Oncol 2012;30(30):3675-3686 [FREE Full text] [doi: 10.1200/JCO.2012.43.0116] [Medline: 23008308]
5. Janelsins MC, Kesler SR, Ahles TA, Morrow GR. Prevalence, mechanisms, and management of cancer-related cognitive impairment. Int Rev Psychiatry 2014;26(1):102-113 [FREE Full text] [doi: 10.3109/09540261.2013.864260] [Medline: 24716504]
6. Wefel JS, Vardy J, Ahles T, Schagen SB. International Cognition and Cancer Task Force recommendations to harmonise studies of cognitive function in patients with cancer. Lancet Oncol 2011;12(7):703-708. [doi: 10.1016/S1470-2045(10)70294-1] [Medline: 21354373]
7. Wefel JS, Schagen SB. Chemotherapy-related cognitive dysfunction. Curr Neurol Neurosci Rep 2012;12(3):267-275. [doi: 10.1007/s11910-012-0264-9] [Medline: 22453825]
8. Wigmore P. The effect of systemic chemotherapy on neurogenesis, plasticity and memory. Curr Top Behav Neurosci 2013;15:211-240. [doi: 10.1007/7854_2012_235] [Medline: 23239468]
9. Vardy J, Wefel J, Ahles T, Tannock I, Schagen S. Cancer and cancer-therapy related cognitive dysfunction: an international perspective from the Venice cognitive workshop. Ann Oncol 2008;19(4):623-629 [FREE Full text] [doi: 10.1093/annonc/mdm500] [Medline: 17974553]
10. Ahles TA, Hurria A. New challenges in psycho-oncology research IV: cognition and cancer: conceptual and methodological issues and future directions. Psychooncology 2018;27(1):3-9. [doi: 10.1002/pon.4564] [Medline: 29044730]
11. Hasmüller S, Bauerfeind I, Debus G. Chemotherapy and post-traumatic stress in the causation of cognitive dysfunction in breast cancer patients. J Natl Cancer Inst 2017;109(10):djx057. [doi: 10.1093/jnci/djx057] [Medline: 28521364]
12. Ahles TA, Root JC. Cognitive effects of cancer and cancer treatments. Annu Rev Clin Psychol 2018;14:425-451 [FREE Full text] [doi: 10.1146/annurev-clinpsy-050817-084903] [Medline: 29345974]
13. Wefel JS, Kesler SR, Noll KR, Schagen SB. Clinical characteristics, pathophysiology, and management of noncentral nervous system cancer-related cognitive impairment in adults. CA Cancer J Clin 2015;65(2):123-138 [FREE Full text] [doi: 10.3322/caac.21258] [Medline: 25483452]
14. Bernstein LJ, McCreath GA, Nyhof-Young J, Dissanayake D, Rich JB. A brief psychoeducational intervention improves memory contentment in breast cancer survivors with cognitive concerns: results of a single-arm prospective study. Support Care Cancer 2018;26(8):2851-2859. [doi: 10.1007/s00520-018-4135-z] [Medline: 29525841]
15. Dumas J. Lifetime hormone exposure and cognition in postmenopausal women with breast cancer. Menopause 2013;20(9):883-885. [doi: 10.1097/GME.0b013e3182a358c4] [Medline: 23942248]
16. Ryan J, Scali J, Carrière I, Amieva H, Rouaud O, Berr C, et al. Impact of a premature menopause on cognitive function in later life. BJOG 2014;121(13):1729-1739. [doi: 10.1111/1471-0528.12828] [Medline: 24802975]
17. Vearncombe K, Pachana N. Is cognitive functioning detrimentally affected after early, induced menopause? Menopause 2009;16(1):188-198. [doi: 10.1097/gme.0b013e3181775eb4] [Medline: 18724262]

XSL•FO
RenderX

18. Paraska KK, Bender CM. Cognitive dysfunction following adjuvant chemotherapy for breast cancer: two case studies. Oncol Nurs Forum 2003;30(3):473-478. [doi: 10.1188/03.ONF.473-478] [Medline: 12719746]
19. Schilder CM, Seynaeve C, Beex LV, Boogerd W, Linn SC, Gundy CM, et al. Effects of tamoxifen and exemestane on cognitive functioning of postmenopausal patients with breast cancer: results from the neuropsychological side study of the tamoxifen and exemestane adjuvant multinational trial. J Clin Oncol 2010;28(8):1294-1300. [doi: 10.1200/JCO.2008.21.3553] [Medline: 20142601]
20. Klemp JR, Myers JS, Fabian CJ, Kimler BF, Khan QJ, Sereika SM, et al. Cognitive functioning and quality of life following chemotherapy in pre- and peri-menopausal women with breast cancer. Support Care Cancer 2018;26(2):575-583 [FREE Full text] [doi: 10.1007/s00520-017-3869-3] [Medline: 28849337]
21. Vega JN, Dumas J, Newhouse PA. Self-reported chemotherapy-related cognitive impairment compared with cognitive complaints following menopause. Psychooncology 2018;27(9):2198-2205 [FREE Full text] [doi: 10.1002/pon.4796] [Medline: 29904964]
22. Liu-Ambrose T, Nagamatsu LS, Graf P, Beattie BL, Ashe MC, Handy TC. Resistance training and executive functions: a 12-month randomized controlled trial. Arch Intern Med 2010;170(2):170-178 [FREE Full text] [doi: 10.1001/archinternmed.2009.494] [Medline: 20101012]
23. Nouchi R, Taki Y, Takeuchi H, Sekiguchi A, Hashizume H, Nozawa T, et al. Four weeks of combination exercise training improved executive functions, episodic memory, and processing speed in healthy elderly people: evidence from a randomized controlled trial. Age (Dordr) 2014;36(2):787-799 [FREE Full text] [doi: 10.1007/s11357-013-9588-x] [Medline: 24065294]
24. Ren X, Wang X, Sun J, Hui Z, Lei S, Wang C, et al. Effects of physical exercise on cognitive function of breast cancer survivors receiving chemotherapy: a systematic review of randomized controlled trials. Breast 2022;63:113-122 [FREE Full text] [doi: 10.1016/j.breast.2022.03.014] [Medline: 35366505]
25. Brunet J, Sharma S. A scoping review of studies exploring physical activity and cognition among persons with cancer. J Cancer Surviv 2024;18(6):2033-2051. [doi: 10.1007/s11764-023-01441-x] [Medline: 37561316]
26. Brunet J, Sharma S, Zadravec K, Taljaard M, LeVasseur N, Srikanthan A, et al. Aerobic exercise and CogniTIVe functioning in women with breAsT cancEr (ACTIVATE): a randomized controlled trial. Cancer 2025;131(1):e35540 [FREE Full text] [doi: 10.1002/cncr.35540] [Medline: 39428863]
27. Mijwel S, Backman M, Bolam KA, Jervaeus A, Sundberg CJ, Margolin S, et al. Adding high-intensity interval training to conventional training modalities: optimizing health-related outcomes during chemotherapy for breast cancer: the OptiTrain randomized controlled trial. Breast Cancer Res Treat 2018;168(1):79-93 [FREE Full text] [doi: 10.1007/s10549-017-4571-3] [Medline: 29139007]
28. Bolam KA, Mijwel S, Rundqvist H, Wengström Y. Two-year follow-up of the OptiTrain randomised controlled exercise trial. Breast Cancer Res Treat 2019;175(3):637-648 [FREE Full text] [doi: 10.1007/s10549-019-05204-0] [Medline: 30915663]
29. Galiano-Castillo N, Arroyo-Morales M, Lozano-Lozano M, Fernández-Lao C, Martín-Martín L, Del-Moral-Ávila R, et al. Effect of an internet-based telehealth system on functional capacity and cognition in breast cancer survivors: a secondary analysis of a randomized controlled trial. Support Care Cancer 2017;25(11):3551-3559. [doi: 10.1007/s00520-017-3782-9] [Medline: 28639097]
30. Koevoets EW, Schagen SB, de Ruiter MB, Geerlings MI, Witlox L, van der Wall E, et al. Effect of physical exercise on cognitive function after chemotherapy in patients with breast cancer: a randomized controlled trial (PAM study). Breast Cancer Res 2022;24(1):36 [FREE Full text] [doi: 10.1186/s13058-022-01530-2] [Medline: 35619188]
31. Myers JS, Erickson KI, Sereika SM, Bender CM. Exercise as an intervention to mitigate decreased cognitive function from cancer and cancer treatment: an integrative review. Cancer Nurs 2018;41(4):327-343 [FREE Full text] [doi: 10.1097/NCC.0000000000000549] [Medline: 29194066]
32. Pinheiro VHG, Pinto SS, Cardozo PS, Andrade LS, Alberton CL. The effect of exercise on cognitive function in breast cancer survivors: systematic review and meta-analysis of randomized controlled trials. Support Care Cancer 2025;33(7):577. [doi: 10.1007/s00520-025-09647-8] [Medline: 40514607]
33. Arch JJ, Stanton AL. Examining the "usual" in usual care: a critical review and recommendations for usual care conditions in psycho-oncology. Support Care Cancer 2019;27(5):1591-1600 [FREE Full text] [doi: 10.1007/s00520-019-04677-5] [Medline: 30788625]
34. Steins Bisschop CN, Courneya KS, Velthuis MJ, Monninkhof EM, Jones LW, Friedenreich C, et al. Control group design, contamination and drop-out in exercise oncology trials: a systematic review. PLoS One 2015;10(3):e0120996 [FREE Full text] [doi: 10.1371/journal.pone.0120996] [Medline: 25815479]
35. Freedland KE, Mohr DC, Davidson KW, Schwartz JE. Usual and unusual care: existing practice control groups in randomized controlled trials of behavioral interventions. Psychosom Med 2011;73(4):323-335 [FREE Full text] [doi: 10.1097/PSY.0b013e318218e1fb] [Medline: 21536837]
36. Smith-Turchyn J, Richardson J, Tozer R, McNeely M, Thabane L. Physical activity and breast cancer: a qualitative study on the barriers to and facilitators of exercise promotion from the perspective of health care professionals. Physiother Can 2016;68(4):383-390 [FREE Full text] [doi: 10.3138/ptc.2015-84] [Medline: 27904238]

XSL•FO
RenderX

37. Phillips SM, Conroy DE, Keadle SK, Pellegrini CA, Lloyd GR, Penedo FJ, et al. Breast cancer survivors preferences for technology-supported exercise interventions. Support Care Cancer 2017;25(10):3243-3252 [FREE Full text] [doi: 10.1007/s00520-017-3735-3] [Medline: 28470368]
38. Sheehan P, Denieffe S, Murphy NM, Harrison M. Exercise is more effective than health education in reducing fatigue in fatigued cancer survivors. Support Care Cancer 2020;28(10):4953-4962. [doi: 10.1007/s00520-020-05328-w] [Medline: 32020356]
39. Denton F, Power S, Waddell A, Birkett S, Duncan M, Harwood A, et al. Is It really home-based? a commentary on the necessity for accurate definitions across exercise and physical activity Programmes. Int J Environ Res Public Health 2021;18(17):9244 [FREE Full text] [doi: 10.3390/ijerph18179244] [Medline: 34501833]
40. Garcia DO, Thomson CA. Physical activity and cancer survivorship. Nutr Clin Pract 2014;29(6):768-779 [FREE Full text] [doi: 10.1177/0884533614551969] [Medline: 25335787]
41. Ramos PG, Júdice PB, Nobre I, Carraça EV. Home-based exercise interventions' impact on breast cancer survivors' functional performance: a systematic review. J Cancer Surviv 2025;19(4):1222-1235. [doi: 10.1007/s11764-024-01545-y] [Medline: 38356019]
42. Galiano-Castillo N, Cantarero-Villanueva I, Fernández-Lao C, Ariza-García A, Díaz-Rodríguez L, Del-Moral-Ávila R, et al. Telehealth system: a randomized controlled trial evaluating the impact of an internet-based exercise intervention on quality of life, pain, muscle strength, and fatigue in breast cancer survivors. Cancer 2016;122(20):3166-3174 [FREE Full text] [doi: 10.1002/cncr.30172] [Medline: 27332968]
43. Hartman SJ, Nelson SH, Myers E, Natarajan L, Sears DD, Palmer BW, et al. Cancer 2018;124(1):192-202 [FREE Full text] [doi: 10.1002/cncr.30987] [Medline: 28926676]
44. Neville AR, Bernstein LJ, Sabiston CM, Jones JM, Trinh L. Feasibility of a remotely-delivered yoga intervention on cognitive function in breast cancer survivors: a mixed-methods study. Front. Cognit 2023;2:12868. [doi: 10.3389/fcogn.2023.1286844]
45. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010;8:18 [FREE Full text] [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
46. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
47. Bentvelzen AC, Crawford JD, Theobald A, Maston K, Slavin MJ, Reppermund S, et al. Validation and normative data for the modified telephone interview for cognitive status: the Sydney Memory and ageing study. J Am Geriatr Soc 2019;67(10):2108-2115. [doi: 10.1111/jgs.16033] [Medline: 31290146]
48. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985;10(3):141-146. [Medline: 4053261]
49. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
50. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
51. Rhodes RE. The evolving understanding of physical activity behavior: a multi-process action control approach. In: Elliot AJ, editor. Advances in Motivation Science. United Kingdom: Academic Press; 2017:171-205.
52. Tabaczynski A, Arbour-Nicitopoulos KP, Rhodes RE, Sabiston CM, Trinh L. Correlates of physical activity participation among individuals diagnosed with Cancer: an application of the multi-process action control framework. Int J Environ Res Public Health 2023;20(5):4345 [FREE Full text] [doi: 10.3390/ijerph20054345] [Medline: 36901355]
53. Trinh L, Rhodes RE, Alibhai SMH, Campbell KL, Langelier DM, Chang E, et al. A randomized controlled trial adding behavioral counseling to supervised physical activity in people living with and beyond cancer (BOOST-UP-): a study protocol for a live remotely-delivered behavior change intervention. BMC Cancer 2025;25(1):847 [FREE Full text] [doi: 10.1186/s12885-025-13904-8] [Medline: 40346460]
54. Rhodes RE. Multi-process action control in physical activity: a primer. Front Psychol 2021;12:797484 [FREE Full text] [doi: 10.3389/fpsyg.2021.797484] [Medline: 34975698]
55. American College of Sports Medicine, Riebe D, Ehrman JK, Liguori G, Magal M. ACSM’s Guidelines for Exercise Testing and Prescription. London: Wolters Kluwer; 2018.
56. BORG GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [doi: 10.1249/00005768-198205000-00012]
57. Hatcher RL, Gillaspy JA. Development and validation of a revised short version of the working alliance inventory. Psychother Res 2006;16(1):12-25. [doi: 10.1080/10503300500352500]
58. Tracey TJ, Kokotovic AM. Factor structure of the Working Alliance Inventory. Psychological assessment: A journal of consulting and clinical psychology 1989;1(3):207-210. [doi: 10.1037/1040-3590.1.3.207]

XSL•FO
RenderX

59. Heaton RK, Akshoomoff N, Tulsky D, Mungas D, Weintraub S, Dikmen S, et al. Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. J Int Neuropsychol Soc 2014;20(6):588-598 [FREE Full text] [doi: 10.1017/S1355617714000241] [Medline: 24960398]
60. Hook JN, Gershon RC. NIH Toolbox for the assessment of neurological and behavioral function (NIHTB). In: Boyle G, Stern Y, Stein D, editors. The SAGE Handbook of Clinical Neuropsychology: Clinical Neuropsychological Assessment and Diagnosis. London: SAGE Publications Ltd; 2023:469-482.
61. Yao C, Bernstein LJ, Rich JB. Executive functioning impairment in women treated with chemotherapy for breast cancer: a systematic review. Breast Cancer Res Treat 2017;166(1):15-28. [doi: 10.1007/s10549-017-4376-4] [Medline: 28707202]
62. Stoet G. PsyToolkit: A software package for programming psychological experiments using Linux. Behav Res Methods 2010;42(4):1096-1104. [doi: 10.3758/brm.42.4.1096]
63. Stoet G. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. Teach Psychol 2016;44(1):24-31 [FREE Full text] [doi: 10.1177/0098628316677643]
64. NIH Toolbox. Remote Administration in NIH Toolbox V2. URL: https://nihtoolbox.zendesk.com/hc/en-us/articles/13434920124820-Remote-Administration-in-NIH-Toolbox-V2 [accessed 2026-01-30]
65. Freedson PS, Melanson E, Sirard J. Calibration of the computer science and applications, Inc. accelerometer. Med Sci Sports Exerc 1998;30(5):777-781. [doi: 10.1097/00005768-199805000-00021] [Medline: 9588623]
66. Tabaczynski A, Bastas D, Whitehorn A, Trinh L. Changes in physical activity and associations with quality of life among a global sample of cancer survivors during the COVID-19 pandemic. J Cancer Surviv 2023;17(4):1191-1201 [FREE Full text] [doi: 10.1007/s11764-021-01156-x] [Medline: 35079964]
67. Trinh L, Kramer AF, Rowland K, Strom DA, Wong JN, McAuley E. A pilot feasibility randomized controlled trial adding behavioral counseling to supervised physical activity in prostate cancer survivors: behavior change in prostate cancer survivors trial (BOOST). J Behav Med 2021;44(2):172-186. [doi: 10.1007/s10865-020-00185-8] [Medline: 32979134]
68. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004;10(2):307-312. [doi: 10.1111/j..2002.384.doc.x] [Medline: 15189396]
69. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]
70. Naumann F, Martin E, Philpott M, Smith C, Groff D, Battaglini C. Can counseling add value to an exercise intervention for improving quality of life in breast cancer survivors? A feasibility study. J Support Oncol 2012;10(5):188-194. [doi: 10.1016/j.suponc.2011.09.004] [Medline: 22169703]
71. Sherman KA, Heard G, Cavanagh KL. Psychological effects and mediators of a group multi-component program for breast cancer survivors. J Behav Med 2010;33(5):378-391. [doi: 10.1007/s10865-010-9265-9] [Medline: 20502954]
72. Campbell KL, Kam JWY, Neil-Sztramko SE, Liu Ambrose T, Handy TC, Lim HJ, et al. Effect of aerobic exercise on cancer-associated cognitive impairment: a proof-of-concept RCT. Psychooncology 2018;27(1):53-60. [doi: 10.1002/pon.4370] [Medline: 28075038]
73. Northey JM, Pumpa KL, Quinlan C, Ikin A, Toohey K, Smee DJ, et al. Cognition in breast cancer survivors: a pilot study of interval and continuous exercise. J Sci Med Sport 2019;22(5):580-585. [doi: 10.1016/j.jsams.2018.11.026] [Medline: 30554923]
74. Gokal K, Munir F, Ahmed S, Kancherla K, Wallis D. Does walking protect against decline in cognitive functioning among breast cancer patients undergoing chemotherapy? Results from a small randomised controlled trial. PLoS One 2018;13(11):e0206874 [FREE Full text] [doi: 10.1371/journal.pone.0206874] [Medline: 30485297]
75. Bender CM, Sereika SM, Gentry AL, Cuglewski C, Duquette J, Grove G, et al. Effects of aerobic exercise on neurocognitive function in postmenopausal women receiving endocrine therapy for breast cancer: The Exercise Program in Cancer and Cognition randomized controlled trial. Psychooncology 2024;33(2):e6298. [doi: 10.1002/pon.6298] [Medline: 38911475]
76. Smith GVH, Myers SA, Fujita RA, Yu C, Campbell KL. Virtually supervised exercise programs for people with cancer. Cancer Nurs 2024;48(6):E397-E407. [doi: 10.1097/ncc.0000000000001353]
77. Groen WG, van Harten WH, Vallance JK. Systematic review and meta-analysis of distance-based physical activity interventions for cancer survivors (2013-2018): we still haven't found what we're looking for. Cancer Treat Rev 2018;69:188-203. [doi: 10.1016/j.ctrv.2018.07.012] [Medline: 30077954]
78. Oyer RA, Hurley P, Boehmer L, Bruinooge SS, Levit K, Barrett N, et al. Increasing racial and ethnic diversity in cancer clinical Trials: an American society of clinical oncology and association of community cancer centers joint research statement. J Clin Oncol 2022;40(19):2163-2171. [doi: 10.1200/jco.22.00754]
79. Myers JS, Koleck TA, Sereika SM, Conley YP, Bender CM. Perceived cognitive function for breast cancer survivors: association of genetic and behaviorally related variables for inflammation. Support Care Cancer 2017;25(8):2475-2484 [FREE Full text] [doi: 10.1007/s00520-017-3654-3] [Medline: 28247126]
80. Akbari PS, Hassan Y, Archibald L, Tajik T, Dunn K, Berris M, et al. Effect of physical activity during chemotherapy on cognitive function in cancer Survivors: a systematic review and meta-analysis. Physiother Can 2023;75(1):12-21 [FREE Full text] [doi: 10.3138/ptc-2021-0032] [Medline: 37250734]

XSL•FO
RenderX

81. Chaddock-Heyman L, Mackenzie MJ, Zuniga K, Cooke GE, Awick E, Roberts S, et al. Higher cardiorespiratory fitness levels are associated with greater hippocampal volume in breast cancer survivors. Front Hum Neurosci 2015;9:465 [FREE Full text] [doi: 10.3389/fnhum.2015.00465] [Medline: 26379528]
82. Jones LW, Haykowsky MJ, Swartz JJ, Douglas PS, Mackey JR. Early breast cancer therapy and cardiovascular injury. J Am Coll Cardiol 2007;50(15):1435-1441 [FREE Full text] [doi: 10.1016/j.jacc.2007.06.037] [Medline: 17919562]
83. Costa DSJ, Fardell JE. Why are objective and perceived cognitive function weakly correlated in patients with cancer? J Clin Oncol 2019;37(14):1154-1158. [doi: 10.1200/jco.18.02363]
84. Kramer AF, Colcombe S. Fitness effects on the cognitive function of older adults: a meta-analytic study-revisited. Perspect Psychol Sci 2018;13(2):213-217. [doi: 10.1177/1745691617707316] [Medline: 29592650]

Abbreviations

ACSM: American College of Sports Medicine
ACTIVATE: Aerobic exercise and CogniTIVe functioning in women with breAsT cancEr
ANCOVA: analyses of covariance
BC: breast cancer
CRCI: cancer-related cognitive impairment
CSEP: Canadian Society of Exercise Physiologists
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
HR: heart rate
ICCTF: International Cognition and Cancer Task Force
MVPA: moderate-to-vigorous intensity physical activity
M-PAC: multi-process action control framework
maSCD: menopause - associated subjective cognitive decline
NIH: National Institutes of Health
PA: physical activity
QEP: Qualified Exercise Professional
QoL: quality of life
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
RPE: rating of perceived exertion
TICS-M: Modified Telephone Interview of Cognitive Status
WAI-SR: Working Alliance Inventory – Short Revised

Edited by M Balcarras; submitted 04.Mar.2025; peer-reviewed by M Knobf, C Curatoli, K Broekhuizen; comments to author 27.Jul.2025; revised version received 23.Oct.2025; accepted 06.Jan.2026; published 23.Feb.2026.

Please cite as:
Trinh L, Cuda N, Sabiston CM, Warner E, Ryan JD, Kramer AF, Ha M, Xiao H, O'Rourke S, McAuley E
The Feasibility and Acceptability of a Remotely Delivered, Combined Exercise Intervention on Cognitive Function in Patients With Breast Cancer Following Chemotherapy: Randomized Controlled Trial
JMIR Cancer 2026;12:e73393
URL: https://cancer.jmir.org/2026/1/e73393
doi:10.2196/73393
PMID:

©Linda Trinh, Natalie Cuda, Catherine M Sabiston, Ellen Warner, Jennifer D Ryan, Arthur F Kramer, Michelle Ha, Hui Xiao, Sarah O'Rourke, Edward McAuley. Originally published in JMIR Cancer (https://cancer.jmir.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Intervention for Electronic Patient-Reported Outcomes in Advanced Cancer: Mixed Methods Study

Christina Sauer[1], PD, PhD; Simeon Sauer[2], Prof Dr; Elena Sophia Doll[3], MSs; Stefanie Zschäbitz[4], MD; Jan Simon Raue[5], PhD; Maik Turni[5], MA; Till Johannes Bugaj[6], MD; Kiriaki Hiller[4], MD; Imad Maatouk[7], MD

[1]University Cancer Center Frankfurt (UCT), University Hospital Frankfurt, Goethe University, Theodor-Stern-Kai 7, Frankfurt, Hesse, Germany

[2]Technische Hochschule Mannheim, Mannheim, Baden-Wurttemberg, Germany

[3]Institute of Medical Psychology, University Hospital Heidelberg, Heidelberg, Baden-Wurttemberg, Germany

[4]Deptartment of Medical Oncology, University Hospital Heidelberg, Heidelberg, Baden-Wurttemberg, Germany

[5]Outcomes4Me Germany GmbH, Berlin, Berlin, Germany

[6]Department of General Internal Medicine, Psychosomatics and Psychotherapy, University Hospital Heidelberg, Heidelberg, Baden-Wurttemberg, Germany

[7]Department of Internal Medicine II, Division of Integrated Psychosomatic Medicine, Universitätsklinikum Würzburg, Wuerzburg, Bavaria, Germany

Corresponding Author:
Christina Sauer, PD, PhD
University Cancer Center Frankfurt (UCT), University Hospital Frankfurt, Goethe University, Theodor-Stern-Kai 7, Frankfurt, Hesse, Germany

Abstract

Background: Digital health interventions are increasingly being integrated into oncological care to support patients in managing treatment-related symptoms and psychological distress. In a randomized controlled pilot trial, we investigated the feasibility and preliminary efficacy of a digital health app (SOFIA) among patients with cancer, including those in palliative care. SOFIA consists of an electronic patient-reported outcome (ePRO) assessment and coaching component. We showed good feasibility and high acceptability of SOFIA in routine clinical care.

Objective: This pilot study aimed to evaluate user experiences and behavior with SOFIA. We applied a mixed methods design, combining a qualitative exploration of patients' experiences and a quantitative analysis of app usage patterns. The integrated findings are intended to inform the refinement and further development of digital health interventions to better address the needs, preferences, and engagement behaviors of this patient population.

Methods: Patients randomly assigned to the intervention group participated in semistructured interviews at 2 time points during the 12-week intervention period: midway through the intervention (at week 6) (T1) and postintervention (at week 12) (T2). Qualitative data were analyzed using content analysis. User data were descriptively analyzed with Python and the Pandas library (version 2.2.3).

Results: Our qualitative analysis of 29 patients revealed benefits regarding the ePRO assessment (empowerment, support, user-friendliness, and facilitation of patient-physician communication), as well as some criticism (inflexible design and insufficient use by physicians). Benefits of the coaching tool included the availability of helpful information and design aspects such as clarity and user-friendliness. Quantitative data from 32 active users of the app showed that 16 (50%) of patients read articles, 28 (87.5%) started journeys, and 15 (46.9%) completed exercises at any point during the study.

Conclusions: The results of this mixed methods study may provide important indications for digital interventions, including ePRO assessment, for patients with cancer. From the patients' perspective, key features include intuitive design, relevant symptom items, reliable reminder functions, the translation of ePROs into clinically actionable information, clear symptom visualizations, seamless integration into electronic health records, and effective physician engagement. Both our qualitative and quantitative data show the importance of therapy-specific content. These results might increase acceptance and usage, and thereby also the clinical benefit, of future digital interventions for patients with cancer.

Trial Registration: German Clinical Trial Register DRKS00021064; https://www.drks.de/search/de/trial/DRKS00021064

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2020-047277

(JMIR Cancer 2026;12:e91416) doi:10.2196/91416

KEYWORDS

cancer; electronic patient-reported outcome; ePRO; digital intervention; eHealth; immunotherapy; immune-checkpoint inhibition

XSL•FO
RenderX

Introduction

Digital health interventions (DHIs) and the recording of electronic patient-reported outcomes (ePROs) have become an important part of psycho-oncological care. The systematic collection of ePROs has demonstrated substantial benefits in oncology, enhancing communication, quality of care, and clinical efficiency. ePROs foster more effective dialogue between patients and healthcare professionals by providing structured, continuous insights into patients' symptoms and well-being [1]. This facilitates more open, targeted, and meaningful communication during clinical encounters [2]. Evidence further indicates that integrating ePROs into oncology care leads to improved outcomes, including higher health-related quality of life (HRQoL) [3], more effective symptom management, and even prolonged survival among patients with cancer [4,5]. The use of ePROs supports a patient-centered approach by enabling treatment decisions that are more closely tailored to the individual needs, preferences, and experiences of patients. Moreover, ePRO systems allow for the early detection of clinical deterioration and deterioration of HRQoL [6]. Finally, regular monitoring enables timely interventions that can prevent complications and reduce emergency room visits [7].

DHIs are increasingly being integrated into oncological care to support patients in managing treatment-related symptoms and psychological distress. Previous qualitative research with patients with breast cancer has shown that such interventions are perceived as acceptable and hold strong potential to reduce anxiety, address gaps in care, and provide reassurance throughout the treatment process [8]. Similarly, in the study by Krakowczyk et al [9], patients with various cancer entities evaluated the usefulness and usability of *Make It Training*, a modular digital mental health intervention designed to reduce psychological distress and improve coping and HRQoL. Patients described the intervention as user-friendly and supportive in developing functional coping strategies, highlighting its potential for integration into daily life and routine oncological care. However, to date, evidence on the acceptability and perceived usefulness of such DHIs [10] among palliative patients remains scarce, although this group faces particularly severe physical and psychological challenges due to the advanced stage of their illness.

In the SOFIA pilot trial, we developed and tested a mobile app for patients undergoing immunotherapy with immune checkpoint inhibitors (ICIs) within a randomized controlled design. The SOFIA app comprised 2 components: *SOFIA Monitoring*, which involved the regular assessment of ePROs, and *SOFIA Coaching*, which provided tailored information related to treatment and illness management. In a previous publication, we demonstrated the high feasibility and acceptance of the app in clinical routine, with an adherence rate of 91% for completing ePRO assessments twice weekly over a 3-month intervention period. Patients in the intervention group (IG) reported significantly lower levels of depression and distress, as well as improved HRQoL, compared to those receiving treatment as usual [11].

Several studies have qualitatively investigated expectations and experiences with DHIs in patients with cancer. In a recent meta-analysis, Zhang et al [12] synthesized 12 qualitative studies on user experiences among older patients with cancer, identifying 3 overarching themes: (1) positive experiences, (2) barriers to use, and (3) user expectations. The included studies varied in terms of app content (eg, ePRO assessment, physiological tracking, or combined DHIs such as ONKOkompass) and also addressed expectations and needs. The review provides an important overview of key themes and topics. However, to compare specific benefits and barriers, it is necessary to investigate these aspects separately for each component (ePRO assessment, coaching tools) in patients who have actually used the app.

Similarly, a recent systematic review by Lai-Kwon et al [10] reported that ePRO-based systems for monitoring immune-related adverse events are generally feasible and acceptable in patients receiving immunotherapy. However, the included studies predominantly focused on symptom monitoring, with limited or no systematic evaluation of additional intervention components such as coaching modules.

Evidence on the usefulness of ePRO-based and combined eHealth interventions in patients receiving ICI therapy remains limited. Tolstrup et al [13] investigated the usefulness of an ePRO assessment in 14 patients with melanoma receiving ICI therapy and found that patients perceived the tool as easy to access and meaningful to complete. In addition, Glaser et al [14] evaluated the CAPABLE system, an eHealth intervention including symptom management and coaching features in patients undergoing ICI therapy. Their mixed methods study reported high overall perceived usefulness, with more than 75% of participants rating the system positively. However, the evaluation was conducted at the system level and did not distinguish between experiences with the ePRO and coaching components.

Overall, to date, no study has included a mixed sample of patients with cancer undergoing ICI therapy and quantitatively investigated their experiences with ePRO assessment and coaching components separately. Furthermore, no study has combined qualitative data with user log data, which provide important insights into actual usage behavior.

This study aimed to investigate acceptability and patient engagement with the app, operationalized as (1) qualitative patient perspectives on using the SOFIA app, assessed separately for the ePRO assessment and coaching components, and (2) quantitative usage behavior. These insights are intended to inform the future development and optimization of DHIs for patients with cancer, including those in palliative care and undergoing ICI therapy.

Methods

Study Design

This study was embedded within a pilot study and used a convergent parallel mixed methods approach [15], combining (1) qualitative exploration of patients' experiences of app use and (2) quantitative analysis of user behavior. The underlying

XSL•FO
RenderX

phenomenon of interest was patient experience with both components of the SOFIA app, conceptualized as the interplay between patients' subjective experiences and their observable usage behavior. The mixed methods design was chosen to capture both dimensions and to better understand the mechanisms shaping engagement.

The study was a prospective single-center, 2-arm randomized controlled external pilot trial. The study protocol and details of the study methods are described elsewhere [11,16].

To estimate the key parameters for the main trial, a total of 70 participants (35 per group) were required, based on the simulation-based approach for determining sample size in external pilot trials with continuous outcome variables [17].

Intervention

Participants in the IG received access to the SOFIA app (for a detailed description, [11]). SOFIA was integrated into the Mika app, a digital therapeutic providing individualized supportive care and aiming to reduce psychological distress in people with cancer. The app entails different components, including symptom monitoring and coaching modules (for a detailed description, [11]).

Monitoring: SOFIA Monitoring comprised 11 physical and up to 9 mental ePROs, assessed twice weekly with email reminders. Physical symptoms were selected by experienced medical oncologists based on European Society for Medical Oncology guidelines [18] and included weakness, diarrhea, melena, dry cough, dyspnea, reduced urinary output, joint and muscle pain, skin toxicity, fever, and jaundice.

Coaching: The coaching component included modules such as Discover and Journeys. The module "Discover" is the content library of the Mika app, consisting of articles and videos with easily understandable and evidence-based information on several cancer-related topics (eg, cancer and treatment types, nutrition in cancer; Table S1 in Multimedia Appendix 1), written and peer-reviewed by medical experts. An integrated statistical content recommendation engine delivers individual content recommendations to the user based on age, gender, cancer type, therapy type, and reported symptoms. By considering the symptoms the patient has reported in the symptom monitoring module, as well as patterns in their reading behavior, the algorithm predicts the likelihood of the user engaging with specific content and, based on the estimated read probability, presents the respective content to the user.

"Journeys" are the resource-activating psycho-oncological training courses of the Mika app that aim to help patients cope with cancer-associated distress in different areas (eg, living with immunotherapy; Table S1 in Multimedia Appendix 1). There are 13 journeys, each of which entails 3 subsequent stages, which in turn include 3 exercises. Exercises can be allocated to 1 of 4 categories: knowledge, mindfulness, relaxation, or physical exercise.

Study Procedures and Analyses

Qualitative Approach

Reporting of the qualitative part adheres to the COREQ (Consolidated Criteria for Reporting Qualitative Research) [19] (Checklist 1). Patients in the IG participated in semistructured interviews at 2 time points during the 12-week intervention period: midway through (at week 6) (T1) and postintervention (T2). The interview guide is available in Multimedia Appendix 2. Questions were developed based on expert discussions and experiences from previous studies, with the aim of gathering information on experienced benefits and suggestions for improving SOFIA Monitoring and Coaching. The 1:1 interviews were conducted in German by 3 female study assistants who were master's-level psychology students. Owing to their psychology background, all had well-developed communication skills and received additional briefing and supervision from the study coordinator (CS) prior to data collection. To minimize the burden on participants, the interviews were conducted by telephone. Some participants were already acquainted with their individual interviewer through the recruitment contact. At the beginning of each interview, the study assistants introduced themselves as study assistants in the SOFIA project; no other interviewer characteristics were shared. The interviews lasted about 10 minutes, depending on how detailed the answers were. Due to the brevity of the questionnaire and the limited selection of questions, no pilot testing or audio recording was conducted. The study assistants wrote down the statements by hand (either verbatim or summarized in bullet points) and then transferred them anonymized into an Excel spreadsheet. Additional field notes were not made. Interview notes were not returned to participants for comment and/or correction.

Statements were analyzed using structured qualitative content analysis, combining deductive and inductive approaches [20]. While the interview guide provided an initial set of categories (deductive), these categories were refined inductively based on the participants' responses, with subcategories and hierarchical levels being developed during the course of the analysis. Initially, meaningful content aspects within each statement were identified as coding units and assigned to the overarching categories from the interview guide by 2 independent raters. As individual statements could contain multiple content aspects, more than 1 category assignment per statement was possible. The allocation was subsequently discussed within the research team. Disagreements were discussed until consensus was reached. In cases where differing opinions persisted, the senior researcher and study coordinator (CS) made the final decision. This procedure for resolving disagreements was applied consistently across all stages of the coding process. In the next step, common groups (subcategories) were formed within the overarching categories based on the assigned content aspects. Where appropriate, these subcategories were further consolidated into generic categories through additional team discussions. The final category structure emerged through an iterative, consensus-based coding process involving multiple rounds of analysis and regular team discussions. During this process, categories and subcategories were continuously reviewed, refined, and reorganized where substantial overlap

XSL•FO
RenderX

was identified. In our overview, we have summarized all statements that were mentioned more than twice.

Quantitative Approach

SOFIA app usage was analyzed using log data of patients' in-app activities, providing detailed insights into the types of content and exercises with which patients were most engaged. After participants in the IG received a study access code and consented to the app's privacy terms, their pseudonymized in-app activities were automatically recorded, allowing the evaluation of engagement with different coaching features of the app. For the current analyses, the app modules "Discover" and "Journeys" were of interest.

The raw log files were preprocessed by the app provider (Fosanis GmbH), resulting in several spreadsheets containing general user data and logs of articles read in the "Discover" module, as well as logs of journeys started or completed and exercises performed. The spreadsheets included the following information: unique user ID; name of the article, journey, or exercise with which the user was engaged; period during which the engagement took place (weeks 0 - 12: intervention phase; weeks 12+: follow-up phase); and the duration (in seconds) of the engagement.

The spreadsheets were then processed further using a customized Python script and the Pandas library (version 2.2.3) for data processing. Sessions of zero duration were removed, and the time unit was converted from seconds to hours. Finally, the descriptive statistics of user engagement presented in the *Results* section were produced.

Ethical Considerations

The Ethics Committee of Heidelberg University Hospital (reference S-581/2018) approved the trial protocol, and the study was registered in the German Clinical Trial Register (reference DRKS00021064). All participants were informed about the study procedures and signed informed consent before being included in the study.

Results

Participant Characteristics

Of the 34 patients in the IG, 2 patients left the study directly after randomization. For the remaining 32 participants, user data were available. Regarding the availability of interview data, 29 patients participated in the midway assessment (T1), while 23 patients participated in the postintervention assessment (T2). One patient discontinued participation due to dissatisfaction with the application, 2 patients died during the study period, 1 patient was physically unable to participate in the interview, and 2 patients could not be reached by phone.

Patient characteristics are presented in Table 1. The study flow chart can be found in Multimedia Appendix 3. For a detailed explanation of the study flow [11].

XSL•FO
RenderX

Table . Patient characteristics (n=29).

Characteristics	Values
Age (y), mean (SD)	61.2 (11.4)
Time since diagnosis (mo), mean (SD)	25.3 (42.6)
Male sex, n (%)	19 (65.5)
Tumor entity, n (%)	
Urological	11 (37.9)
Gastrointestinal	7 (24.1)
Hepatobillary	5 (17.2)
Head and neck	1 (3.4)
Lung	1 (3.4)
Gynecological	2 (6.9)
Squamous cell, not other specified	2 (6.9)
Metastasis (yes), n (%)	27 (93.1)
ECOG[a], n (%)	
0	13 (44.8)
1	16 (55.2)
Immunotherapy type of regimen, n (%)	
ICI[b] (monotherapy)	13 (44.8)
ICI + ICI	8 (27.6)
ICI + TKI[c]	6 (20.7)
ICI + CHX[d]	2 (6.9)
Previous cancer treatments (yes), n (%)	25 (86.2)
Previous treatments, n (%)	
Radiation	2 (6.9)
Surgery	6 (20.7)
Chemotherapy	3 (10.3)
Immunotherapy	5 (17.2)
>1 previous cancer treatments	16 (55.2)
Psychiatric side diagnosis, n (%)	1 (3.4)

[a]ECOG: Eastern Cooperative Oncology Group Performance Status

[b]ICI: immune checkpoint inhibitor.

[c]TKI: tyrosine kinase inhibitor.

[d]CHX: chemotherapy.

In our IG sample, 6 patients received steroids due to adverse events after the start of immunotherapy. Five patients had brain metastases, resulting in 1 patient undergoing surgery, 3 receiving radiotherapy, and 1 death. No patients had meningeal metastases. Five patients had thyroid dysfunction: 3 alone and 2 combined with other side effects.

Qualitative Results

For both SOFIA Monitoring and SOFIA Coaching, we identified two main categories: (1) perceived benefits and (2) criticism or recommendations. In the following, we describe the corresponding generic categories and subcategories for each and present illustrative quotations from the interview notes. The quotations were translated into English by CS. The numbering is based on the order in which the statements (eg, #1) and their documentation appear in the Excel spreadsheet (eg, *a*=first data spreadsheet), as well as the measurement time (T1 vs T2). Figures 1 and 2 summarize the results of the qualitative analysis by presenting the coding structure with main categories and subcategories.

XSL•FO
RenderX

Figure 1. Main categories, generic categories, and subthemes of SOFIA Monitoring. ICI: immune checkpoint inhibitor.

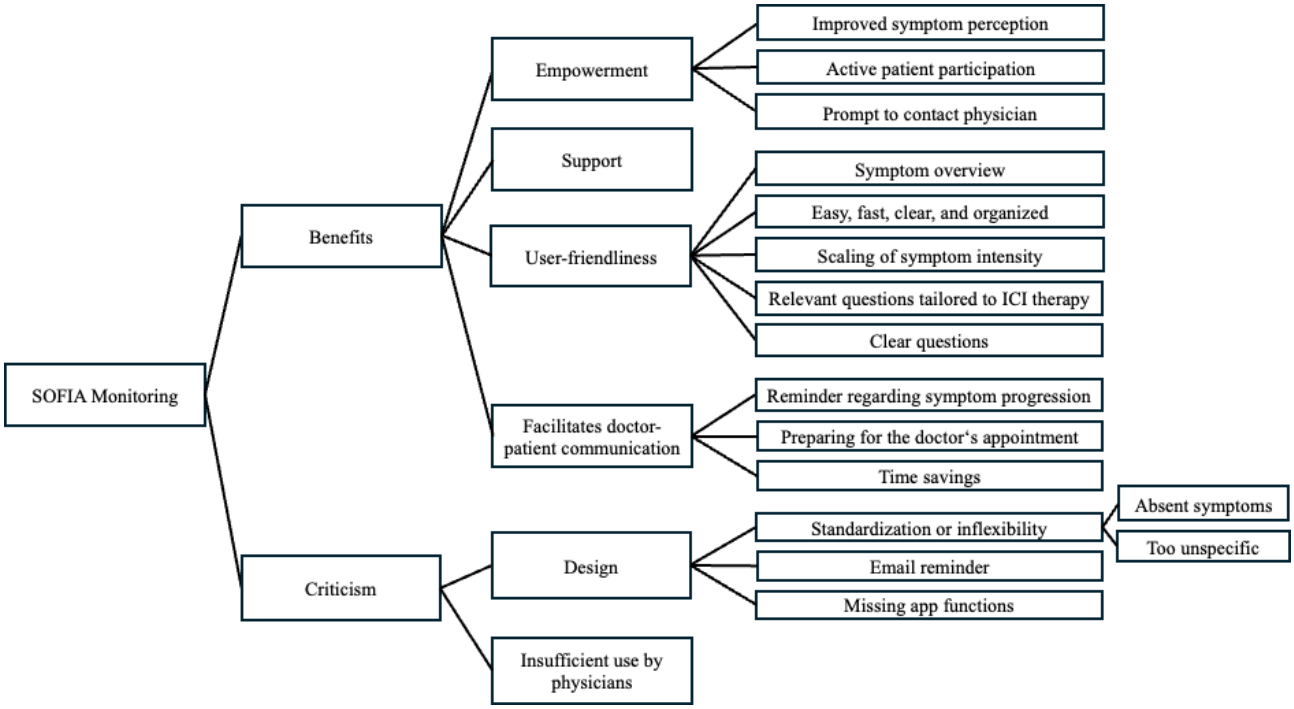
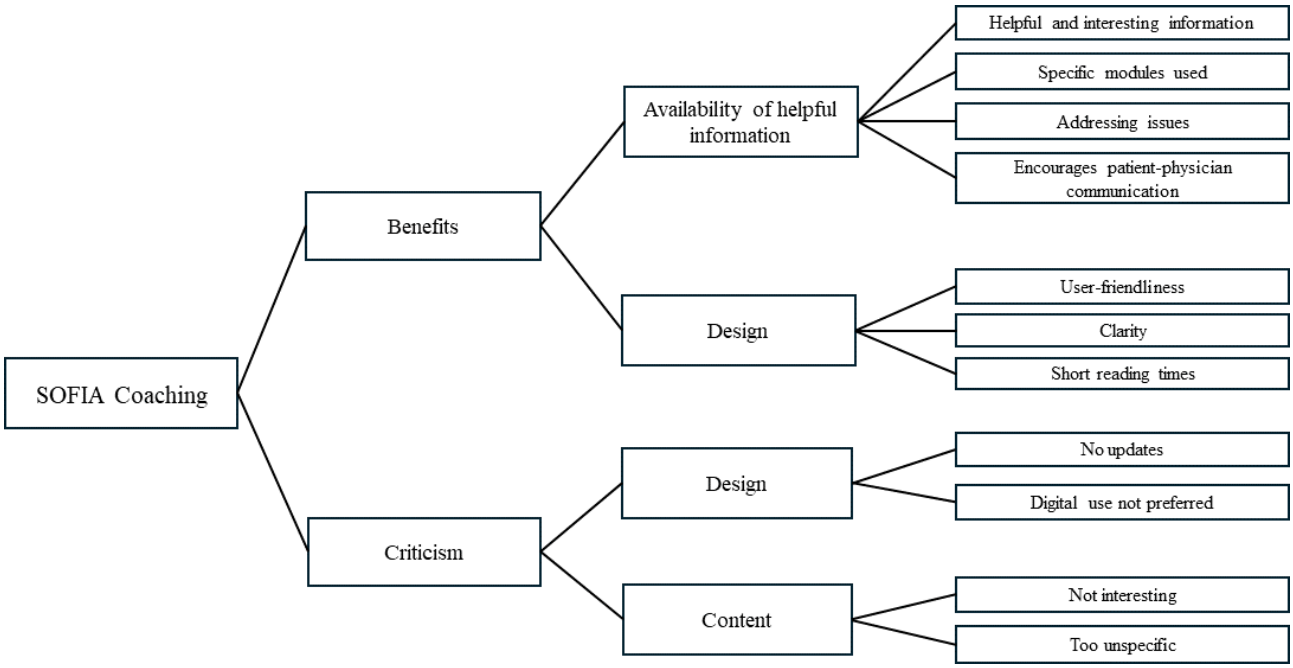


Figure 2. Main categories, generic categories, and subthemes of SOFIA Coaching.



SOFIA Monitoring

Main Category 1: Benefits of SOFIA Monitoring (Assessment of ePROs)

Generic Category: Empowerment

Patients reported at T1 and T2 that the ePRO assessment improved their symptom perception (subcategory: improved symptom perception), strengthened their patient participation (subcategory: active patient participation), and the prompt to contact the physician (subcategory: prompt to contact the physician) encouraged them to become active or more aware of severe symptoms. These subcategories were combined to create the generic category called "empowerment," since empowerment is the process by which patients gain the knowledge, skills, confidence, and control necessary to make informed decisions about their health and to play an active role in their own care [21].

Subcategory: Improved Symptom Perception

Patients reported that the ePRO assessment actively encouraged them to recognize their symptoms, including psychological ones, and that this helped them classify the symptoms:

XSL•FO
RenderX

This is very helpful for assessing yourself better. Example: When you subjectively feel unwell and then objectively evaluate that feeling in the app. [#1a, T1]

Also good for recording changes [was it just a one-off?]. If you just wrote it down, you wouldn't have such a clear overview. [#3a, T1]

You become aware of the questions and symptoms. Otherwise, I would not concern myself with them, but rather suppress them, especially when they concern depression. [#29a, T1]

This helped me realize that I am doing well. [#1a, T2]

Paying more attention to side effects ['listening to my body more closely,' 'checking more often for certain symptoms,' being more attentive to my body]. [#38a, T2]

Deciding whether to go to the doctor or not. [#40a, T2]

Subcategory: Active Patient Participation

Patients reported that the ePRO assessment enabled them to play an active role in managing their condition:

Being actively involved in documenting one's own recovery and being able to observe progress and see where the journey is headed. [#25a, T1]

Being part of something bigger and, therefore, being able to help future patients. [#26a, T1]

Subcategory: Prompt to Contact the Physician

Patients found the warning notice, along with the contact details, helpful in determining when to contact a physician:

[...] A notification to contact a doctor, because not everyone does that ['there's time/it's not urgent'], but when the notification keeps appearing, you decide to take action. [#4a, T1]

You can see when something is above range and are informed to contact your doctor in that case, or that everything is OK. [#45a, T1]

Generic Category: Support

Patients reported that they had the feeling that somebody was interested in their symptoms and that they were not alone with their fears. The patients felt reassured knowing what the typical symptoms were, which gave them a sense of security.

I don't feel left alone with my fears; I feel like someone cares about me. [#52a, T1]

It's good that someone is asking about my symptoms twice a week, not just every 3 weeks. [#2a, T2]

Generic Category: User-Friendly Application

The patients appreciated the clarity of the presentation, the scaling of symptom intensity, and the relevant questions tailored to immunotherapy. They found the application easy and quick to use. See Table 2 for subcategories and examples.

Table . Subcategories of user-friendliness.

Subcategory	Statement (example)
Symptom overview	#5a (T1): You can view the scale yourself over a period of 30 days to get a good overview.
Easy, fast, clear, and organized handling	#42a (T2): Questions clear, given answer options good. #10a (T1): […] fast, clear questions.
Scaling of symptom intensity	#10a (T1): The range of 1 - 10, because it can fluctuate daily. #23a (T1): The alignment with immunotherapy is satisfactory.
Relevant questions for ICI-therapy[a]	#12a (T1): You don't have to figure out what is relevant yourself.
Clear questions	#73-74a (T2): Very easy, works well. #43a (T2): The questions are clearly formulated, and the predefined response options are well constructed.

[a]ICI: immune checkpoint inhibitor.

Generic Category: Facilitation of Patient-Physician Communication

Some patients reported that the ePRO helped them remember symptoms for the physician-patient consultation and to report them during the consultation (subcategory: reminder regarding symptom course). Patients also used it to prepare for consultations with their physicians (subcategory: preparation of questions for the medical consultation). They also saw it as a time saver, as it meant that physicians were already able to obtain a detailed overview, facilitating communication and the identification of relevant symptoms (subcategory: time savings). Examples are depicted in Table 3.

XSL•FO
RenderX

Table . Subcategories of facilitation of patient-physician communication.

Subcategory	Statement (example)
Reminder regarding symptom course	#1c (T2): Yes, in the sense that you don't forget anything and have accurate information (even if something only happened once, which you would otherwise have forgotten to mention).
Preparation of questions for the medical consultation	#10c (T1): Yes, because you are more aware of what happened and what you want to say.
Time savings	#3c (T2): Yes, because the doctor already had my medical history, I didn't have to explain everything again and could focus on what was important.

Main Category 2: Criticism or Recommendations of SOFIA Monitoring

Generic Category: Design

Subcategory: Standardization or Inflexibility

The main criticism from patients concerned the standardization of the ePRO assessment in the app. Some patients reported missing symptoms (eg, hypertension or hypotension, weight gain/loss) and expressed a desire for greater specificity of items or a free text field for entering symptoms.

> *Option to enter an individual symptom.* [#2e, T1]
>
> *Precise definition, what kind of pain, more specific.* [#26e, T1]
>
> *Difficult 'in recent days' because I always have muscle pain [unrelated to cancer].* [#29d, T1]

Subcategory: Email Reminder

Patients reported that they were unsatisfied with the email reminder function. They missed a direct reminder of the ePRO assessment via the app because, in some cases, there was no connection between the email reminder and the app (eg, #21d(T2): "Difficulties remembering to enter Mondays and Thursdays (as there is no link between mobile phone), did not make good use of the reminder function in the app.").

Subcategory: Missing App-Function

Some patients reported that they missed some functions in the SOFIA Monitoring component, for example,

> *You cannot enter anything if there are changes in the therapy.* [#13f, T1]
>
> *Doctors should be more proactive in approaching patients; if abnormal values are detected, it should not be up to the patient [who is already under stress] to report them.* [#15f, T1]

Generic Category: No Usage of Physicians

At T1, 62% (18/29) of patients reported that physicians did not use their input during the consultation; 17% (5/29) indicated that physicians used it partially, and 7% (2/29) reported that their input was used; 14% (4/29) were unable to judge that. At T2 (n=21 answers were available for this question), 14 (67%) patients reported no use of their input, 3 (14%) reported partial use, and 4 (19%) reported use. Patients expressed a desire for more feedback and a better integration of the ePRO assessment into their treatment.

SOFIA Coaching

Main Category 3: Benefits of SOFIA Coaching

Generic Category: Availability of Helpful Information

Patients appreciated the availability of helpful and interesting information about their illness and treatment (subcategory: helpful and interesting information), as well as specific modules such as thematic journeys, the contact module, and a video interview (with an immunotherapy expert from the National Center for Tumor Diseases Heidelberg) (subcategory: specific modules used). Patients also valued being made aware of relevant issues (subcategory: addressing issues), which in turn facilitated communication with their physician (subcategory: encourages patient-physician communication). Examples are depicted in Table 4.

XSL•FO
RenderX

Table . Subcategories of availability of helpful information.

Subcategory	Statement (example)
Helpful and interesting information	#2f (T1): One can glean useful tips from it.
	#7f (T1): I liked it. Explanation of immunotherapy, themed trips, finding your way to cope with cancer-Positive: Content seemed very diverse, eg, something for all age groups (children/pensioners)
Specific modules used	#33f (T1): Thematic journeys, approximately every 2 days for half an hour to an hour, are helpful.
	#13f (T1): Video interview was good.
	#3f (T1): It's good to know who to turn to (contact persons just a click away), in good hands. (contact module)
Addressing issues	#1f (T1): It's good that it's included in the app, because otherwise you might not bother with it. Depending on what's happening that day, you can take another look at it
Encourages patient-physician communication	#40f (T1): I looked at thematic journeys, which raised questions that I then asked during my consultation with the doctor

Generic Category: Design

We identified 3 subcategories (user-friendliness, clarity, and short reading times) in the Design category. Patients appreciated the user-friendliness of SOFIA Coaching. They highlighted its ease of use, clarity, and short reading time (eg, #15f (T1): "The reading time is very helpful because my ability to concentrate is limited by my illness").

Main Category 4: Criticism or Recommendations for SOFIA Coaching

Generic Category: Design

Patients criticized missing updates (subcategory: missing updates). Some patients reported they generally preferred an analog information (subcategory: digital use not preferred). Examples are depicted in Table 5.

Table . Generic categories of improvable design aspects.

Subcategory	Statement (example)
Missing updates	#13f (T1): There are no updates/news that would encourage one to revisit the site.
	#15f (T1) Push notifications for new (would be helpful)
Digital use not preferred	#21g (T1): Does not use Coaching because he does not spend much time in front of the screen and is not a "computer freak."
	#45g (T2): Technical side: He has to print out anything he finds interesting. There is no offline/paper function for marking things or making notes. He would prefer a kind of "wiki" to better manage the information (information on study results). (…)

Generic Category: Content Not Interesting or Unspecific

Few patients reported they experienced the content and wished partly more detailed information about specific topics (eg, cancer and genetics).

> *Individual to each person; article very general, much of it did not apply.* [#52g, T2]
>
> *I would like more information about the specific type of cancer.* [#44f, T2)]

Quantitative Results

Based on the log data of the 32 active users of the coaching features of the SOFIA app, we found that 16 (50%) of them had read at least 1 article from the "Discover" module. The median number of articles read per user in this group was 7.5. The corresponding mean value of 14.8 (SD 19.9) was distorted by 2 outliers (52 and 72 articles). As seen from the more detailed analysis in Table 6, more articles were read in the intervention period of the study than in the follow-up period.

Table . Analysis of articles read per user in the "Discover" module of the SOFIA app (N=32).

Time period	Users with ≥1 articles read, n (%)	Number of articles read in this group, median (IQR)	Number of articles read in this group, mean (SD)
Any period	16 (50)	7.5 (12.5)	14.8 (19.9)
Intervention period only	16 (50)	4 (10)	10.6 (14.6)
Follow-up period only	7 (21.9)	4 (9)	9.6 (11.6)

XSL•FO
RenderX

The app organizes articles in different categories. The most frequently read category was "cancer therapy" (*k*=47), closely followed by "COVID and cancer" (*k*=38), "relieving symptoms" (*k*=37), "healthy lifestyle" (*k*=33), and "nutrition and cancer" (*k*=28). All data are depicted in figures in Multimedia Appendix 4 (pie chart, bar chart, and box plot).

The log data of the "Journeys" module of the SOFIA app revealed that 87.5% (28/32) users started at least 1 journey; 25% (8/32) of them completed at least 1 journey. Table 7 shows that most of the engagement in journeys took place in the intervention period.

Table . Analysis of journeys started or completed in the "Journeys" module of the SOFIA app (N=32).

Time period	Users with ≥1 journey started, n (%)	Number of journeys started in this group, median (IQR)/mean (SD)	Users with ≥1 journey completed, n (%)	Number of journeys completed in this group, median (IQR)/mean (SD)
Any period	28 (87.5)	1 (1.25)/1.96 (1.62)	8 (25)	1 (2.00)/1.75 (1.04)
Intervention period only	26 (81.3)	1 (1.75)/1.96 (1.46)	6 (18.8)	1.5 (1.75)/1.83 (0.98)
Follow-up period only	4 (12.5)	1 (0.00)/1.00 (0.00)	3 (9.4)	1 (0.00)/1.00 (0.00)

Table 8 shows the popularity of journeys among users. By far, the most popular journey was "Living with immuno-therapy." Figures in Multimedia Appendix 5 show the titles and frequencies of journeys completed.

Table . Frequency of journeys started.

	Users, n
Living with immunotherapy	24
Find your way	8
Relieve exhaustion	6
Gain control	5
Activate power source	4
Reduce stress	4
Nutrition and cancer	2
Managing emotions	1
Yoga and cancer	1

In the SOFIA app, journeys include exercises. Fifteen out of 32 (46.9%) users were engaged in at least 1 such exercise. In this group, an average of 9.1 (SD 8.8) exercises were started (median 9, IQR 6.5). Figures in Multimedia Appendix 6 show the most frequently started exercises. Multimedia Appendix 7 includes a joint display table linking the main qualitative themes of SOFIA Coaching to their respective quantitative usage patterns.

Discussion

Principal Results

DHIs are an integral part of oncological care. The evidence for this is very clear. Many efficacy studies do not provide accurate user information. The results of our convergent parallel mixed methods approach showed that, qualitatively, patients reported feelings of empowerment and support, appreciated the user-friendly design, and reported perceived improvement of patient-physician communication (ePRO assessment) as well as access to helpful and validated information (coaching component). Simultaneously, they wished for more flexibility in the ePRO assessments and more integration of the ePRO assessment into the clinical routine, as well as more specific data. Regarding the coaching component, some participants wished for more updates or did not prefer digital use. Quantitative user data indicated that over 50% (16/32) of patients read articles, engaged with journeys, and performed exercises, with immunotherapy-relevant articles, the management of side effects, and living with immunotherapy being of particular interest.

Implications for ePRO Assessments

The results of our main study revealed high acceptability and feasibility of the ePRO assessment in patients undergoing ICI therapy, with an adherence rate of 91% [11]. We conducted qualitative interviews to gain deeper insights into user perspectives and needs. The findings of our accompanying qualitative study corroborate those of a previous investigation involving 186 patients with various cancer entities undergoing diverse cancer therapies, as well as 55 clinicians [22]. That study examined patient and clinician perspectives on the use of a DHI for collecting ePROs. Three overarching themes were identified for both stakeholder groups: (1) acceptability and functionality (relevance of symptom items, usability, reminders, perceived health status); (2) impact on clinical care (information and resource provision, contribution to research, self-monitoring, reassurance, connection to the hospital, and support for guided decision-making); (3) personal value of using eRAPID—the tool used for ePRO assessment (clinician engagement with eRAPID, facilitation of consultations, and support for medication or treatment adjustments) [22]. Furthermore, the

XSL•FO
RenderX

results are in line with the recently published systematic review of qualitative data [12] identifying "improved self-management," "safe and accessible support," and "better access to health information" as key benefits, and "inadequate content design," "technical difficulties," and "negative user perceptions and behaviors" as key barriers. Consistent with our generic categories—user-friendly application (acceptability and functionality), facilitation of patient-physician communication (impact on clinical care), and empowerment, support, or availability of helpful and interesting information (personal value)—these findings provide several implications for the development of future DHIs including ePRO assessments in oncology. First, ePRO systems should prioritize high acceptability and usability by ensuring intuitive design, relevant symptom items, and reliable reminder functions (eg, directly via the app) [2]. Second, features that translate PROs into clinically actionable information—such as automated alerts or structured summaries—are essential to support timely decision-making. Third, future tools should actively facilitate patient-physician communication through clear symptom visualizations and seamless integration into electronic health records. In this context, ePROs could be particularly useful for gathering information about how patients tolerate therapies such as tyrosine kinase inhibitors. These therapies are often used in combination with ICIs and are frequently associated with chronic side effects that are difficult to assess during routine consultations, but which can have a significant impact on HRQoL [23].

Moreover, ePRO systems should promote patient empowerment by enabling self-monitoring, offering feedback on symptom trajectories, and providing access to supportive resources. Successful implementation further requires strong clinician engagement and incorporation into routine workflows. The results of our previously published study [11] revealed that 65% (n=28) of physicians (who treated patients in the IG) never or seldom used the graph for monitoring patients' care—even though over 50% of physicians evaluated ePROs as useful and saw significant benefits in managing treatment with ICI (eg, in preparing for consultations or detecting side effects). These reports are in line with the results of the answers of our patients, as over 60% of patients reported no usage of their ePRO assessments in the clinical routine, but patients wished that their reports would be better integrated into the patient-physician consultation. It is important to note that these implications are derived from qualitative perceptions and descriptive usage patterns and should not be interpreted as evidence of causal mechanisms at an individual level.

In future studies, the integration of ePRO data into clinical workflows should be systematically evaluated, including the role of physician training. In line with studies showing that the assessment of ePROs and screenings should have consequences and aftercare [24], this is important for patients' adherence to reporting ePROs regularly. It would then be particularly interesting to gain a deeper understanding of patients' and physicians' perspectives when ePROs are more integrated into clinical encounters.

Finally, offering flexible assessments while simultaneously enabling standardized data capture for research and quality improvement will be crucial for maximizing the clinical utility of next-generation ePRO solutions.

Furthermore, the factor of personal value (empowerment, support, availability of helpful information) might have led to improved HRQoL and reduced psychological burden in patients in the IG compared with the control group in our sample [11]. However, this interpretation should be viewed with caution, as our small sample size did not allow us to demonstrate the mediating effects of these factors.

Implication of Coaching Components and Integration of Quantitative Findings

Patients reported benefiting from the availability of relevant and helpful information, as well as from the user-friendly and clear design. This is consistent with previous studies showing that access to helpful information and safe, accessible support is a key factor in eHealth interventions for patients with cancer [12,22]. These findings underline the importance of providing reliable and accessible information. In particular, patients benefited from the journeys and topics addressing side effects and preparation for physician consultations. These findings are further supported by our user-data analyses, which show that patients actively engaged with the app's content. During the observation period, 81% (26/32) started a Journey module, and 50% (16/32) accessed written articles. In the follow-up period, engagement declined, but 22% (7/32) read articles, and 12% (4/32) started journeys. The frequent use of Journey modules reflects the Journey format's strong appeal. Content analyses indicated that users were particularly interested in topics related to immunotherapy, current developments, and the management of treatment-related side effects, which was underlined by our qualitative data (ie, availability of helpful information; dealing with issues). These patterns are consistent with recent studies demonstrating that patients with cancer seek timely, therapy-specific information and guidance [25]. Future studies should take this into account when developing digital contents for patients with cancer.

However, 50% (16/32) of users did not engage with the SOFIA app's articles. Criticisms identified in the qualitative evaluation of the SOFIA Coaching program may explain the limited use (eg, low preference for digital formats, lack of updates, or articles perceived as too general or insufficiently relevant). Future studies should provide cancer-entity-specific information, ensure regular content updates, and consider that DHIs alone may not reach all patients. To avoid disadvantaging individuals who are less digitally inclined, alternative, nondigital information formats, which are often provided by cancer-specific nongovernmental funding and nonprofit-making organizations, should continue to be offered in parallel.

Limitations

Our study provides important insights into the assessment of ePROs and the use of DHIs in patients with cancer and represents an important addition to our feasibility study. In particular, the mixed methods approach, combining qualitative and quantitative data, represents a major strength of this study. However, several limitations have to be discussed. We only assessed the patients' perspective in our qualitative approach.

XSL•FO
RenderX

Although we quantitatively assessed physicians' views on the benefits, usability, and use of the ePRO assessment in a sample of 33 physicians involved in treating patients in the IG [11], a complementary qualitative study would likely have provided deeper insights into perceived benefits and barriers. Due to anonymization processes, it was not possible to match the user data with other PROs and the qualitative data. Furthermore, we were unable to provide a detailed analysis of user behavior, such as engagement over time, weekly trajectories, dropout curves, and patterns of sustained vs early disengagement. Nevertheless, the data provide important insights in the usage of DHIs.

The sample for the qualitative analysis was constrained by the number of participants in the intervention, as determined by the sample size calculation for the primary outcomes of the main trial. Therefore, data saturation was not explicitly addressed in the qualitative part. However, recurring themes were identified across interviews, and our sample is larger than in comparable previous studies [12]. The concept of saturation has also been critiqued, as it assumes that no additional insights can emerge beyond a certain point [26]. Further, since the interviewers had prior contact with the participants during the recruitment phase of the study, a potential influence of social desirability has to be taken into account. However, we consider this influence to be minor, as the first contact was brief and participants also raised critical points during the phone interviews.

Conclusions

This mixed methods study revealed the benefits and criticisms of ePRO assessments and digital coaching tools for patients with cancer under ICI-therapy. Results indicated that a user-friendly app, perceived facilitation of patient-physician communication, and several dimensions of personal value (eg, empowerment, support, or availability of helpful and interesting information) (personal values) may represent pivotal aspects for patients with cancer using DHIs. Quantitative data underline the importance of therapy-specific contents. In sum, we derived several indications for future DHIs in patients with cancer, which may contribute to the improvement of acceptance and usage and thereby further support the clinical benefit.

Acknowledgments

The authors thank all patients who participated in this study. They also thank Andreas Siebert for his contribution to the qualitative analysis. Parts of the manuscript were proofread by ChatGPT (GPT-5; OpenAI) to increase readability. Following the use of this tool, the authors examined and revised the content as required, assuming complete responsibility for the publication's material.

Funding

The study was financially supported by the National Center for Tumor Diseases Elevator Pitch 2017.

Data Availability

Data are available upon reasonable request from the corresponding author.

Authors' Contributions

Conceptualization: CS, IM
Data curation: CS, SS, MT
Formal analysis: CS, SS, ESD, MT, KH
Funding acquisition: CS, IM
Investigation: CS, ESD, SZ
Methodology: CS, SS, ESD, IM
Project administration: CS, IM
Resources: CS, IM, JSR, TJB
Software: SS, MT
Supervision: CS, IM
Validation: CS, SS
Visualization: CS, SS
Writing – original draft: CS
Writing – review & editing: SS, ESD, SZ, JSR, MT, TJB, KH, IM

Conflicts of Interest

JSR and MT are employed by Outcomes4Me Germany GmbH, which develops and provides the application (Mika) investigated in this publication. CS, SS, ESD, SZ, TJB, and IM have no conflicts of interest to declare.

Multimedia Appendix 1
List of content categories of the Mika "Discover" and "Journeys" modules.
[DOCX File, 16 KB - cancer_v12i1e91416_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
SOFIA interview guide.
[DOCX File, 24 KB - cancer_v12i1e91416_app2.docx]

Multimedia Appendix 3
Study flow chart, reproduced from Sauer et al [11], Cancer, 2024, licensed under CC BY-NC-ND 4.0.
[PDF File, 103 KB - cancer_v12i1e91416_app3.pdf]

Multimedia Appendix 4
Supplement S5 for articles read.
[PDF File, 1801 KB - cancer_v12i1e91416_app4.pdf]

Multimedia Appendix 5
Supplement S6 for journeys.
[PDF File, 849 KB - cancer_v12i1e91416_app5.pdf]

Multimedia Appendix 6
Supplement S7 for exercises.
[PDF File, 769 KB - cancer_v12i1e91416_app6.pdf]

Multimedia Appendix 7
Joint display table.
[DOCX File, 21 KB - cancer_v12i1e91416_app7.docx]

Checklist 1
COREQ checklist.
[PDF File, 385 KB - cancer_v12i1e91416_app8.pdf]

References

1. Cracchiolo JR, Arafat W, Atreja A, et al. Getting ready for real-world use of electronic patient-reported outcomes (ePROs) for patients with cancer: a National Comprehensive Cancer Network ePRO Workgroup paper. Cancer 2023 Aug 15;129(16):2441-2449. [doi: 10.1002/cncr.34844] [Medline: 37224181]
2. Basch E, Stover AM, Schrag D, et al. Clinical utility and user perceptions of a digital system for electronic patient-reported symptom monitoring during routine cancer care: findings from the PRO-TECT trial. JCO Clin Cancer Inform 2020 Oct;4:947-957. [doi: 10.1200/CCI.20.00081] [Medline: 33112661]
3. Hang L, Zhang J, Lu Z, Xu J, Chen Y. Effectiveness of ePRO-based symptom management for cancer patients: a systematic review and meta-analysis of randomized controlled studies. Support Care Cancer 2024 Dec 2;32(12):842. [doi: 10.1007/s00520-024-09026-9] [Medline: 39621108]
4. Li Y, Li J, Hu X. The effectiveness of symptom management interventions based on electronic patient-reported outcomes (ePROs) for symptom burden, quality of life, and overall survival among patients with cancer: a meta-analysis of randomized controlled trials. Int J Nurs Stud 2023 Nov;147:104588. [doi: 10.1016/j.ijnurstu.2023.104588] [Medline: 37690275]
5. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565. [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
6. Harbeck N, Fasching PA, Wuerstlein R, et al. Significantly longer time to deterioration of quality of life due to CANKADO PRO-React eHealth support in HR+ HER2- metastatic breast cancer patients receiving palbociclib and endocrine therapy: primary outcome analysis of the multicenter randomized AGO-B WSG PreCycle trial. Ann Oncol 2023 Aug;34(8):660-669. [doi: 10.1016/j.annonc.2023.05.003] [Medline: 37201751]
7. Basch E, Deal AM, Dueck AC, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017 Jul 11;318(2):197-198. [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
8. Sansalone E, White J, Zucca A, et al. Acceptability of the RecoverEsupport digital health intervention among patients undergoing breast cancer surgery: qualitative study. J Med Internet Res 2025 Oct 17;27:e77567. [doi: 10.2196/77567] [Medline: 41105868]
9. Krakowczyk JB, Truijens F, Teufel M, et al. Evaluation of the e-Mental health intervention Make It Training from patients' perspectives: qualitative analysis within the Reduct Trial. JMIR Cancer 2024 Apr 9;10:e53117. [doi: 10.2196/53117] [Medline: 38592764]

XSL•FO
RenderX

10. Lai-Kwon J, Cohen JE, Lisy K, et al. The feasibility, acceptability, and effectiveness of electronic patient-reported outcome symptom monitoring for immune checkpoint inhibitor toxicities: a systematic review. JCO Clin Cancer Inform 2023 May;7:e2200185. [doi: 10.1200/CCI.22.00185] [Medline: 37220322]
11. Sauer C, Zschäbitz S, Krauss J, et al. Electronic health intervention to manage symptoms of immunotherapy in patients with cancer (SOFIA): results from a randomized controlled pilot trial. Cancer 2024 Jul 15;130(14):2503-2514. [doi: 10.1002/cncr.35300] [Medline: 38564338]
12. Zhang J, Son Y, Jiang Y, et al. User experiences of digital health interventions among older adults with cancer: a systematic review of qualitative data. Support Care Cancer 2026 Apr 6;34(4):407. [doi: 10.1007/s00520-026-10631-z] [Medline: 41936643]
13. Tolstrup LK, Pappot H, Bastholt L, Zwisler AD, Dieperink KB. Patient-reported outcomes during immunotherapy for metastatic melanoma: mixed methods study of patients' and clinicians' experiences. J Med Internet Res 2020 Apr 9;22(4):e14896. [doi: 10.2196/14896] [Medline: 32271150]
14. Glaser SLC, Fraterman I, van Brummelen N, et al. Usability and usefulness of a symptom management coaching system for patients with cancer treated with immune checkpoint inhibitors: comparative mixed methods study. JMIR Form Res 2025 Jan 23;9:e57659. [doi: 10.2196/57659] [Medline: 39847771]
15. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
16. Sauer C, Krauß J, Jäger D, et al. eHealth intervention to manage symptoms for patients with cancer on immunotherapy (SOFIA): a study protocol for a randomised controlled external pilot trial. BMJ Open 2021 Jul 19;11(7):e047277. [doi: 10.1136/bmjopen-2020-047277] [Medline: 34281921]
17. Teare MD, Dimairo M, Shephard N, Hayman A, Whitehead A, Walters SJ. Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. Trials 2014 Jul 3;15:264. [doi: 10.1186/1745-6215-15-264] [Medline: 24993581]
18. Haanen JBAG, Carbonnel F, Robert C, et al. Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2017 Jul 1;28(suppl_4):iv119-iv142. [doi: 10.1093/annonc/mdx225] [Medline: 28881921]
19. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
20. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
21. Health promotion glossary of terms 2021. : World Health Organization; 2021 URL: https://www.who.int/publications/i/item/9789240038349 [accessed 2026-06-04]
22. Warrington L, Holmes M, Gibson A, et al. Patient and clinician perspectives of an eHealth intervention for supporting cancer treatment in the UK: mixed methods evaluation of the eRAPID randomised controlled trial. BMJ Open 2024 Nov 7;14(11):e078283. [doi: 10.1136/bmjopen-2023-078283] [Medline: 39515868]
23. Sodergren SC, White A, Efficace F, et al. Systematic review of the side effects associated with tyrosine kinase inhibitors used in the treatment of gastrointestinal stromal tumours on behalf of the EORTC Quality of Life Group. Crit Rev Oncol Hematol 2014 Jul;91(1):35-46. [doi: 10.1016/j.critrevonc.2014.01.002] [Medline: 24495942]
24. Mitchell AJ, Vahabzadeh A, Magruder K. Screening for distress and depression in cancer settings: 10 lessons from 40 years of primary-care research. Psychooncology 2011 Jun;20(6):572-584. [doi: 10.1002/pon.1943] [Medline: 21442689]
25. Springer F, Mehnert-Theuerkauf A, Gebhardt C, Stolzenburg JU, Briest S. Unmet supportive care needs among cancer patients: exploring cancer entity-specific needs and associated factors. J Cancer Res Clin Oncol 2024 Apr 12;150(4):190. [doi: 10.1007/s00432-024-05715-4] [Medline: 38607426]
26. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]

Abbrevations

COREQ: Consolidated Criteria for Reporting Qualitative Research
DHI : digital health intervention
ePROs: electronic patient-reported outcomes
HRQoL : health-related quality of life
ICI: immune checkpoint inhibitor
IG: intervention group
SOFIA : eHealth intervention for managing symptoms of immunotherapy

XSL•FO
RenderX

Edited by M Balcarras; submitted 14.Jan.2026; peer-reviewed by D Bracken-Clarke, Y Chu; revised version received 06.May.2026; accepted 07.May.2026; published 18.Jun.2026.

Please cite as:
Sauer C, Sauer S, Doll ES, Zschäbitz S, Raue JS, Turni M, Bugaj TJ, Hiller K, Maatouk I
Digital Intervention for Electronic Patient-Reported Outcomes in Advanced Cancer: Mixed Methods Study
JMIR Cancer 2026;12:e91416
URL: https://cancer.jmir.org/2026/1/e91416
doi:10.2196/91416

© Christina Sauer, Simeon Sauer, Elena Sophia Doll, Stefanie Zschäbitz, Jan Simon Raue, Maik Turni, Till Johannes Bugaj, Kiriaki Hiller, Imad Maatouk. Originally published in JMIR Cancer (https://cancer.jmir.org), 18.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Real-World Impact of Electronic Patient-Reported Outcomes on Early Intervention Among Older Patients With Lung Cancer: Prospective Cohort Study

Yasuhiro Nakajima[1,2], MD, PhD; Tetsushi Maesako[1,2], MD; Takashi Murata[1,2], MD; Daisuke Minami[1], MD; Toshiyuki Minami[2], MD, PhD; Arihiko Kanehiro[1], MD, PhD; Kozo Kuribayashi[2], MD, PhD; Takashi Kijima[2], MD, PhD

[1]Department of Respiratory Medicine, Himeji St. Mary's Hospital, 650 Nibuno, Himeji, Hyogo, Japan

[2]Department of Respiratory Medicine, Hyogo Medical University Hospital, Nishinomiya, Hyogo, Japan

Corresponding Author:
Yasuhiro Nakajima, MD, PhD
Department of Respiratory Medicine, Himeji St. Mary's Hospital, 650 Nibuno, Himeji, Hyogo, Japan

Abstract

Background: Electronic patient-reported outcomes (ePROs) have demonstrated greater sensitivity in detecting adverse events than clinician-reported outcomes. However, conventional care faces challenges related to recall bias, and evidence from real-world populations of older adults is limited. This study evaluates the feasibility of an ePRO system integrated with ecological momentary assessment and ecological momentary intervention (EMI) to facilitate a concordance model within the framework of digital transformation in health care.

Objective: This study aimed to evaluate the satisfaction, usability, and clinical impact of an ePRO system integrated with ecological momentary assessment and EMI in a real-world cohort of older patients with lung cancer and to investigate its ability to facilitate a patient-centered concordance model within the framework of digital transformation in health care.

Methods: In this single-center prospective cohort study, 35 patients with lung cancer used the Welby My Carte ONC app to record daily symptoms and visual data. Satisfaction and usability were evaluated using 10-point scales at the end of the first and third treatment cycles. The study focused on the clinical impact of EMI triggered by ePRO data, such as early detection of toxicity.

Results: The cohort included a high proportion of older patients (20/35, 57%) aged 70 to 89 years. Satisfaction and usability scores significantly improved over time (satisfaction: 6.0 to 8.0, $P<.001$; usability: 5.0 to 8.0, $P<.001$). ePRO monitoring enabled the early detection of critical events, including immune-related uveitis (case 1), Stevens-Johnson syndrome (case 2), rare drug-induced purpura (case 3), and the monitoring of infusion-related reactions (case 4).

Conclusions: ePRO monitoring is feasible for older patients with lung cancer and enhances safety through timely EMI. It serves as a core component of a comprehensive patient support program and supports the transition to a more active concordance model, ultimately improving patient well-being.

(JMIR Cancer 2026;12:e97890) doi:10.2196/97890

KEYWORDS

lung cancer; electronic patient-reported outcome; ePRO; digital transformation; geriatric oncology; ecological momentary assessment; EMA

Introduction

Patient-reported outcomes are measures of a patient's health status that come directly from the patient, without interpretation by a clinician. In oncology, electronic patient-reported outcome (ePRO) systems have gained significant attention for their potential to enhance symptom management and clinical outcomes. Previous research, most notably the landmark study by Basch et al [1], demonstrated that ePRO systems are more sensitive and reliable in detecting adverse events, such as fatigue, nausea, and diarrhea, than clinician-reported outcomes. Furthermore, subsequent large-scale randomized trials have shown that ePRO-based symptom monitoring during routine cancer treatment can significantly improve overall survival and quality of life [2]. Recent evidence has further expanded on this by highlighting that high patient satisfaction and acceptability are critical for successful ePRO implementation, particularly among older populations who may initially perceive digital tools as burdensome. Despite these advantages, conventional outpatient care is inherently limited by recall bias, as patients are often required to retrospectively report their symptoms over several weeks during periodic clinic visits. To address this limitation, ecological momentary assessment (EMA) and ecological momentary intervention (EMI) have been proposed [3]. EMA allows real-time symptom recording in a patient's

XSL•FO
RenderX

natural environment, thereby enhancing ecological validity and eliminating recall bias, while EMI uses these data to provide immediate, personalized clinical support [3]. For older patients, who frequently face specific challenges such as cognitive decline, visual impairment, limited digital literacy, and social isolation, the real-time nature of EMA and EMI is particularly important for preventing the oversight of subtle but critical symptom changes. The integration of these digital tools is a cornerstone of digital transformation in health care. In this study, the ePRO system is positioned not merely as a stand-alone monitoring tool but as a central component of a comprehensive patient support program (PSP) [4]. This program integrates education, symptom management, and multidisciplinary collaboration to bridge the communication gap between clinic visits, similar to integrated care models that have shown benefits for patients with lung cancer [5]. In the context of lung cancer treatment, digital transformation in health care aims to move beyond passive patient compliance toward a patient-centered concordance model [6]. This model emphasizes a partnership between health care providers and patients, in which decisions are made based on shared goals and the patient's lived experience. Such shared decision-making is further supported by the integration of outcome information into clinical conversations [7]. However, a gap remains in the evidence regarding the longitudinal feasibility and patient satisfaction associated with such systems in real-world aging populations with cancer. Therefore, this study aimed to bridge the gap between clinical trial evidence and real-world practice. By evaluating the satisfaction, usability, and clinical impact of an ePRO-EMI system in older patients with lung cancer, we investigated whether digital transformation in health care can effectively support a patient-centered concordance model and improve patient well-being in a real-world setting.

Methods

Study Design and Setting

This single-center prospective cohort study was conducted at Himeji St Mary's Hospital in Himeji, Japan, between 2023 and 2024. The institution was selected as it serves as a regional core center for oncology, providing a high volume of systemic therapies for a diverse, real-world population of patients with lung cancer. The study analyzed longitudinal data obtained from routine clinical practice to bridge the gap between clinical trial evidence and real-world implementation.

Participants and Enrollment

Eligible participants included patients with lung cancer receiving systemic therapy, including cytotoxic chemotherapy, immune checkpoint inhibitors (ICIs), and molecularly targeted therapies. Sampling was conducted on a consecutive basis for all eligible patients at our institution during the study period. As the ePRO system was implemented as part of standard care for self-monitoring, all eligible patients were encouraged to participate. The decision to use or withdraw from the app did not affect the standard of medical care provided.

The ePRO Software and Workflow

The ePRO system used was Welby My Carte ONC (Welby Inc). The software interface allowed patients or their caregivers to record daily symptoms (eg, fatigue and appetite loss), psychological status, and physical findings through photographs (eg, skin rashes). To support adherence, particularly among older patients, the app provided push-notification reminders and visual tracking of symptom trends. Training and onboarding were provided by health care staff at the time of treatment initiation. Data were transmitted in real time to a secure dashboard accessible by the clinical team, facilitating immediate clinical decision-making through an integrated EMA-EMI framework. Representative screenshots of the Welby My Carte ONC platform are provided in Multimedia Appendix 1. A schematic overview of the EMA-EMI clinical workflow is shown in Multimedia Appendix 2.

Measures and Clinical Data

Patient-reported outcomes were evaluated using structured questionnaires at the end of the first cycle (interim evaluation) and the third cycle (final evaluation). The questionnaire assessed four quantitative items: (1) satisfaction with the app, (2) ease of use (usability), (3) willingness to continue using the app, and (4) clarity of the instructions provided by staff. Each item was scored on a 10-point Likert scale, where 1 represented "strongly disagree" or "very poor" and 10 represented "strongly agree" or "excellent." Open-ended responses were also collected to identify perceived advantages and challenges. Clinical covariates, including age, sex, Eastern Cooperative Oncology Group performance status (PS), disease stage, and line of therapy, were extracted from electronic medical records. Because the questionnaire was designed as a pragmatic usability assessment tool for routine clinical implementation, formal psychometric validation was not performed. The complete questionnaire items used for patient evaluation are provided in Multimedia Appendix 3.

Clinical Impact and Case Selection

The clinical impact of the system was evaluated based on the frequency and nature of physician-initiated interventions (EMIs) triggered by ePRO data, such as telephone consultations, early clinic visits, and dose modifications. To illustrate the clinical utility of the system, 4 representative cases were selected based on their diversity of toxicity types (eg, immune-related and drug-induced) and their demonstration of how real-time monitoring can lead to improved safety.

Statistical Analysis

Descriptive statistics were used to summarize patient characteristics and clinical data. For longitudinal comparisons of ePRO scores between cycle 1 and cycle 3, the Wilcoxon signed-rank test was used. This nonparametric approach was selected because of the ordinal nature of the 10-point scale data and the relatively small sample size. Missing questionnaire data were handled using complete case analysis. Qualitative data from open-ended responses were analyzed using a descriptive thematic analysis framework. The process involved an inductive approach in which 2 researchers independently reviewed the free-text comments to identify recurring patterns and codes.

XSL•FO
RenderX

These codes were then grouped into overarching themes. Discrepancies in thematic assignment were resolved through consensus among the research team. The qualitative findings were intended to provide exploratory supplementary insights that complemented the quantitative results.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and was approved by the institutional review board of Himeji St Mary's Hospital (S024-001). All participating patients provided written informed consent prior to data collection.

Reporting Guidelines

This study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guideline for observational cohort studies. The completed STROBE checklist is provided in Multimedia Appendix 4.

Results

Patient Characteristics

A total of 35 patients with lung cancer were included in the analysis (Table 1).

The cohort was characterized by a high proportion of older individuals, with 20 (57.1%) of the patients aged ≥70 years (n=11 aged 70-79 years and n=9 aged 80-89 years). Most participants were male (24/35, 68.6%) and maintained a good PS (PS 0: n=23; PS 1: n=10). Treatment regimens were diverse, reflecting real-world practice: 19 patients received chemotherapy combined with ICIs, while others received cytotoxic chemotherapy (n=6), ICI monotherapy (n=4), tyrosine kinase inhibitors (TKIs; n=4), or TKI-chemotherapy combinations (n=2).

Table . Patient characteristics (N=35).

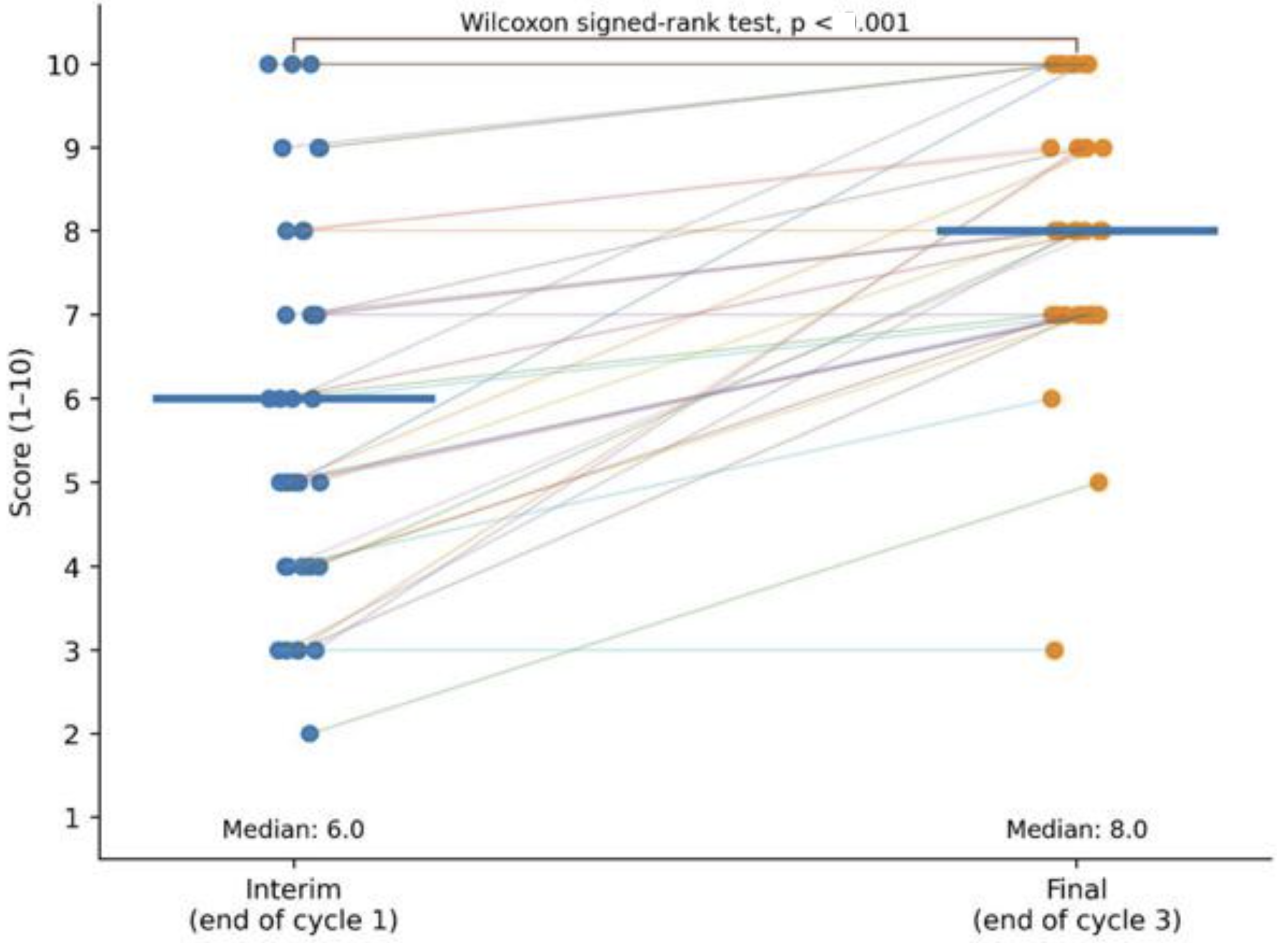
Characteristics	Patients, n (%)
Age group (years)	
40-49	1 (2.9)
50-59	5 (14.3%)
60-69	9 (25.7)
70-79	11 (31.4)
80-89	9 (25.7)
Sex	
Male	24 (68.6)
Female	11 (31.4)
Performance status	
0	23 (65.7)
1	10 (28.6)
≥2	2 (5.7)
Treatment regimens	
Cytotoxic chemotherapy	6 (17.1)
Chemotherapy+immune checkpoint inhibitor	19 (54.3)
Immune checkpoint inhibitor–based therapy	4 (11.4)
Tyrosine kinase inhibitor–based therapy	4 (11.4)
Chemotherapy+tyrosine kinase inhibitor	2 (5.7)

Quantitative Patient-Reported Outcomes

Quantitative evaluations revealed that patients successfully adapted to the ePRO system over the course of 3 treatment cycles. All measured items showed significant improvement from the interim evaluation to the final evaluation. Specifically, the median patient satisfaction score increased from 6.0 at the interim evaluation (cycle 1) to 8.0 at the final evaluation (cycle 3; P<.001; Figure 1). Usability scores, reflecting ease of use, increased from a median of 5.0 to 8.0 (P<.001; Figure 2). Willingness to continue using the app improved from a median of 6.0 to 8.0 (P<.001; Figure 3). Furthermore, the clarity of the instructional guidance provided by the clinical staff showed a significant upward trend, with median scores increasing from 5.0 to 8.0 (P<.001; Figure 4). These findings suggest that even older patients can achieve high levels of proficiency and satisfaction with digital monitoring tools through repeated use.

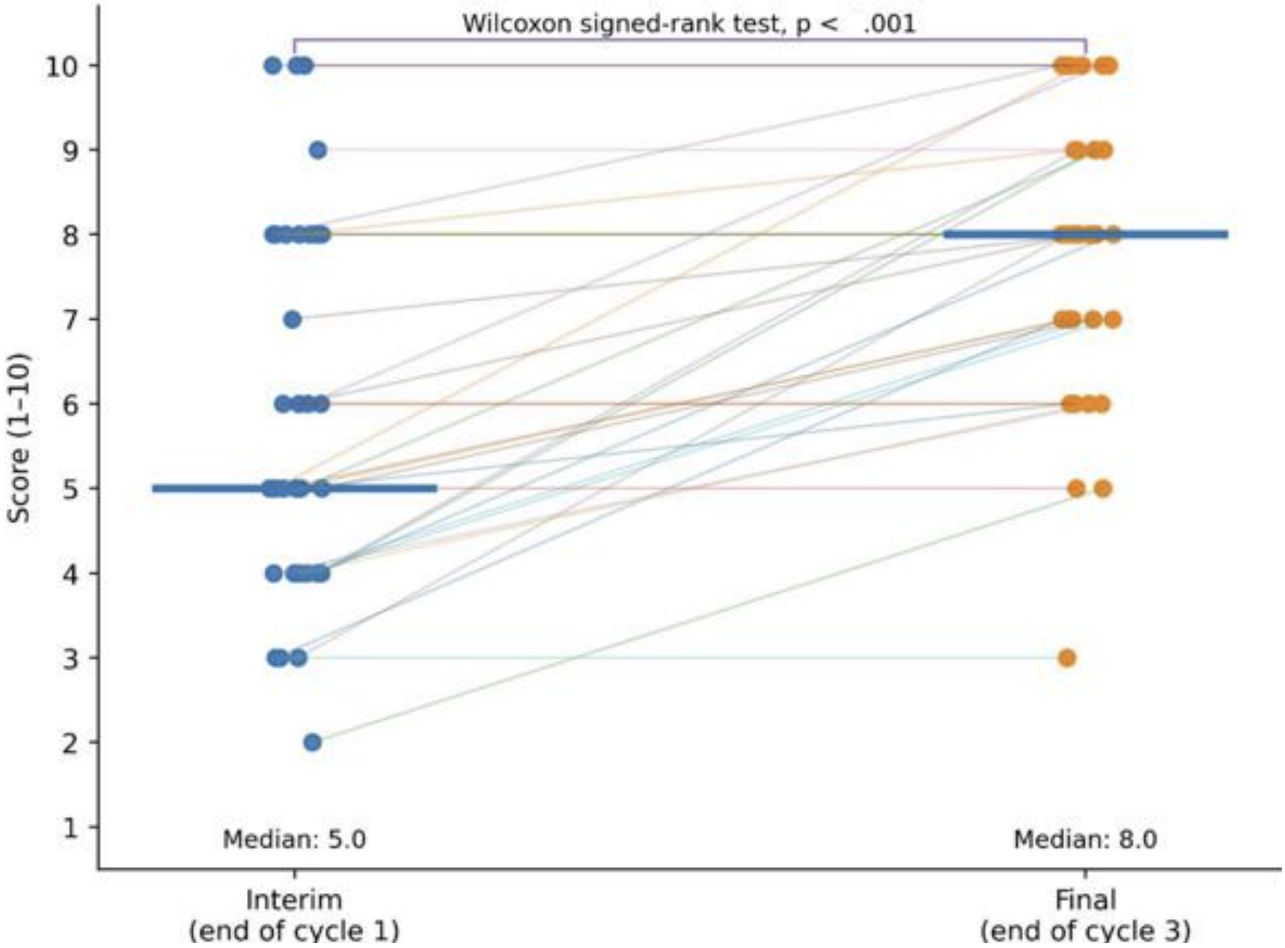
XSL•FO
RenderX

Figure 1. Satisfaction with the electronic patient-reported outcome app.



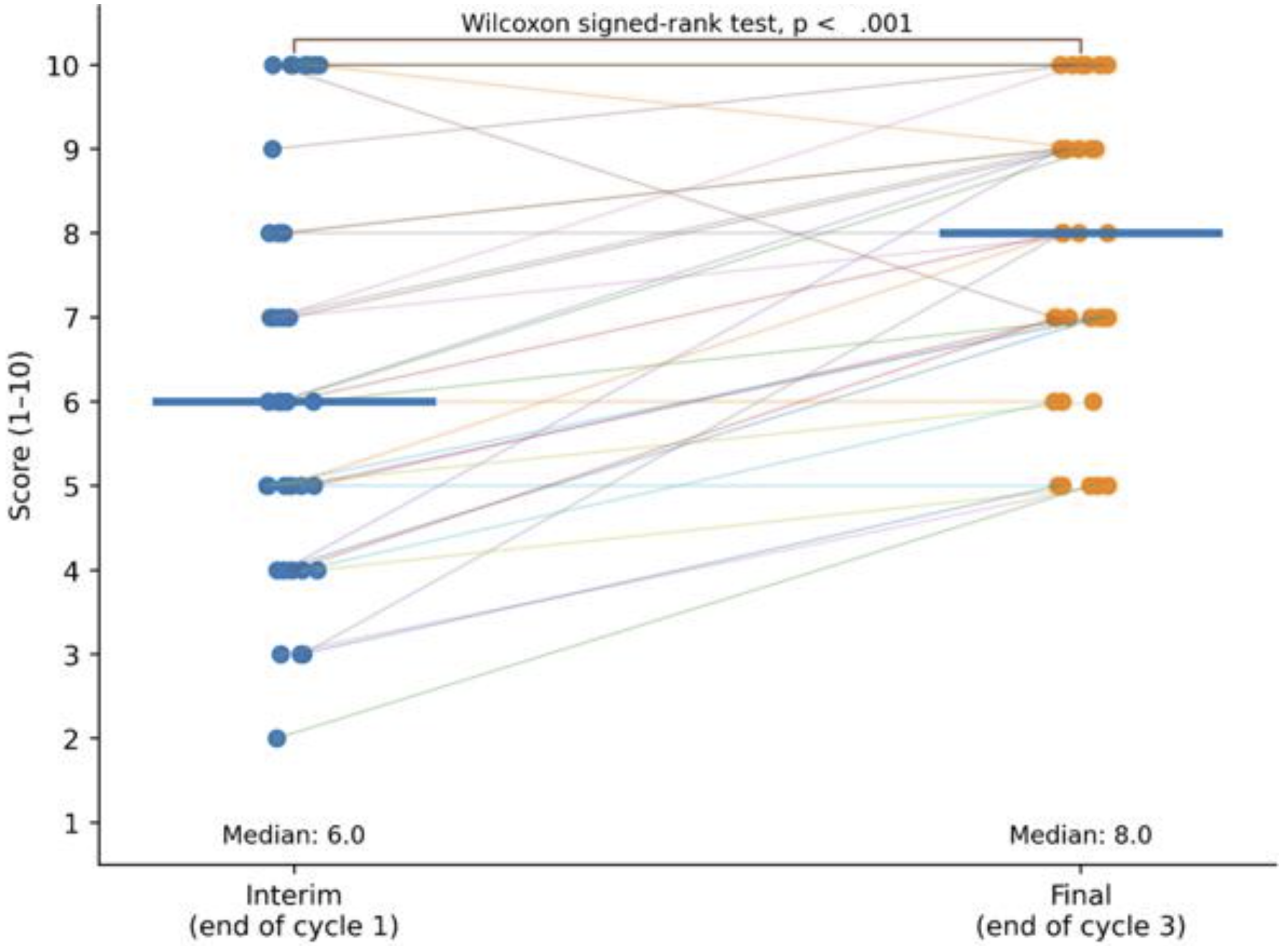
XSL•FO
RenderX

Figure 2. Usability of the electronic patient-reported outcome app.



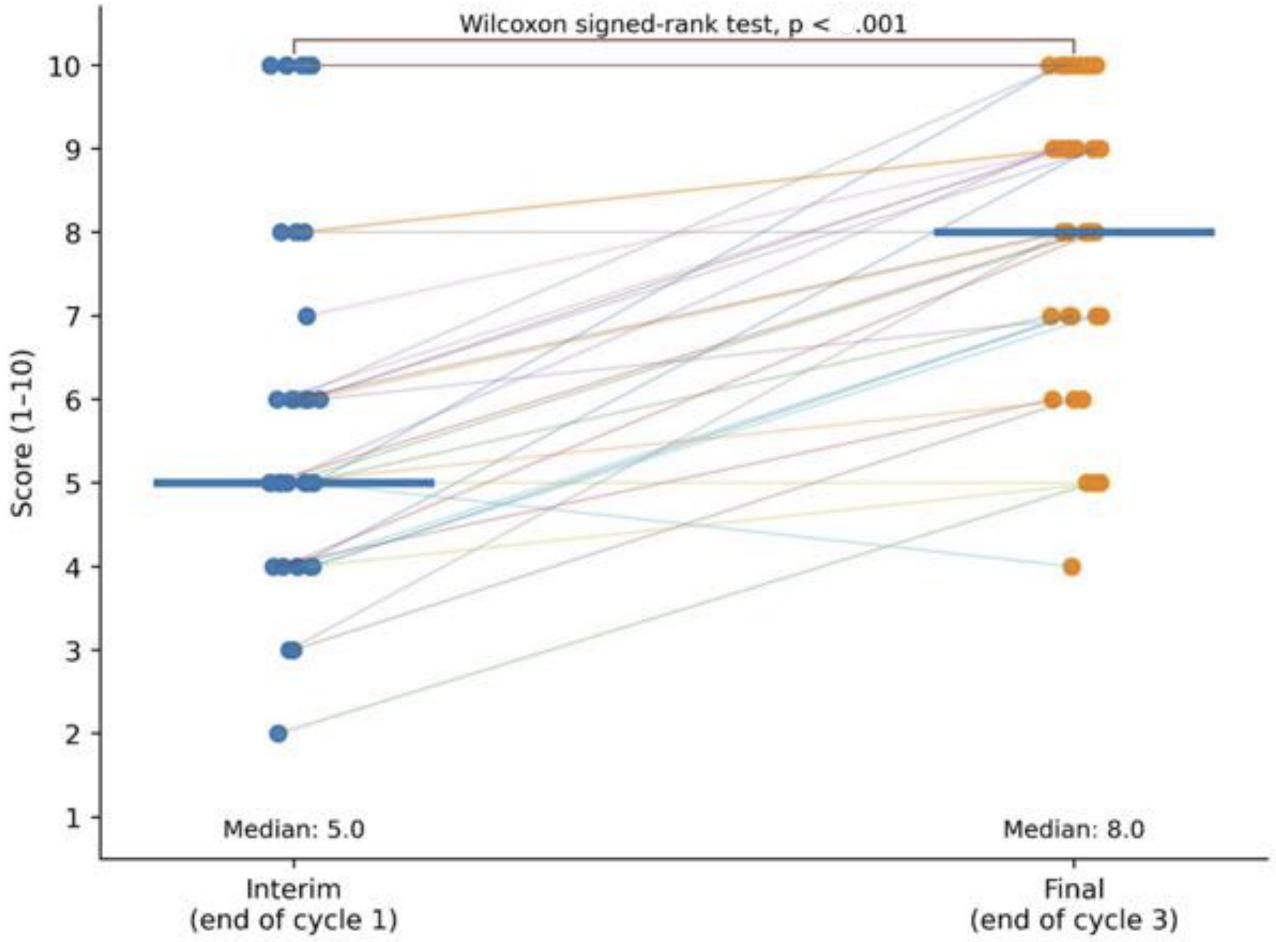
XSL•FO
RenderX

Figure 3. Willingness to continue using the electronic patient-reported outcome app.



XSL•FO
RenderX

Figure 4. Patient-reported clarity of instructions provided by clinical staff.



Qualitative Feedback: Advantages and Challenges

The thematic analysis of qualitative feedback identified 4 primary areas of focus regarding the patient experience. First, daily ePRO input was found to enhance self-monitoring and provide a significant sense of psychological security. By recording symptoms daily, patients reported that they could more effectively distinguish between expected treatment-related adverse events and general daily fatigue, thereby reducing unnecessary anxiety. Real-time connectivity with the clinical team further contributed to a perceived sense of security, particularly during the early phases of treatment. Second, patients consistently preferred the digital interface over traditional paper-based documentation. Participants highlighted the convenience and low burden of mobile data entry, as well as the clinical utility of longitudinal symptom visualizations, which enabled clearer communication of symptom trajectories during outpatient visits and reduced reliance on retrospective recall. Third, the system facilitated multigenerational support, particularly among the oldest participants. Family members frequently assisted with data entry, allowing them to become more actively involved in the care process while also providing a practical means of monitoring the patient's condition remotely. Finally, despite these benefits, several technical and functional barriers were identified. Some patients expressed uncertainty regarding successful data transmission and suggested the inclusion of automated confirmation messages following submission. Participants also indicated that the integration of educational resources, such as a built-in reference guide to treatment-related adverse events, would enhance their ability to interpret symptom changes and manage treatment expectations. Collectively, these findings suggest that the perceived benefits of ePRO extended beyond symptom monitoring to include psychological reassurance and family engagement, while the principal challenges were related to communication feedback mechanisms and patient education.

Clinical Impact and Representative Case Narratives

The clinical utility of the ePRO-EMI framework was demonstrated through its ability to capture clinically significant adverse events that might otherwise have remained undetected between standard outpatient visits. The system facilitated rapid physician-initiated interventions (EMI) across a variety of toxicity profiles, as illustrated by the following 4 representative cases selected to show the diversity of detected events. In the first case, a 71-year-old male reported "foggy vision" via the app on the first day of pembrolizumab treatment. This early EMA data led to a prompt hospital visit on day 9, resulting in the diagnosis and successful management of grade 2 immune-related uveitis without treatment interruption. Similarly, a 40-year-old female used the image upload function to report a skin rash. Real-time monitoring of these images allowed clinicians to detect the rapid progression of toxicity from grade 2 to grade 3, leading to an immediate diagnosis of Stevens-Johnson syndrome and life-saving corticosteroid intervention.

XSL•FO
RenderX

The system also proved effective in identifying rare toxicities, as demonstrated by the case of a 83-year-old female who uploaded photographs of lower-extremity skin lesions. This led to the diagnosis of epidermal growth factor receptor-TKI–associated purpura and enabled a flexible dose modification to maintain treatment continuity. Finally, the ePRO system provided a continuous layer of safety monitoring for a 50-year-old female who had experienced severe infusion-related reactions to amivantamab. Postdischarge monitoring confirmed the absence of symptom exacerbation, allowing safe continuation of therapy in an outpatient setting and reducing patient anxiety. These cases highlight how the integration of real-time patient data into clinical workflows can directly enhance patient safety.

Discussion

Principal Findings

This study demonstrated that an ePRO system integrated with the EMA-EMI framework significantly enhanced patient satisfaction and usability over time while enabling the early detection of critical adverse events in a real-world population of older adults with lung cancer. Specifically, the longitudinal improvement in median scores from 6.0 to 8.0 across satisfaction and usability metrics suggests that older patients can effectively adapt to digital health tools when provided with continuous institutional support. Our findings regarding feasibility and satisfaction align with recent evidence highlighting the potential of ePRO systems in older adults. While previous studies have often noted lower completion rates in older adult cohorts, Riedl et al [8] observed that home-based ePRO systems are feasible for >80% of older adults who survived cancer when family and staff support are integrated. The significant increase in scores for clarity of instructions (Figure 4) in our cohort suggests that the initial digital barrier can be overcome through repeated use and proactive staff engagement, leading to high levels of acceptance. The clinical impact of the ePRO-EMI system was particularly evident in the early detection of potentially life-threatening toxicities, bridging the gap between periodic outpatient visits. For instance, the detection of immune-related uveitis (case 1) within the first week of treatment is crucial, as this condition can manifest much earlier than the reported median latency of 12 weeks [9]. Similarly, real-time monitoring of image data in case 2 allowed the immediate diagnosis of Stevens-Johnson syndrome, which is characterized by a high mortality risk among patients receiving combination immunotherapy [10]. For targeted therapies, ePRO facilitated the management of rare toxicities, such as osimertinib-induced cutaneous vasculitis (case 3) [11], and provided an additional safety layer for outpatient monitoring of infusion-related reactions (case 4) [12]. These cases collectively illustrate how the EMA-EMI framework mitigates recall bias and enhances ecological validity by capturing subtle symptom changes in the patient's natural environment. Furthermore, our results suggest that ePRO should not be viewed as a stand-alone monitoring tool but as a core component of a comprehensive PSP [4]. The integration of real-time data into clinical workflows supports a transition from passive compliance to an active concordance model, in which treatment decisions are made through shared decision-making based on the patient's lived experience and real-time status [6,7]. Despite these promising results, several challenges remain. Adherence tended to decline during periods of symptom stability, suggesting that future iterations of the system could benefit from behavioral economics or gamification strategies to maintain long-term engagement [13]. Additionally, as 57.1% of our cohort was aged ≥70 years, family members often assisted with data entry. This introduces the potential for interrater variability, as proxies may overestimate symptom severity compared to patient self-reports [14]. This study has several limitations. First, as a single-center prospective cohort study with a relatively small sample size (N=35), the generalizability of our findings to broader populations may be limited. Second, the absence of a control group receiving conventional care makes it difficult to definitively determine whether the ePRO-EMI approach is superior to conventional care. Finally, although we used a descriptive thematic analysis to categorize patient feedback, the qualitative analysis was conducted on a limited dataset from a single-center cohort. Therefore, these findings should be interpreted as exploratory and illustrative of this specific clinical setting, rather than as fully generalizable to all older adults with cancer. In addition, the long-term sustainability and scalability of digital health interventions remain challenging, particularly in aging populations with heterogeneous levels of digital literacy and varying support systems [15]. In conclusion, this study provides real-world evidence that ePRO-based monitoring is feasible and beneficial for older patients with lung cancer. Future research should integrate artificial intelligence to provide automated, predictive warnings of toxicities and to foster deeper collaboration with regional health care systems [16]. Ultimately, digital transformation in health care, centered on patient-reported data, holds the potential to significantly improve patient well-being and safety in an aging society.

Conclusions

ePRO integrated with EMA and EMI is feasible and clinically valuable for older patients with lung cancer in real-world settings. Real-time monitoring enabled the early detection of clinically significant adverse events and supported patient-centered care beyond routine outpatient visits. However, successful implementation requires continuous institutional support, improved digital literacy, and integration into comprehensive PSPs. Future research should focus on artificial intelligence–assisted toxicity prediction and broader regional health care collaboration to optimize sustainable implementation in aging societies.

Acknowledgments

The authors thank all participating patients and collaborating institutions for their contributions to this study. A generative artificial intelligence tool (ChatGPT; OpenAI) was used solely for English-language editing and for the conversion of figure files from

XSL•FO
RenderX

PDF to PNG format during manuscript preparation. All scientific content, clinical interpretations, data analyses, references, and final manuscript decisions were independently reviewed and verified by the authors.

Funding

The authors declared that no financial support was received for this work.

Data Availability

The datasets used and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

YN conceived and designed the study. YN, T Maesako, T Murata, DM, AK, KK, and TK collected the clinical data. YN performed the statistical analyses and drafted the manuscript. All authors contributed to data interpretation, critically revised the manuscript for important intellectual content, and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of the electronic patient-reported outcome system (Welby My Carte ONC).
[PNG File, 336 KB - cancer_v12i1e97890_app1.png]

Multimedia Appendix 2
Ecological momentary assessment–ecological momentary intervention workflow and clinical response framework.
[PNG File, 189 KB - cancer_v12i1e97890_app2.png]

Multimedia Appendix 3
Questionnaire items for patient evaluation.
[PNG File, 225 KB - cancer_v12i1e97890_app3.png]

Multimedia Appendix 4
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist for cohort studies.
[PNG File, 520 KB - cancer_v12i1e97890_app4.png]

References

1. Basch E, Jia X, Heller G, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. J Natl Cancer Inst 2009 Dec 2;101(23):1624-1632. [doi: 10.1093/jnci/djp386] [Medline: 19920223]
2. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565. [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
3. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39. [doi: 10.1348/135910709X466063] [Medline: 19646331]
4. Ganguli A, Clewell J, Shillington AC. The impact of patient support programs on adherence, clinical, humanistic, and economic patient outcomes: a targeted systematic review. Patient Prefer Adherence 2016;10(1):711-725. [doi: 10.2147/PPA.S101175] [Medline: 27175071]
5. Temel JS, Greer JA, El-Jawahri A, et al. Effects of early integrated palliative care in patients with lung and GI cancer: a randomized clinical trial. J Clin Oncol 2017 Mar 10;35(8):834-841. [doi: 10.1200/JCO.2016.70.5046] [Medline: 28029308]
6. Jordan JL, Ellis SJ, Chambers R. Defining shared decision making and concordance: are they one and the same? Postgrad Med J 2002 Jul;78(921):383-384. [doi: 10.1136/pmj.78.921.383] [Medline: 12151651]
7. Mohindra NA, Garcia SF, Kircher S, et al. Development of an electronic health record-integrated patient-reported outcome-based shared decision-making dashboard in oncology. JAMIA Open 2024 Oct;7(3):ooae056. [doi: 10.1093/jamiaopen/ooae056] [Medline: 39049991]
8. Riedl D, Lehmann J, Rothmund M, et al. Usability of electronic patient-reported outcome measures for older patients with cancer: secondary analysis of data from an observational single center study. J Med Internet Res 2023 Sep 21;25:e49476. [doi: 10.2196/49476] [Medline: 37733409]
9. Wu Z, Sun W, Wang C. Clinical characteristics, treatment, and outcomes of pembrolizumab-induced uveitis. Invest New Drugs 2024 Oct;42(5):510-517. [doi: 10.1007/s10637-024-01464-w] [Medline: 39141261]

XSL•FO
RenderX

10. Zheng Z, Shen Z. Clinical characteristics of Stevens-Johnson syndrome/toxic epidermal necrolysis-like reactions induced by immune checkpoint inhibitors. Oncologist 2025 Jun 4;30(6):oyaf143. [doi: 10.1093/oncolo/oyaf143] [Medline: 40491130]
11. Zhang Y, Ling M, Wang M, Chen Y, Zhang L. Rare skin adverse reactions induced by osimertinib: a case report and literature review. Front Oncol 2025;15:1523541. [doi: 10.3389/fonc.2025.1523541] [Medline: 39917172]
12. Park K, Sabari JK, Haura EB, et al. Management of infusion-related reactions (IRRs) in patients receiving amivantamab in the CHRYSALIS study. Lung Cancer 2023 Apr;178:166-171. [doi: 10.1016/j.lungcan.2023.02.008] [Medline: 36868177]
13. Huang X, Xiang X, Liu Y, Wang Z, Jiang Z, Huang L. The use of gamification in the self-management of patients with chronic diseases: scoping review. JMIR Serious Games 2023 Dec 22;11(2):e39019. [doi: 10.2196/39019] [Medline: 38133907]
14. Roydhouse JK, Gutman R, Keating NL, Mor V, Wilson IB. Proxy and patient reports of health-related quality of life in a national cancer survey. Health Qual Life Outcomes 2018 Jan 5;16(1):6. [doi: 10.1186/s12955-017-0823-5] [Medline: 29304818]
15. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
16. Machine learning models predict adverse events and survival in patients receiving immunotherapy. ESMO Daily Reporter. 2025. URL: https://dailyreporter.esmo.org/esmo-targeted-anticancer-therapies-congress-2025/highlights-from-tat-2025/machine-learning-models-predict-adverse-events-and-survival-in-patients-receiving-immunotherapy [accessed 2026-06-21]

Abbreviations

EMA: ecological momentary assessment
EMI: ecological momentary intervention
ePRO: electronic patient-reported outcome
ICI: immune checkpoint inhibitor
PS: performance status
PSP: patient support program
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TKI: tyrosine kinase inhibitor

Edited by M Balcarras; submitted 12.Apr.2026; peer-reviewed by I Kudel, S Nasiri; revised version received 05.Jun.2026; accepted 08.Jun.2026; published 06.Jul.2026.

Please cite as:
Nakajima Y, Maesako T, Murata T, Minami D, Minami T, Kanehiro A, Kuribayashi K, Kijima T
Real-World Impact of Electronic Patient-Reported Outcomes on Early Intervention Among Older Patients With Lung Cancer: Prospective Cohort Study
JMIR Cancer 2026;12:e97890
URL: https://cancer.jmir.org/2026/1/e97890
doi:10.2196/97890

© Yasuhiro Nakajima, Tetsushi Maesako, Takashi Murata, Daisuke Minami, Toshiyuki Minami, Arihiko Kanehiro, Kozo Kuribayashi, Takashi Kijima. Originally published in JMIR Cancer (https://cancer.jmir.org), 6.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

“Like a Clinical Nurse Consultant in Your Pocket”—Parents’ and Health Care Providers’ Perceptions of a Co-Designed Smartphone App Following Treatment for Pediatric Cancer: Mixed Methods Evaluation Study

Hannah Pring[1,2], BA, MYMH; Cinzia R De Luca[2,3,4], BA, BSc, MA, PhD; Georgia Taylor[1], BSc, MHA; Chris Williams[1,2], BSN, MA; Megan Nagarajah[5], BND; David D Eisenstat[3,4,6], MD, MA, FRCPC, FRACP; Maria C McCarthy[2,3,4], BAppSc, MA, PhD

[1]Paediatric Integrated Cancer Service, The Royal Children's Hospital, South Building, 50 Flemington Road, Melbourne, Victoria, Australia

[2]Clinical Sciences, Murdoch Children's Institute, The Royal Children's Hospital, 50 Flemington Road, Melbourne, Victoria, Australia

[3]Children's Cancer Centre, The Royal Children's Hospital, Melbourne, Victoria, Australia

[4]Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia

[5]Innovation and Commercialisation, Murdoch Children's Research Institute, Melbourne, Victoria, Australia

[6]Stem Cell Medicine, Murdoch Children's Research Institute, Melbourne, Victoria, Australia

Corresponding Author:
Maria C McCarthy, BAppSc, MA, PhD
Clinical Sciences, Murdoch Children's Institute, The Royal Children's Hospital, 50 Flemington Road, Melbourne, Victoria, Australia

Abstract

Background: Pediatric cancer presents many challenges for young patients and families that extend beyond active treatment. Families report significant unmet information and supportive care needs at the end of treatment (EOT). Digital health solutions, such as patient apps, offer a potential solution to address current gaps in care.

Objective: This study aimed to develop and test a smartphone app called LotusLab, designed to improve access to survivorship information for parents of childhood cancer survivors aged 0 to 18 years, and for young people (aged 12 years and older) at the EOT. In addition, this study aimed to examine parents’ and health care providers’ (HCPs) perceptions of LotusLab during β-testing to inform further app development and refinement.

Methods: Participants were invited to explore the content and key functionality of LotusLab, including an embedded patient-reported outcome measure (PROM). Evaluation of the usability, acceptability, usefulness, and feasibility of the app was conducted through quantitative surveys and qualitative semistructured interviews.

Results: Participants comprised 8 parents whose children had completed cancer treatment, and 9 HCPs. All participants (100%) “agreed” or “strongly agreed” that the app was acceptable; however, HCPs reported more neutral acceptability responses compared with parents. Parents scored app usability more highly, with a mean System Usability Scale (SUS) score of 86.88 (SD 8.74), than HCPs’ mean SUS score of 66.39 (SD 10.16). A common perceived benefit of LotusLab was the provision of trusted information and psychosocial resources. Suggestions for improvement included adding a search function, enhancing the layout, and adding additional topics. All parents completed the app-embedded PROM. Both cohorts noted the value of the PROM in enhancing EOT consultations, while highlighting considerations for the effective integration of and response to PROM data.

Conclusions: Initial data indicate that LotusLab is easy for parents to use and provides value through the provision of EOT information and the opportunity to enhance the use of the PROM. Consistent with the co-design methodology, these findings are guiding the ongoing development and implementation of LotusLab.

(JMIR Cancer 2026;12:e90177) doi:10.2196/90177

KEYWORDS

digital health; childhood cancer; survivorship; co-design; pediatric; app

XSL•FO
RenderX

Introduction

A cancer diagnosis in childhood presents extraordinary challenges for patients and families who must navigate complex treatments and profound impacts on family life. The end-of-treatment (EOT) period represents a critical transition period for families, as they re-establish regular family routines while managing treatment side effects and fear of relapse. Multiple studies highlight that this period is marked by significant unmet information and supportive care needs for families [1-4]. Provision of relevant and timely information at the EOT represents a critical strategy for bridging identified gaps in care and mitigating associated psychological distress [5,6]. However, families continue to report barriers to accessing information when needed [2,7-9]. Research examining the optimal delivery of early survivorship information indicates that such information should be tailored to the needs and preferences of the individual, presented in multimodal formats (eg, short videos, animations, and testimonials), and available electronically [10-12].

Mobile apps have been found to be useful and acceptable support tools for young people and caregivers experiencing cancer [13-21]. The potential benefits of apps include increased accessibility to trusted information due to high levels of mobile phone use and opportunities to actively engage users in their health care, including through the delivery of personalized care plans and patient-reported outcome measures (PROMs) [19,21-26]. Notably, there are examples of co-designed digital PROM tools in pediatric oncology that have demonstrated strong and sustained user engagement, particularly during active treatment, providing symptom monitoring and delivery of tailored information [21,27].

Although there are positive examples of progress in the development of apps and digital PROMs, real-world uptake and ongoing engagement with digital health tools remain low, often due to a lack of robust co-design and testing [28,29]. Therefore, to ensure ongoing sustainability, it is essential to actively engage users in all phases of the development and testing of mobile interventions [24,30].

Despite an increasing number of apps developed for families experiencing a cancer diagnosis [23], none were identified that provide information for both young people and parents, and specifically target the EOT transition within a pediatric setting. This paper reports the findings from a study undertaken during the β-testing phase of the co-design of a digital survivorship companion app called LotusLab. The aim of this study was to examine parent caregivers' and health care providers' (HCPs) perceptions of the app, including early-phase testing and perspectives on the use of a proxy-reported outcome measure embedded within the app.

Methods

Setting

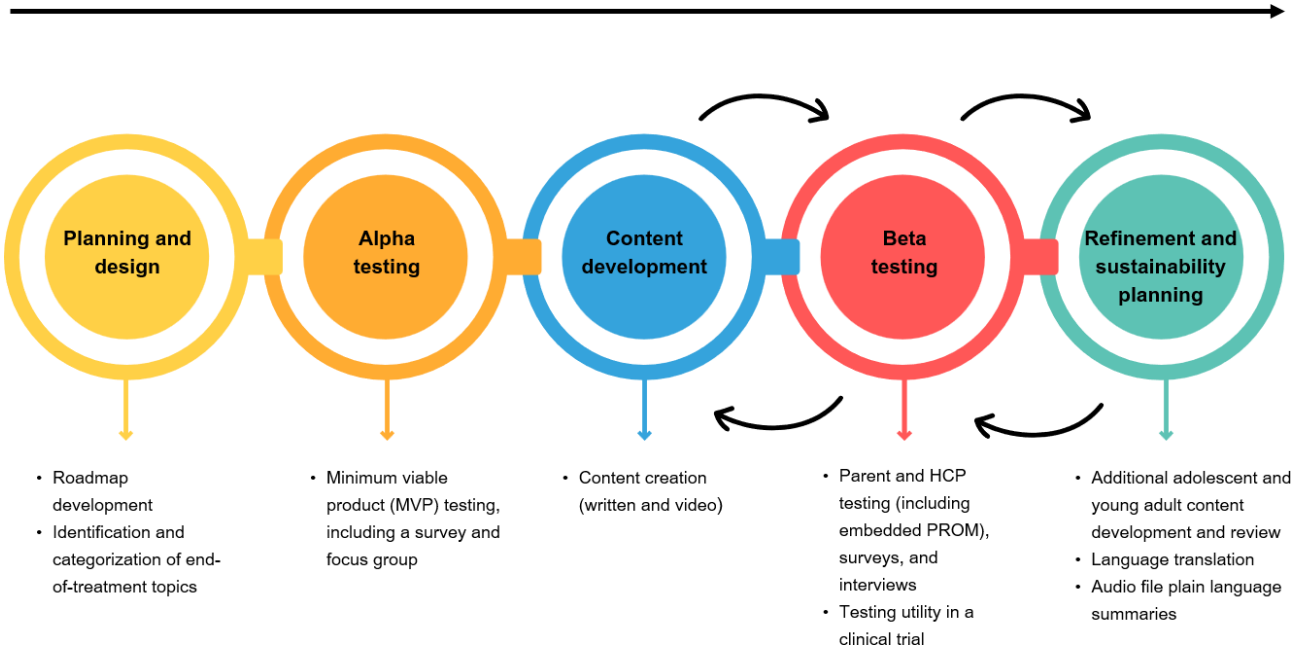
This study was conducted at The Royal Children's Hospital (RCH), a major tertiary pediatric hospital in Melbourne, Australia.

Study Design

Study Phases

Our research team has undertaken a 3-year mixed method study to co-design, develop, and test LotusLab, a digital hub for information after childhood cancer. LotusLab is designed for caregivers of childhood cancer survivors and for young people (aged 12 years and older) at EOT. LotusLab has been co-designed with young people, caregivers, and HCPs in alignment with participatory design principles [31]. Figure 1 shows an overview of the co-design activities across the app development and testing process.

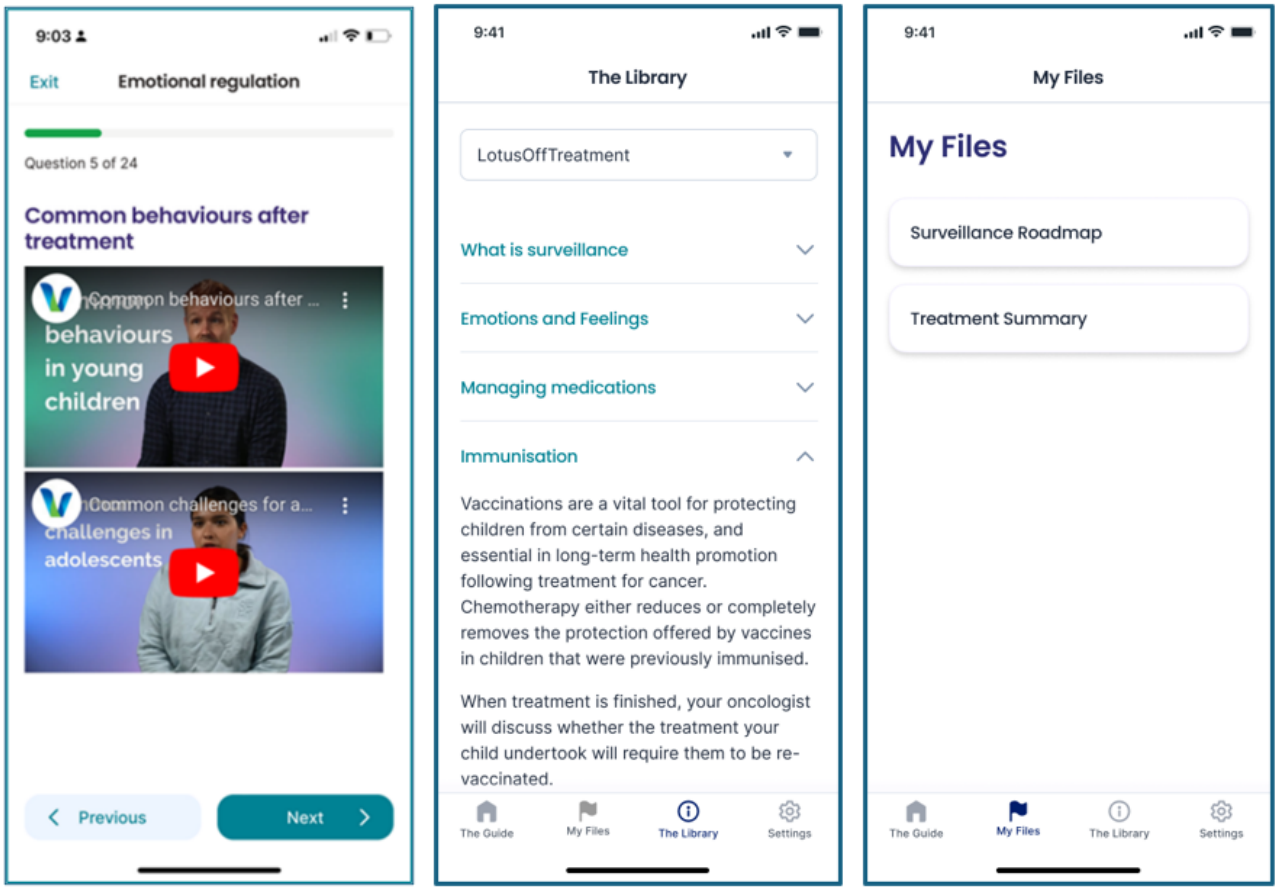
Figure 1. LotusLab co-design and development process. Co-design activities with parents, young people, and HCPs. HCP: health care provider; PROM: patient-reported outcome measure.



XSL•FO
RenderX

As a result of this iterative development process, the LotusLab prototype, available for β-testing, provided tailored information and videos, surveys, PROMs, notifications and reminders, and storage of personal survivorship documentation (manuscript under review). Figure 2 shows screenshots of the prototype.

Figure 2. LotusLab screenshots. Both individuals in the screenshots have provided written consent for their images to be used for the purposes of this manuscript.



Following planning, design, α-testing (ie, early-phase internal testing), and content development phases, the current study was conducted during the β-testing phase, which focused on testing performance and experience in "real-world" environments. The aim was to evaluate parents' and HCPs' perceptions of LotusLab and the use of a proxy-reported outcome measure embedded within the app. Usability, acceptability, usefulness, and feasibility (refer to Table 1 for descriptions) were evaluated through quantitative surveys and qualitative semistructured interviews. Descriptions of these 4 domains were informed by digital health evaluation and implementation science literature [32-34] to ensure that we captured technical outcomes related to the app itself, as well as implementation factors.

Table . Template analysis theme descriptions.

A priori themes	Description theme inclusions
Theme 1: usability	User-friendliness, ease of use, and efficiency, including navigation, interface, intuitiveness, and ideas for improved functionality
Theme 2: acceptability	Content suitability, relevance, alignment with expectations, and likelihood of recommending. Encompasses all content feedback
Theme 3: usefulness	Ability of the app to provide benefit/value add/impact. Usefulness in meeting needs/requirements. Effectiveness in delivering desired outcomes
Theme 4: feasibility	Practicalities of implementation including timing, barriers, enablers, and ongoing sustainability

Participants

Parents or caregivers (hereafter referred to as parents) were eligible to participate if their child (aged 0 - 18 y) had finished cancer treatment for any cancer diagnosis within the previous 14 months and was receiving medical cancer surveillance at The RCH's Children's Cancer Center (RCH CCC). Eligible clinicians included RCH CCC medical or specialist nursing staff who provided direct clinical care and coordination to families recruited into the study.

XSL•FO
RenderX

Procedure

Parent participants were identified by RCH CCC clinical staff, who introduced the study and provided an information letter. A member of the research team contacted the parent within 2 weeks to ascertain their willingness to participate. Medical and specialist nursing staff who were directly involved in the care of the recruited families were approached via email to participate in the study. All participants provided written consent. Following parent and HCP consent, the study team used the hospital appointment booking system to identify and record the date of the patient's next medical surveillance appointment.

Parents and HCPs were sent instructional emails detailing how to download and explore LotusLab, including how to enroll in the "Lotus Off Treatment" program. HCPs are integral to the successful implementation of digital health interventions. Therefore, as potential deliverers and trusted advocates, HCPs were asked to explore the parent (end user) interface and to evaluate the potential value and use of LotusLab to support their patients and families.

Once enrolled, participants had access to 15 topics, comprising videos and written information and resources tailored to parents and related to the EOT period. If available, the patient's personalized treatment summary and surveillance roadmap were uploaded to the app. All participants were encouraged to explore the LotusLab layout and content. In addition, parents were asked to complete an in-app survey called "About You and Your Child" to provide basic demographic data and to self-rate their digital literacy. Within the survey, parents provided details about their child's age, which triggered the in-app release of the corresponding parent proxy-reported Pediatric Quality of Life Inventory (PedsQL; version 4.0), a widely used, validated 23-item measure assessing child health-related quality of life across 4 dimensions: physical, emotional, social, and school functioning [35]. Participants received email and mobile notifications alerting them to the availability of the PedsQL for in-app completion.

Parents were asked to complete the PedsQL within 2 weeks prior to their child's appointment. Although the PedsQL is validated for self-reporting by children aged 5 to 18 years [35], this early-phase testing was designed to capture parents' experiences only and therefore included proxy reports for children aged 2 to 4, 5 to 7, 8 to 12, and 13 years and older. Further development and testing of the adolescent and young adult interface are ongoing.

PedsQL results were exported from the app by the study team, and subscale and total scores were manually calculated as per standardized scoring instructions. This procedure was undertaken due to app capability limitations at this phase of testing and the ongoing development of the clinician user interface. Preliminary testing focused on initial perceptions of parents and clinicians regarding the use of and discussion of the PROM in the real-world clinical setting. To support HCP understanding and interpretation of PedsQL domains and scores, and to aid consultation discussions, a custom PedsQL report was shared via email with the relevant HCP prior to the patient's upcoming appointment (Multimedia Appendix 1). Parents and HCPs were encouraged to discuss the results of the PedsQL during the medical consultation. A retrospective note audit was undertaken to capture any documentation of the PedsQL results and discussion.

Following the consultation, participants were invited to complete usability and acceptability surveys administered via REDCap (REDCap Consortium) software [36] and to take part in an online semistructured interview to discuss their experience with LotusLab.

Measures

Demographics

Brief demographic data related to patient and family factors (eg, diagnosis, age of the child, etc) were captured via an inbuilt survey titled "About You and Your Child." Additional treatment information or relevant dates were accessed via the electronic medical record (EMR) and stored securely. Information related to HCPs (eg, discipline, role, primary institution, and years of oncology experience) was gathered during the consent process, either verbally or via email.

Digital Health Care Literacy Scale

Parents' digital literacy level was captured to enhance the accuracy of usability testing and to interpret differences between user abilities and issues with the LotusLab design and functionality. The validated 3-item Digital Health Care Literacy Scale was selected to assess participants' comfort and confidence with mobile technology, specifically the ability to install and use apps and solve technical issues on their own [37].

System Usability Scale

Usability was measured and benchmarked with the System Usability Scale (SUS), a validated and widely used measure for assessing the usability of digital products [38,39]. The short 10-item questionnaire shows strong internal consistency and construct validity [40], and the benchmark mean usability score of 68 (SD 12.5) is considered appropriate for assessing the usability of digital health apps across different contexts [39]. Scores greater than 51 can be interpreted as "okay," scores of 75 and above can be interpreted as "good," and scores of 86 and above as "excellent" [41].

Adapted Abbreviated Acceptability Rating Profile

Acceptability was measured using a study-specific scale that was adapted from the 8-item abbreviated acceptability rating profile (AARP) [42]. Wording of the items was adapted to assess the app (eg, "I would be willing to recommend this app to other parents"), and items were scored on a 5-point Likert scale (strongly disagree, disagree, neutral, agree, and strongly agree; Multimedia Appendix 2).

Semistructured Interviews

Semistructured interviews were conducted with parents and HCPs to elicit in-depth perceptions of the app. Interviews were recorded, transcribed, and deidentified for analysis in NVivo (version 15), a software program for data analysis [43] (Multimedia Appendix 3).

XSL•FO
RenderX

Data Analysis

Descriptive statistics were used to summarize data from the measures outlined above. For the Digital Health Care Literacy Scale, frequencies of responses were calculated. SUS mean scores and score ranges were calculated to provide an overall indicator of usability; frequencies and percentages from the adapted AARP were summarized, describing the proportion of participants who rated the intervention as acceptable.

Template analysis was selected for qualitative analysis of semistructured interviews. Template analysis involves the development of a predetermined coding template that can be adapted in response to the data. It is considered a practical and rapid method well suited to iterative design cycles [44]. In our study, 4 a priori themes were selected to guide analysis and address the study objectives: usability, acceptability, usefulness, and feasibility. Although some additional subthemes were predetermined (eg, barriers and enablers to implementation), others were generated and validated using thematic analysis. To enhance reliability, 2 research team members (HP and MCM) independently coded 3 interviews with parents and HCPs and compared their coding. Consensus on the type of content that related to each a priori theme was reached (see Table 1 for definitions). Remaining transcripts were analyzed by HP with ongoing review by MCM.

Ethical Considerations

The study was approved by the RCH Human Ethics and Research Committee (#85334). Informed consent was obtained from all participants. All data were securely stored on servers and software approved for the use of protected health information. Additionally, data were deidentified upon collection and stored under a study identification number. Participants did not receive compensation for participation in this study.

Results

Study Participants

Eight parents (7 mothers) of children with a range of diagnoses, aged between 3 and 17 years, were approached and consented to participate in the study. All participants (100%) reviewed the app content and provided feedback via online surveys and semistructured interviews.

Twelve HCPs consented to the study; 3 did not respond to follow-up and did not participate in interviews or surveys (75% response rate). Nine HCPs (5 oncologists and 4 advanced practice nurses) reviewed the app content and completed the online usability and acceptability surveys. Eight participated in interviews; 1 HCP was unavailable due to extended leave. Participant characteristics are presented in Table 2.

XSL•FO
RenderX

Table . Participant characteristics.

Participant characteristics	Value, n
Parent characteristics (n=8)	
Parent/caregiver age (y)	
25 - 34	1
35 - 44	5
45 - 54	2
Parent sex	
Female	7
Male	1
Child's diagnosis	
Acute lymphoblastic leukemia	3
Brain/CNS[a] tumor	1
Solid tumors	4
Child's current age (y)	
0 - 5	2
6 - 12	2
13 - 17	4
Child's treatment type[b]	
Chemotherapy	8
Immunotherapy	1
Radiation therapy	3
Surgery	6
Region	
Metropolitan	7
Regional	1
Language other than English spoken at home	
No	7
Yes	1
Education level	
High school diploma	2
Bachelor's degree	3
Graduate certificate	1
Postgraduate degree	2
Phone type	
Android	2
iOS	6
Health care provider characteristics (n=9)	
Role	
Nurse practitioner	2
Consultant oncologist	5
Clinical nurse consultant	2
Age (y)	

XSL•FO
RenderX

Participant characteristics	Value, n
35 - 44	6
45 - 54	2
55 - 64	1
Sex	
Female	7
Male	2
Years of oncology experience (y)	
0 - 5	1
6 - 10	2
11 - 15	3
16 - 20	1
>20	2
Phone type	
Android	3
iOS	6

[a]CNS: central nervous system.

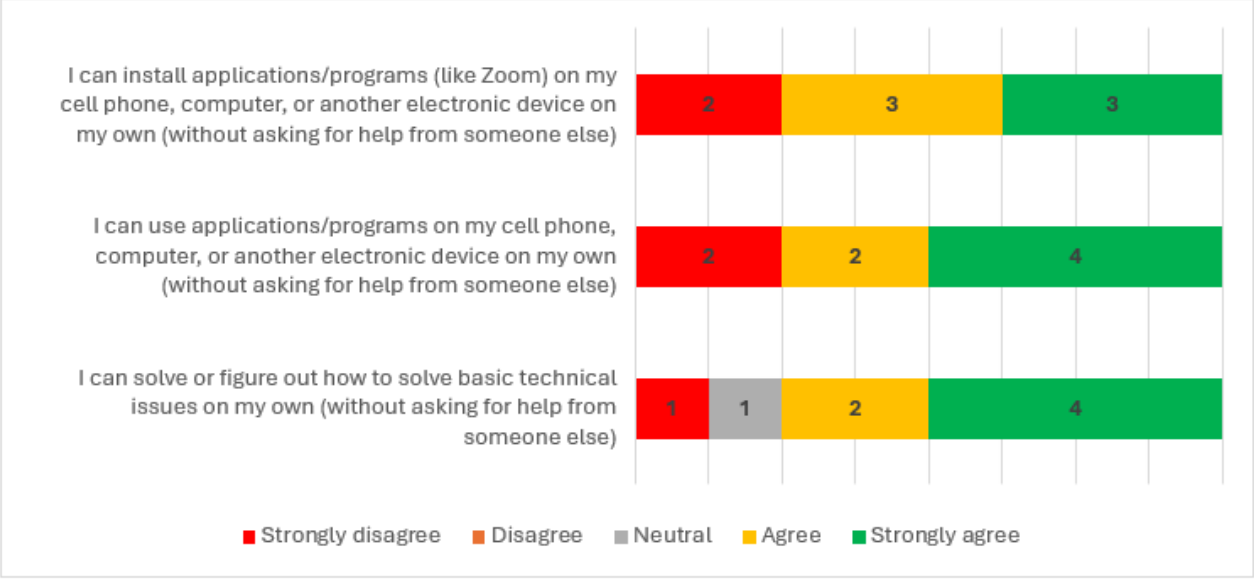
[b]A child can undergo multiple treatment types.

Parent Results

Digital Health Literacy Scale

Figure 3 presents digital literacy results. Six of the 8 parents (75%) "agreed" or "strongly agreed" that they could use or install an app on their device without asking for help and indicated that they could solve basic technical problems on their device. Two (25%) parents "strongly disagreed" that they could do this.

Figure 3. Parent Digital Health Literacy Scale summary.



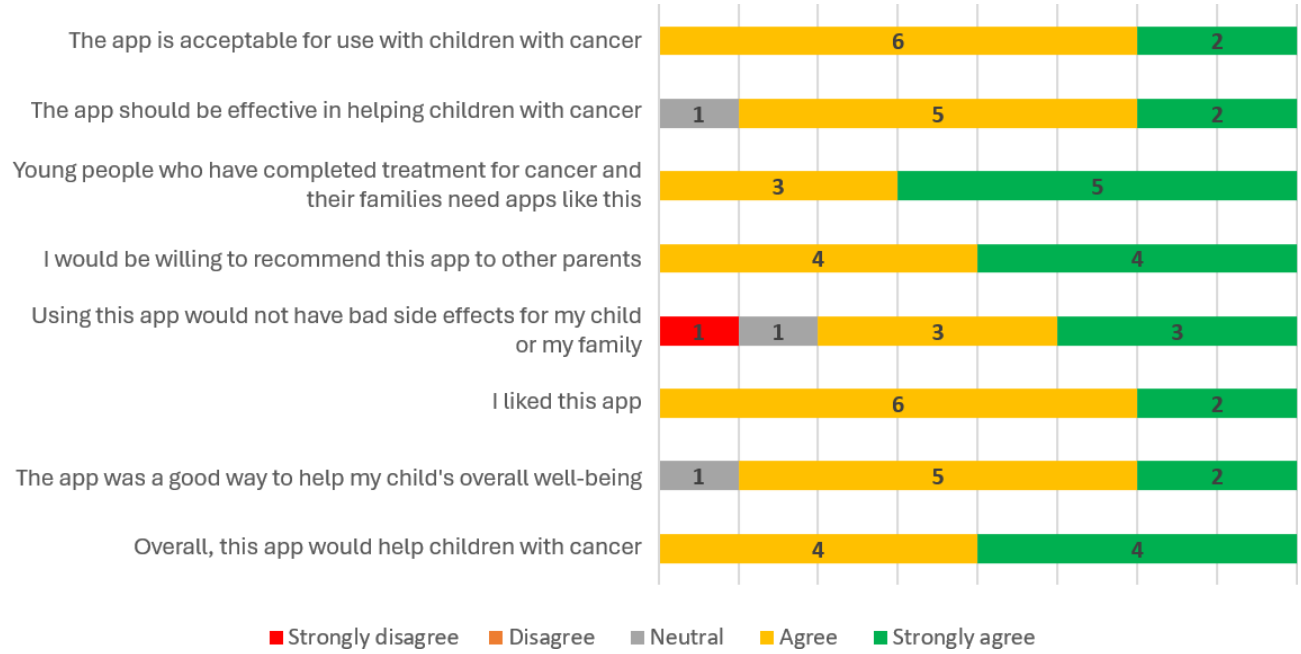
Usability and Acceptability of the App

Parents responded with a high mean SUS usability score of 86.88 (SD 8.74; range 75 - 100) compared with the benchmark usability score of 68 [39]. All parents "agreed" or "strongly agreed" that the app was "quick to learn" and that they felt confident using the app. Parents individually rated usability as either "excellent" (5/8) or "good" (3/8).

All parents "agreed" or "strongly agreed" that the app was fit for purpose and that a digital solution such as LotusLab was needed for young people who had recently completed treatment for cancer. Figure 4 shows parent ratings of LotusLab's acceptability.

XSL•FO
RenderX

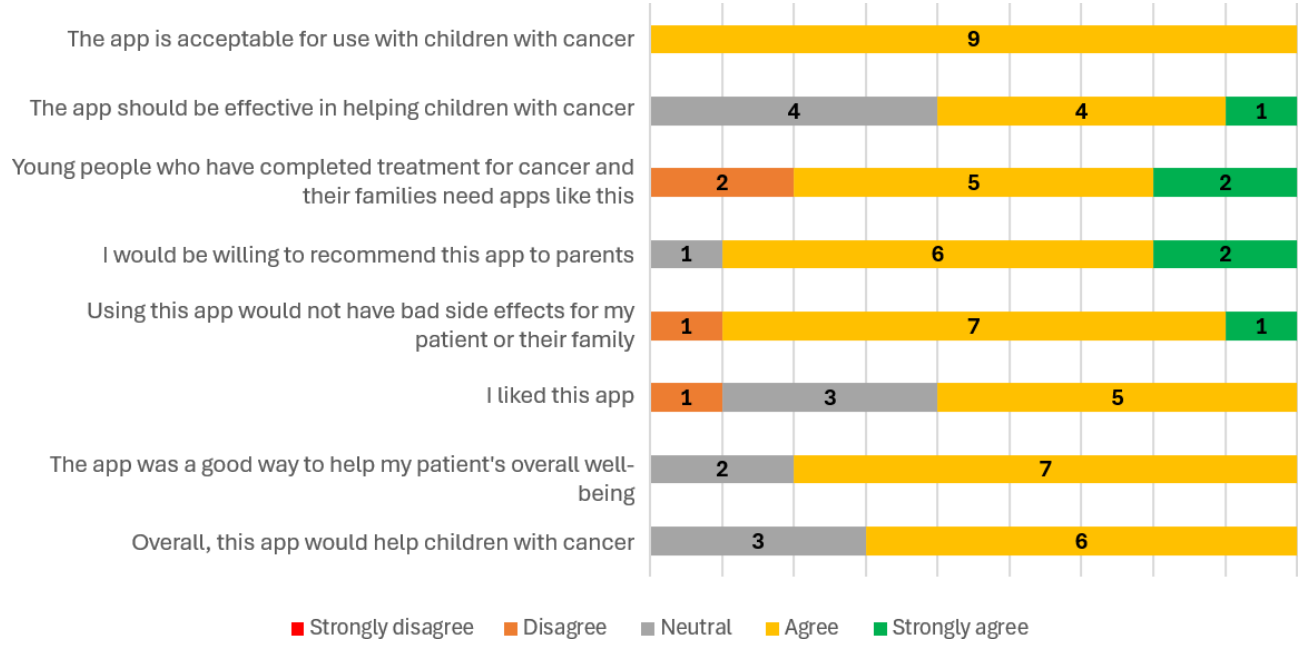
Figure 4. Parents' ratings of LotusLab acceptability.



HCP Results: Usability and Acceptability of the App

HCPs rated the usability of the parent-facing app in the "okay" range, with an overall mean SUS score of 66.39 (SD 10.16; range 45.0 - 77.5). All providers "agreed" that the app was "acceptable for use with children with cancer." Seven of 9 HCPs (77.8%) "agreed" or "strongly agreed" that young people who had completed treatment and their families need apps like this, and 8 (88.9%) "agreed" or "strongly agreed" that they would be willing to recommend the app to parents. Figure 5 shows HCP ratings of LotusLab acceptability.

Figure 5. Health care providers' ratings of LotusLab's acceptability.



Usage of the App-Embedded PROM (PedsQL)

All 8 parents completed the PedsQL measure embedded within the app. Four participants had a discussion with the oncologist regarding their questionnaire results. One parent stated that, although they did not discuss the questionnaire specifically, they felt the results informed the consultation, as areas of concern were brought up by the consultant and addressed. For 1 parent, the survey was brought up for discussion by the consultant; however, the parent "didn't bother going into it," as they felt a different HCP would be better placed to discuss the results. One parent was unsure whether the PedsQL was discussed, as their partner attended the consultation. One parent stated that the PedsQL was not discussed.

A retrospective review of EMR notes found that 4 of 8 clinicians documented the areas of need identified by the PedsQL and the outcomes of the discussion. One clinician documented the parent's participation in the study but did not directly reference

XSL•FO
RenderX

the PedsQL, and 3 clinicians did not reference the PedsQL in the medical record. Parents' experiences of completing the questionnaire within the app and discussing the PedsQL are captured in the semistructured interviews, as are HCPs' experiences of receiving and using the PedsQL results. Reporting of PedsQL scores is outside the scope of this study.

Qualitative Interview Results

Interview Characteristics

Semistructured individual interviews were conducted online (using Microsoft Teams) with both parents and HCPs. Interviews ranged from 19 to 41 minutes and were conducted between October 2023 and April 2024. Refer to Table 1 for a priori theme definitions and Table 3 for themes, subthemes, and exemplar quotes from parents and HCPs.

XSL•FO
RenderX

Table . Exemplar quotes from semistructured interviews with parents and health care providers (HCPs).

Themes and subthemes	Quotes
Theme 1: usability	
Navigation	"Usability...was spot-on...it seemed to be...very clear and clean." (Parent 6)
	"I liked the idea of all the activities which allowed people to pick and choose the ones that they liked to do." (HCP 10)[a]
Look and feel/design	"...whilst aesthetically some people might think it's a bit basic I actually just found that the basic works so why mess with it?" (Parent 1)
	"...more navigational like a 'Choose Your Own Adventure'..." (HCP 7)
Functionality	"...I can't see the completed activities. I think that's – definitely should be fixed because it would be good to go back and redo some of these or just as a bit of a memory trigger." (Parent 6)
	"I do think that search function would be valuable." (HCP 5)
Theme 2: acceptability	
New content suggestions	"Quick links...I've got some patients now off therapy a year and they're asking me about immunisation...on a weekly basis." (HCP 11)[a]
Content relevancy	"All the videos are really good, especially all the videos that had parents...just made everything so much more relatable and also really calming..." (Parent 3)
	"...the things that are more of value in the app...you know the emotional development of kids or the emotional responses that they might have after finishing treatment." (HCP 1)
Content accessibility	"The videos were good perhaps from another language perspective too. Some people find it easier to listen than to read." (Parent 2)
	"...obviously the topics are good, but if someone is not as health literate then the text heavy thing actually might be a barrier to them..." (HCP 2)
AYA[b] acceptability	"...it depends on the kid...if it's really wordy with modules and things they might feel like it's homework..." (Parent 4)
	"...I think there's some really good key bits but...for a young person I think would look very different." (HCP 7)
Overall experience	"...I was like oh god, I wish I had this app 12 months ago because...I could have just gone to the app and navigated my way through and got the information that I needed." (Parent 1)
Theme 3: usefulness	
Benefits/value add	"I think that this is like a really nice CNC [Clinical Nurse Consultant] in your pocket...you can just refer back to it." (Parent 3)
	"...a nice kind of point of connection still between our service and then coming off treatment...I think, to give them a bit of security..." (HCP 2)
PedsQL[c] usefulness	"It's quite meaningful to be asked those kinds of questions...really nice to feel like...someone's checking in with us." (Parent 3)
	"...a really nice kind of segway...I've got this survey I can see that sleep, pain and mood are a problem...So I think that using them in a consult...would be super helpful..." (HCP 3)
	"...it can take up a lot of time and you're not in the right position to solve a lot of those issues either..." (HCP 6)
Treatment summary and surveillance roadmap	"But if it's there on an app and you can say look, they said I was to do this or do this...and if you can show them it's going to make things so much easier. So it would be really good to have access." (Parent 6)
	"...he can just go like this and say this is my diagnosis and my treatment. Wouldn't that empower him? Wouldn't that have been amazing?" (HCP 11)[a]
Theme 4: feasibility	

XSL•FO
RenderX

Themes and subthemes	Quotes
Barriers to implementation and ongoing engagement	"I know that my son has been resistant to...Who have I got to talk to now?...I'm ready to move on with the next phase of my life." (Parent 4) "...like I've got loads of apps on my phone...like I turn off a lot of notifications...I think that's the tricky thing about an app." (HCP 3)
Enablers to implementation and ongoing engagement	"...I don't know if she [CNC] has like an orientation checklist...or maybe even like a welcome email from her where maybe she...can just send out the link to this app maybe." (Parent 3) "I think it's just the exposure to it...the CNCs who have the relationships with the family...getting it into the Facebook groups and all of those sorts of things, that's where it'll all take-off..." (HCP 10)[a]
Timing and introduction	"I know it's an app for after treatment but for me it was very lonely during treatment...I think if there was a whole aspect of the app where it was what to expect going into it 'cause it's so scary and you've got no idea." (Parent 5)

[a]A total of 12 HCPs were recruited, and 8 participated in interviews; the HCP number aligned with the quote represents the number allocated during sequential recruitment.

[b]AYA: adolescent and young adult.

[c]PedsQL: Pediatric Quality of Life Inventory.

Theme 1: Usability

Parents felt that LotusLab was easy to navigate and appreciated the "clean" and "clear" look and feel. They valued the ability to choose topics according to personal interest and importance. Practical suggestions for improvement centered on the ability to view completed topics and PROM results once submitted, the ability to share content with others (eg, friends, medical team, and community health), as well as a search facility.

Overall, HCPs also found the app easy to navigate; however, they had more suggestions about the app layout, suggesting topic groupings such as "emotions" and a "more colorful and tiles layout" could help make it easier to navigate. HCPs echoed parents' comments that the ability to share content and reaccess information through a search function and "Quick Links" sections would enhance the user experience.

Theme 2: Acceptability

Parents commented positively on the relevant and relatable content. Two parents stated they wished they had access to the app earlier. There were many suggestions for additional content (eg, development of a frequently asked questions section, information on how parents can support adolescents, and the need for additional videos from the perspectives of children and young people). Specifically, 2 parents reflected on the need for additional content for young people diagnosed with brain tumors, addressing the impacts of cognitive impairment on returning to school and maintaining friendship groups. Parents provided positive feedback about the video content and its ability to enhance accessibility for users with low general and health literacy levels.

HCPs provided positive feedback on the app content, particularly in relation to the videos and psychosocial and emotional needs information at the EOT. There was feedback that additional app content would be needed to enhance its relevance and suitability for patients who had undergone bone marrow transplant and those with brain tumors, due to specific medical and psychosocial needs. HCPs felt that while an app is an appropriate mechanism for young people to access information, LotusLab required enhancements to make it more engaging for young people, such as reducing the amount of text, reordering content to ensure videos always appear at the top, and making it look less like a "learning module." HCPs suggested that information relating to the LotusLab sign-up process should also be available in languages other than English to ensure equity of access.

Theme 3: Usefulness

Parents felt the app provided additional value by collating trusted information in one place (taking out the "legwork"), providing education, alleviating concerns, and allowing people to digest information in their own time.

All parents reported that they found the technical process of completing the PedsQL within the app easy. In addition, most found the process meaningful as a mechanism for "checking in." Two participants did not discuss the survey results with their oncologists, as they felt that these medical practitioners were not best placed to support their identified needs and that it would have been more beneficial to discuss the results with a different member of the health care team, for example, an occupational therapist. One parent found some of the specific survey questions challenging to answer. They felt some questions were not appropriate to their child's current situation due to the long-term and evolving cognitive impacts of their diagnosis and treatment. This participant cautioned about the importance of delivering surveys like the PedsQL on a case-by-case basis to ensure relevance to a family's current situation.

Some participants had difficulty accessing their treatment summaries and roadmaps uploaded to the "My Files" section of the app. Despite this challenge, all parents felt that access to these documents would be useful, especially when visiting other health professionals or when they were unable to access the existing EMR patient portal.

HCPs felt the value of LotusLab centered on its ability to address identified gaps by providing connections to information at the

XSL•FO
RenderX

EOT, and they regarded the app as "a good bridge." In addition, HCPs considered that the app could help normalize experiences. Multiple HCPs felt that LotusLab could potentially reduce the need for parents to consult Google and save HCPs time in answering common questions.

HCPs responded positively to the concept of using the app for PROMs, streamlining the data collection process, and using the results as a mechanism for meaningful discussion of broader quality of life issues and concerns. However, HCPs flagged potential concerns about their capacity to review results captured in the app prior to a consultation, and about the risk of identifying needs but feeling they were "not in the right position to solve a lot of those issues." They acknowledged that identifying a need that is not addressed introduces risk and could potentially result in a negative experience for patients and caregivers.

HCPs noted that access to personalized survivorship documents, such as a treatment summary, could be empowering, especially for young people who do not always have ownership of these documents. One clinician cited specific examples from their career when access to a treatment summary would have been beneficial for both the treating clinician and the patient, for example, when a childhood cancer survivor presented as an adult to an emergency department but had "no idea" which cancer they had or what treatment they had received for their cancer.

Theme 4: Feasibility

The interviews explored in detail the potential barriers to engagement and sustainability of the app. Some parents felt they would be less likely to continue engaging with the app the further off treatment they were, or when feeling "ready to get back to normal" and "not wishing to revisit the past." Two participants noted that there was a risk the app could feel like "homework," which could reduce engagement among young people.

Parent-identified enablers to engagement and sustainability included promotion of the app by staff, such as specialist nurses, who know the families well and are best placed to discuss what to expect during the EOT period. Parents noted that embedding the app within existing discussions, such as "EOT reviews," would be helpful and suggested the app could be advertised through posters in waiting areas.

There were mixed perspectives on the timing of the introduction to LotusLab. Although LotusLab was designed to provide EOT information, many parents felt that access to the app would be beneficial in the months leading up to completion of treatment. One parent felt that introduction to the app should be on a case-by-case basis, as the EOT period can be complex and a specific time point can be hard to determine, but most agreed on the value of knowing what to expect. It was also suggested by several participants that the app could have value during active treatment, lending itself well to "different starts."

HCPs identified barriers to engagement and sustainability, including the risk that only a self-selecting cohort of people would download and access the app and that those with the greatest unmet information needs might require additional support to engage with the app. One HCP noted that many people turn off their notifications; therefore, ongoing engagement could be challenging. In relation to the collection of PROMs, multiple HCPs raised concerns about the capacity to review survey results prior to consultations due to time constraints, as well as safety concerns centered on ensuring that all results are reviewed and acted upon by an appropriate professional in a timely manner.

HCP-suggested enablers also centered on embedding the implementation of LotusLab within existing roles or services, for example, integrating it into a survivorship clinical nurse coordinator role and including it in "new patient packs." One HCP suggested that the app could be shared through family networks, such as social media groups. One HCP noted that, in order to streamline LotusLab implementation, it could be helpful to undertake robust change management processes with staff. As per parent feedback, HCPs also suggested creating "a version of this at diagnosis" so that the app could become a longer-term companion across the cancer pathway.

Discussion

Principal Findings

This study assessed parents' and HCPs' perceptions and experiences of a newly developed app and generated key insights for app refinement, further development, and successful implementation within real-world clinical settings.

Overall, survey and interview data indicated that LotusLab is easy to use and acceptable, and is particularly valuable for the provision of survivorship information. Exploration of usefulness and feasibility highlighted the potential benefits of using LotusLab in bridging care gaps and building engagement during the EOT period. Study participants recommended practical suggestions to enhance the user experience and encourage ongoing engagement.

Throughout the development of LotusLab, parent perspectives as end users have been central; however, effective co-design incorporates input from all stakeholders [45]. In this study, we found similarities and differences in feedback between parents and HCPs. Both cohorts found the app easy to use, even for 2 parents who self-rated their digital literacy as low, and both cohorts suggested similar enhancements to improve usability: primarily a search function, additional content for specific cohorts, and quick links. Interestingly, the HCPs scored overall app usability more variably and acceptability lower than the parents, and 1 clinician did not like LotusLab, providing feedback that they personally did not like using apps in general. HCPs provided more feedback on the design and layout of LotusLab, whereas parent feedback focused heavily on the app content. These findings could reflect the fact that the interface and content were designed with and for parents and not HCPs. Higher parent scores could also reflect the significant unmet information needs experienced by families at the EOT [2,3,8], resulting in parents prioritizing access to information over appearance. This notion of significant unmet needs was further supported by several parents who stated they wished LotusLab had been available to them earlier in the care delivery pathway.

XSL•FO
RenderX

PROMs and proxy-reported outcome measures are increasingly identified as critical components of pediatric cancer care. During this early-phase testing, our aim was to explore the potential benefits of developing a PROM interface and how this could be used to support families at the EOT. Although both parents and HCPs placed significant value on being able to complete the PedsQL within LotusLab, both raised concerns about the ability of HCPs to review and adequately act on identified needs. These findings are echoed in other studies exploring the use of PROMs in pediatric oncology settings, identifying challenges with HCP time pressures, lack of workflow integration, and concerns over the inability to translate results into tangible actions and improved outcomes for families [46-48]. Although clinicians found the provision of meaningful summaries helpful in guiding consultations, this approach did present an additional administrative burden for other team members. Although not specific to the EOT, to address these challenges, apps such as Oncology Hub [21] have opted to deliver algorithm-driven tailored information directly to patients, based on PROM scores at the point of completion and time of need.

From a local implementation perspective, further exploration of the LotusLab algorithm's capabilities and efficient information exchange is warranted. Our findings highlight that health service capacity and approaches to sharing identified issues and concerns are key considerations for the safe, sustainable, and best-practice management of PROMs through apps such as LotusLab, particularly during the EOT and survivorship phases.

Overall, positive findings relating to LotusLab's usability and acceptability align with other studies of mobile apps for cancer survivors [14,49]. There was agreement that the most significant value of LotusLab was its provision of practical and psychosocial information at the EOT. Parents and HCPs felt the app had the potential to reduce the burden on young people, parents, and clinicians by providing easy access to trusted information during a time associated with significant uncertainty.

Both cohorts identified barriers to long-term app engagement in alignment with a systematic review of smartphone applications for cancer survivors [14], for example, the risk of diminishing engagement over time. This study suggests that the unique feature of inclusion of a personalized treatment summary and roadmap within the app could encourage repeat use and empower young people to actively manage their future health care when no longer in the pediatric setting. Other barriers included users turning off key app features (eg, notifications) and treatment side effects (eg, cognitive impacts) affecting users' ability to engage with the app [14].

The impact of digital health apps on patient and family survivorship outcomes is poorly understood, often attributed to a lack of relevant randomized controlled trials [14,50]. In addition, many apps in health care settings do not progress past the initial pilot phase. This is often due to limited, short-term funding that affects the capacity to collect long-term outcomes and cost-effectiveness data, preventing the demonstration of value for both health services and families [14,51-53]. As part of our co-design and evaluation process, we aimed to proactively identify and mitigate potential barriers to sustainability and maximize future impact. A larger "real-world" implementation phase is required to fully quantify the value and impact of LotusLab. Implementing LotusLab within existing systems and workflows, as suggested by participants and similar studies [29], and embedding the app within the standard of care could provide an avenue for sustainable operational costs and increased app success. In addition, LotusLab has the potential to support families across the treatment trajectory (not just EOT), thereby providing a user-friendly way to collect important long-term outcomes data and build the case to support ongoing delivery.

Limitations

This is a small, single-institution study limited to parents and HCPs. Adolescent and young adult (AYA) perceptions were not evaluated during this specific phase. Engagement with young people to develop age-appropriate content and tailor the app to meet their specific needs is currently being undertaken. In addition, a proxy-reported tool was used in this study; therefore, exploration of AYA perspectives on the use of LotusLab for the collection of PROMs is warranted.

Although some existing content is relevant to all families, further stakeholder engagement would be required for implementation in other health service settings and to increase relevance to specific tumor streams and diagnoses. While there are advantages in using generic quality of life measures like the PedsQL, these tools may lack sensitivity related to disease-specific symptoms [26], a finding corroborated by one parent in this study. This finding corresponds to the broader challenge of developing one digital health intervention for a heterogeneous population of childhood cancer survivors and highlights the significance of including tailored measures and content to provide value and improve the user experience.

All study participants were English-speaking; therefore, feedback is not generalizable to families from culturally and linguistically diverse backgrounds. To support the broader LotusLab project goal to increase equitable access to information at the EOT, the translation of app content into other languages is a key component of "Phase 5: Refinement and Sustainability" (Figure 1). Phase 5 also includes the development and testing of audio summaries in languages other than English.

Although the collection and sharing of outcomes data from the PedsQL were considered valuable, further work is required to build LotusLab's capability to effectively house and address PROM results. At this stage of β-testing, the app had technical limitations related to automated scoring and the generation of clinically meaningful reports. At a systems level, there are known limitations that affect the collection, sharing, and clinical use of electronic PROMs, such as poor platform interoperability and difficulty integrating PROMs into the EMR [48]. We will continue exploring the optimal approach to using LotusLab to support the integration of PROMs in the pediatric EOT setting.

Conclusions

Initial β-testing data indicate that LotusLab is easy to use, acceptable, and valuable for parents. Parents and HCPs responded positively to using the app to provide trusted

XSL•FO
RenderX

information and contributed actionable insights and practical feedback related to enhancements, PROM use, and app implementation and sustainability. Next steps include app refinement, ongoing engagement with AYA, expansion of content development, integration within real-world EOT models of care, and exploration of effectiveness and outcomes.

Acknowledgments

The authors acknowledge the support provided by the Victorian Paediatric Integrated Cancer Service and sincerely thank the participating parents and health care providers. No AI tools were used in this study.

Funding

The funding for this study is provided by a grant awarded to MCM from the Cancer Trials Management Scheme Competitive Grants Program (2021-2025), administered by Cancer Council Victoria and the Victorian Government, through the Victorian Cancer Agency. MCM is supported by a Victorian Cancer Agency Fellowship (MCRF 20042) and the My Room Children's Cancer Charity.

Data Availability

The datasets will not be publicly available for privacy reasons but may be obtained from the project's principal investigator (MCM) upon reasonable request.

Authors' Contributions

MCM and CRDL conceived the study design, devised the study protocol, assembled the measures, and provided scientific project management. HP conducted the study and drafted the manuscript, with critical review from MCM. CW, CRDL, GT, and DDE contributed to the study design and implementation, provided content expertise and project management, and critically reviewed the manuscript. MN contributed expertise in digital innovation and sustainability planning. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Custom Pediatric Quality of Life Inventory (PedsQL) template.
[PNG File, 408 KB - cancer_v12i1e90177_app1.png]

Multimedia Appendix 2
Adapted acceptability survey for parents and health care providers.
[PDF File, 103 KB - cancer_v12i1e90177_app2.pdf]

Multimedia Appendix 3
Interview guide for parents and health care providers.
[PDF File, 135 KB - cancer_v12i1e90177_app3.pdf]

References

1. Signorelli C, Wakefield CE, Fardell JE, et al. The impact of long-term follow-up care for childhood cancer survivors: a systematic review. Crit Rev Oncol Hematol 2017 Jun;114:131-138. [doi: 10.1016/j.critrevonc.2017.04.007] [Medline: 28477741]
2. The hidden health crisis: children's cancer needs more than medicine. : Redkite; 2021 URL: https://www.redkite.org.au/wp-content/uploads/2021/10/Redkite-Report-The-Hidden-Health-Crisis.pdf [accessed 2025-12-08]
3. McCarthy MC, McNeil R, Drew S, Orme L, Sawyer SM. Information needs of adolescent and young adult cancer patients and their parent-carers. Support Care Cancer 2018 May;26(5):1655-1664. [doi: 10.1007/s00520-017-3984-1] [Medline: 29222598]
4. Wakefield CE, Butow P, Fleming CAK, Daniel G, Cohn RJ. Family information needs at childhood cancer treatment completion. Pediatr Blood Cancer 2012 Apr;58(4):621-626. [doi: 10.1002/pbc.23316] [Medline: 21910212]
5. McCarthy MC, McNeil R, Drew S, et al. Psychological distress and posttraumatic stress symptoms in adolescents and young adults with cancer and their parents. J Adolesc Young Adult Oncol 2016 Dec;5(4):322-329. [doi: 10.1089/jayao.2016.0015] [Medline: 27214245]

XSL•FO
RenderX

6. Wakefield CE, McLoone J, Butow P, Lenthen K, Cohn RJ. Support after the completion of cancer treatment: perspectives of Australian adolescents and their families. Eur J Cancer Care (Engl) 2013 Jul;22(4):530-539. [doi: 10.1111/ecc.12059] [Medline: 23730980]
7. Lea S, Martins A, Bassett M, et al. Issues experienced and support provided to adolescents and young adults at the end of active treatment for cancer: a rapid review of the literature. Eur J Cancer Care (Engl) 2018 Nov;27(6):30485604. [doi: 10.1111/ecc.12972] [Medline: 30485604]
8. Lea S, Martins A, Fern LA, et al. The support and information needs of adolescents and young adults with cancer when active treatment ends. BMC Cancer 2020 Jul 28;20(1):697. [doi: 10.1186/s12885-020-07197-2] [Medline: 32723357]
9. Keller MC, Needham A, Holden E, et al. We are done! Now what? Exploring end of treatment needs of childhood cancer survivors and their parents. J Pediatr Hematol Oncol Nurs 2024;41(2):96-106. [doi: 10.1177/27527530231194598] [Medline: 38238967]
10. Psihogios AM, Schwartz LA, Deatrick JA, et al. Preferences for cancer survivorship care among adolescents and young adults who experienced healthcare transitions and their parents. J Cancer Surviv 2019 Aug;13(4):620-631. [doi: 10.1007/s11764-019-00781-x] [Medline: 31273639]
11. Myers L, Johnston EA, Zajdlewicz L, et al. What are the mechanisms underlying the delivery of survivorship care information in Australia? A realist review. Psychooncology 2024 Mar;33(3):e6321. [doi: 10.1002/pon.6321] [Medline: 38488825]
12. Sievers Y, Roser K, Scheinemann K, Michel G, Ilic A. The information needs of relatives of childhood cancer patients and survivors: a systematic review of quantitative evidence. Patient Educ Couns 2024 Sep;126:108316. [doi: 10.1016/j.pec.2024.108316] [Medline: 38788309]
13. Landier W, Campos Gonzalez PD, Henneberg H, et al. Children's Oncology Group KidsCare smartphone application for parents of children with cancer. Pediatr Blood Cancer 2023 Jun;70(6):e30288. [doi: 10.1002/pbc.30288] [Medline: 36945123]
14. Changrani K, Chima S, Sharma A, et al. A systematic review of smartphone applications for cancer survivors. J Cancer Surviv 2024 Dec;18(6):1951-1973. [doi: 10.1007/s11764-023-01435-9] [Medline: 37700151]
15. Slater PJ, Fielden PE, Bradford NK. The Oncology Family App: providing information and support for families caring for their child with cancer. J Pediatr Oncol Nurs 2018;35(2):94-102. [doi: 10.1177/1043454217741874] [Medline: 29161974]
16. Mueller EL, Cochrane AR, Campbell ME, Nikkhah S, Miller AD. An mHealth app to support caregivers in the medical management of their child with cancer: co-design and user testing study. JMIR Cancer 2022 Mar 16;8(1):e33152. [doi: 10.2196/33152] [Medline: 35293867]
17. Schwartz LA, Psihogios AM, Henry-Moss D, et al. Iterative development of a tailored mHealth intervention for adolescent and young adult survivors of childhood cancer. Clin Pract Pediatr Psychol 2019;7(1):31-43. [doi: 10.1037/cpp0000272] [Medline: 38882593]
18. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: a systematic review of interventions across the cancer continuum. Psychooncology 2020 Jan;29(1):17-37. [doi: 10.1002/pon.5280] [Medline: 31692183]
19. King-Dowling S, Psihogios AM, Hill-Kayser C, et al. Acceptability and feasibility of survivorship care plans and an accompanying mobile health intervention for adolescent and young adult survivors of childhood cancer. Pediatr Blood Cancer 2021 Mar;68(3):e28884. [doi: 10.1002/pbc.28884] [Medline: 33416214]
20. Elsbernd A, Hjerming M, Visler C, et al. Cocreated smartphone app to improve the quality of life of adolescents and young adults with cancer (Kræftværket): protocol for a quantitative and qualitative evaluation. JMIR Res Protoc 2018 May 10;7(5):e10098. [doi: 10.2196/10098] [Medline: 29748162]
21. Bradford N, Slater P, Fielden P, et al. Care at your fingertips: codesign, development, and evaluation of the Oncology Hub app for remote symptom management in pediatric oncology. JCO Clin Cancer Inform 2023 Jan;7:e2200134. [doi: 10.1200/CCI.22.00134] [Medline: 36706346]
22. Wesley KM, Fizur PJ. A review of mobile applications to help adolescent and young adult cancer patients. Adolesc Health Med Ther 2015;6(141–8):141-148. [doi: 10.2147/AHMT.S69209] [Medline: 26316835]
23. Charbonneau DH, Hightower S, Katz A, et al. Smartphone apps for cancer: a content analysis of the digital health marketplace. Digit Health 2020;6:2055207620905413. [doi: 10.1177/2055207620905413] [Medline: 32110428]
24. Ardito V, Golubev G, Ciani O, Tarricone R. Evaluating barriers and facilitators to the uptake of mHealth apps in cancer care using the Consolidated Framework for Implementation Research: scoping literature review. JMIR Cancer 2023 Mar 30;9:e42092. [doi: 10.2196/42092] [Medline: 36995750]
25. Horan MR, Sim JA, Krull KR, et al. Ten considerations for integrating patient-reported outcomes into clinical care for childhood cancer survivors. Cancers (Basel) 2023 Feb 6;15(4):36831370. [doi: 10.3390/cancers15041024] [Medline: 36831370]
26. Adams M, Taylor RM. Developing patient reported outcome measures for children with cancer: a current perspective. EJC Paediatr Oncol 2025 Dec;6:100312. [doi: 10.1016/j.ejcped.2025.100312]
27. Jibb LA, Sivaratnam S, Hashemi E, et al. Parent and clinician perceptions and recommendations on a pediatric cancer pain management app: a qualitative co-design study. PLOS Digit Health 2023 Nov;2(11):e0000169. [doi: 10.1371/journal.pdig.0000169] [Medline: 38019890]

XSL•FO
RenderX

28. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
29. Poort H, Ryan A, MacDougall K, et al. Feasibility and acceptability of a mobile phone app intervention for coping with cancer as a young adult: pilot trial and thematic analysis. J Med Internet Res 2021 Jun 11;23(6):e25069. [doi: 10.2196/25069] [Medline: 34114957]
30. Noorbergen TJ, Adam MTP, Teubner T, Collins CE. Using co-design in mobile health system development: a qualitative study with experts in co-design and mobile health system development. JMIR mHealth uHealth 2021 Nov 10;9(11):e27896. [doi: 10.2196/27896] [Medline: 34757323]
31. Clemensen J, Rothmann MJ, Smith AC, Caffery LJ, Danbjorg DB. Participatory design methods in telemedicine research. J Telemed Telecare 2017 Oct;23(9):780-785. [doi: 10.1177/1357633X16686747] [Medline: 28027678]
32. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
33. ISO 9241-11:2018 Ergonomics of human-system interaction—Part 11: Usability: definitions and concepts. : International Organization for Standardization (ISO); 2018 URL: https://preview.sist.si/sist-preview/63500/33c267a5a7564f298f02bbd65721a181/ISO-9241-11-2018.pdf [accessed 2026-08-18]
34. Nielson J. Usability Engineering: Academic Press; 1993.
35. Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQL in pediatric cancer: reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. Cancer 2002 Apr 1;94(7):2090-2106. [doi: 10.1002/cncr.10428] [Medline: 11932914]
36. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
37. Nelson LA, Pennings JS, Sommer EC, Popescu F, Barkin SL. A 3-item measure of digital health care literacy: development and validation study. JMIR Form Res 2022 Apr 29;6(4):e36043. [doi: 10.2196/36043] [Medline: 35486413]
38. Brooke J. SUS: a “quick and dirty” usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry: CRC Press; 1996:189-194.
39. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale benchmarking for digital health apps: meta-analysis. JMIR mHealth uHealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
40. Khan Q, Hickie IB, Loblay V, et al. Psychometric evaluation of the System Usability Scale in the context of a childrearing app co-designed for low- and middle-income countries. Digit Health 2025;11:20552076251335413. [doi: 10.1177/20552076251335413] [Medline: 40357428]
41. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009 Apr;4(3):114-123 [FREE Full text]
42. Tarnowski KJ, Simonian SJ. Assessing treatment acceptance: the abbreviated acceptability rating profile. J Behav Ther Exp Psychiatry 1992 Jun;23(2):101-106. [doi: 10.1016/0005-7916(92)90007-6] [Medline: 1460097]
43. NVivo (Version 15). Lumivero. 2023. URL: https://lumivero.com/products/nvivo/ [accessed 2026-06-24]
44. Brooks J, McCluskey S, Turley E, King N. The utility of template analysis in qualitative psychology research. Qual Res Psychol 2015 Apr 3;12(2):202-222. [doi: 10.1080/14780887.2014.955224] [Medline: 27499705]
45. Sacre A, Godfrey A. Co-designing the future: end-user involvement in digital interventions. NPJ Digit Med 2026 Feb 14;9(1):160. [doi: 10.1038/s41746-026-02444-5] [Medline: 41688791]
46. McCabe E, Rabi S, Bele S, Zwicker JD, Santana MJ. Factors affecting implementation of patient-reported outcome and experience measures in a pediatric health system. J Patient Rep Outcomes 2023 Mar 9;7(1):24. [doi: 10.1186/s41687-023-00563-1] [Medline: 36892738]
47. Lyu J, Zhang H, Wang H, et al. Facilitators and barriers to implementing patient-reported outcomes in clinical oncology practice: a systematic review based on the consolidated framework for implementation research. Implement Sci Commun 2024 Oct 29;5(1):120. [doi: 10.1186/s43058-024-00654-0] [Medline: 39473015]
48. Rammant E, Ramsey I, M de Ligt K, et al. Barriers and facilitators for implementing patient-reported outcome measures in oncology practices: an umbrella review. Qual Life Res 2026 Apr 10;35(5):129. [doi: 10.1007/s11136-026-04205-z] [Medline: 41961151]
49. Schilstra CE, Sansom-Daly UM, Schaffer M, et al. “We have all this knowledge to give, so use us as a resource”: partnering with adolescent and young adult cancer survivors to determine consumer-led research priorities. J Adolesc Young Adult Oncol 2022 Apr;11(2):211-222. [doi: 10.1089/jayao.2021.0052] [Medline: 34297633]
50. Livingston PM, Heckel L, Orellana L, et al. Outcomes of a randomized controlled trial assessing a smartphone Application to reduce unmet needs among people diagnosed with CancEr (ACE). Cancer Med 2020 Jan;9(2):507-516. [doi: 10.1002/cam4.2718] [Medline: 31769226]
51. An Q, Kelley MM, Hanners A, Yen PY. Sustainable development for mobile health apps using the human-centered design process. JMIR Form Res 2023 Aug 25;7:e45694. [doi: 10.2196/45694] [Medline: 37624639]

XSL•FO
RenderX

52. Shaffer KM, Turner KL, Siwik C, et al. Digital health and telehealth in cancer care: a scoping review of reviews. Lancet Digit Health 2023 May;5(5):e316-e327. [doi: 10.1016/S2589-7500(23)00049-3] [Medline: 37100545]
53. Davis SW, Oakley-Girvan I. Achieving value in mobile health applications for cancer survivors. J Cancer Surviv 2017 Aug;11(4):498-504. [doi: 10.1007/s11764-017-0608-1] [Medline: 28342093]

Abbreviations

AARP: Abbreviated Acceptability Rating Profile
AYA: adolescent and young adult
CCC: Children's Cancer Centre
EMR: electronic medical record
EOT: end-of-treatment
HCP: health care provider
PedsQL: Pediatric Quality of Life Inventory
PROM: patient-reported outcome measure
RCH: The Royal Children's Hospital
SUS: System Usability Scale

Edited by M Balcarras; submitted 22.Dec.2025; peer-reviewed by M Doig; revised version received 12.Jun.2026; accepted 03.Jul.2026; published 07.Sep.2026.

Please cite as:
Pring H, De Luca CR, Taylor G, Williams C, Nagarajah M, Eisenstat DD, McCarthy MC
"Like a Clinical Nurse Consultant in Your Pocket"—Parents' and Health Care Providers' Perceptions of a Co-Designed Smartphone App Following Treatment for Pediatric Cancer: Mixed Methods Evaluation Study
JMIR Cancer 2026;12:e90177
URL: https://cancer.jmir.org/2026/1/e90177
doi:10.2196/90177

© Hannah Pring, Cinzia R De Luca, Georgia Taylor, Chris Williams, Megan Nagarajah, David D Eisenstat, Maria C McCarthy. Originally published in JMIR Cancer (https://cancer.jmir.org), 7.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Acceptance, Perceived Usefulness, and Data Sharing in Mobile Health Apps Among Patients With Breast Cancer: Cross-Sectional Survey Study

Madeleine Flaucher[1], MSc; Tabea Ott[2,3], Dr theol; Michael Nissen[1], Dr-Ing; Peter Dabrock[4], Prof Dr theol; Hanna Huebner[5,6,7,8], PD Dr rer nat; Matthias W Beckmann[5,6,7,8], Prof Dr med; Peter A Fasching[5,6,7,8], Prof Dr med; Heike Leutheuser[9], Prof Dr-Ing; Bjoern M Eskofier[1,10*], Prof Dr; Alexander Mocker[5,6,7,8*], Dr med

[1]Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nürnberger Straße 74, Erlangen, Germany

[10]Translational Digital Health Group, Institute of AI for Health, Helmholtz Zentrum München, Neuherberg, Germany

[2]Department of Ethics and Law in Medicine, University of Vienna and Medical University of Vienna, Austria

[3]Department of Systematic Theology and the Study of Religions, University of Vienna, Austria

[4]Department of Theology, Chair of Systematic Theology (Ethics), Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

[5]Department of Gynecology and Obstetrics, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

[6]Comprehensive Cancer Center Erlangen (CCC ER-EMN), Erlangen, Germany

[7]Comprehensive Cancer Center Alliance (CCC WERA), Erlangen, Germany

[8]Bavarian Cancer Research Center (BZKF), Erlangen, Germany

[9]Department of Computer Science, Chair of Machine Learning in Medicine, University of Bayreuth, Bayreuth, Germany

[*]these authors contributed equally

Corresponding Author:
Madeleine Flaucher, MSc
Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nürnberger Straße 74, Erlangen, Germany

Abstract

Background: Mobile health (mHealth) apps promise to enhance patient empowerment, enable real-time health monitoring, and support self-management. Patients with breast cancer stand to benefit particularly from these capabilities, given the demanding pre- and posttreatment procedures they face. However, the effective adoption of these tools is challenged by issues including accessibility, usability limitations, and privacy concerns.

Objective: This study aimed to investigate the current usage behavior and attitudes toward mHealth apps among patients with breast cancer in Germany, focusing on acceptance, usefulness, empowerment, and data-sharing. Specifically, we aimed to identify the incentives and barriers influencing use and the intention to use mHealth apps at the time of this writing to guide their further implementation and use.

Methods: A cross-sectional online survey was conducted between November 2023 and June 2024 using a structured questionnaire developed in German. Participants with a self-reported breast cancer diagnosis were recruited via university hospital channels, patient associations, and social media. Topics included technology and health app usage, perceived usefulness, personal empowerment, and attitudes toward data sharing. The survey comprised both closed- and open-ended questions to capture quantitative patterns and qualitative insights into participants' experiences and perspectives.

Results: We collected responses from 90 participants with breast cancer. Participants demonstrated high engagement with digital technology; 97% (87/90) reported daily smartphone use, and 70% (63/90) owned a smartwatch or fitness tracker. Approximately 67% (60/90) of respondents had used health apps in the past year, with many relying on general fitness and lifestyle apps to support their health management. Most participants (62/90, 69%) indicated that health apps contribute to a better understanding of their personal health. Health apps were generally perceived as useful, with users highlighting their ability to provide continuous support and timely, reliable health information (n=31). This, in turn, was seen as instrumental in promoting self-management and empowerment. Respondents also showed a conditional willingness to share personal health data for research, particularly when clear benefits for treatment advancements were evident (n=82). Nonetheless, concerns were raised regarding data privacy, app usability, the need for multiple apps to cover all necessary features, and the potential increase in fear of illness (cyberchondria).

XSL•FO
RenderX

Conclusions: The findings underscore the potential of mHealth apps to empower patients with breast cancer by improving health literacy and supporting self-management. To maximize their impact, future digital health tools must align closely with patient needs by incorporating adaptive, interactive features and integrating seamlessly with clinical care. Addressing key barriers—particularly data privacy and usability issues—is essential for broader adoption. Longitudinal studies are warranted to evaluate the long-term effects of health app usage on patient empowerment, quality of life, and clinical outcomes.

(JMIR Cancer 2026;12:e77750) doi:10.2196/77750

KEYWORDS

mHealth applications; women's health; smartphone; digital health; data sharing

Introduction

The intersection of technology and health care has introduced new pathways in mobile health (mHealth) apps. These digital tools promise enhanced patient autonomy, real-time health monitoring, and personalized care, capabilities particularly appealing for managing chronic diseases as well as long-lasting, complex, and potentially recurring conditions [1]. For patients with breast cancer, mHealth apps could represent a vital resource during their health treatment process [2,3].

Breast cancer is one of the most commonly diagnosed cancers among females worldwide, with approximately 2.3 million new cases annually [4-6]. The diverse nature of the disease requires targeted and individualized treatment [7]. Beyond the physical health implications, breast cancer impacts the emotional and psychological well-being of patients, often leading to anxiety, depression, and a reduced quality of life (QoL) [8]. These effects frequently extend to their families and support networks, who may experience emotional distress and caregiver burden [9]. Hence, effective management and care for breast cancer demand a holistic approach that incorporates medical, psychological, and social support to mitigate these effects [10,11].

By helping navigate the complexity of treatment, emotional distress, and long-term care, mHealth apps have great potential. The technology has the potential to help patients better understand their personal health status, contribute to self-management strategies, and thus improve health outcomes and QoL. Yet the promise of technological innovation often collides with realities of patient experience, where acceptance and adoption are shaped by more than functionality. On the one hand, existing solutions in women's health available in app stores often lack trustworthy resources, a user-centered design with high usability, and research-supported effectiveness [12,13]. Some might even represent risks for patients through outdated health information. Apps developed within a thorough research process and evaluation, on the other hand, usually represent island solutions that were never made available to the general public [14].

Furthermore, existing literature primarily addresses technical and usability aspects, often overlooking the multifaceted needs of patients. While ensuring technical functionality and user-friendly design is critical, these aspects alone do not guarantee the effective integration of mHealth tools into the daily lives of patients. The development requires a comprehensive, patient-centered research approach that prioritizes emotional, cultural, and ethical considerations alongside technological development.

The primary aim of this survey study is to examine this usage behavior, acceptance, perceived usefulness, empowerment, and data-sharing attitudes related to mHealth apps among patients with breast cancer. Secondary objectives were to identify strategies to enhance adoption, engagement, and effective data sharing for future mHealth apps.

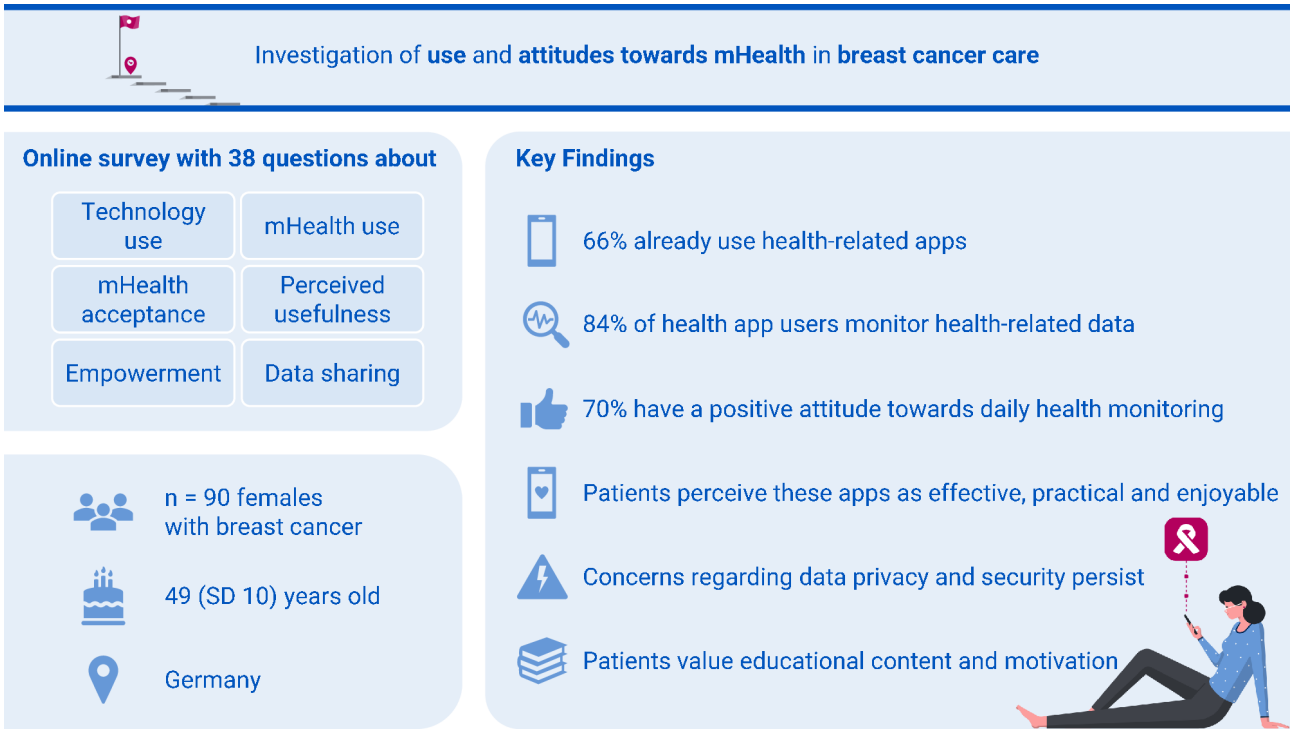
Methods

Survey Design

A questionnaire was designed to investigate several aspects related to the use of health apps among individuals with breast cancer. This study is embedded in the DigiOnKo project [15], an interdisciplinary initiative aimed at advancing digitally supported oncology care pathways by developing, evaluating, and implementing digital tools to support patients and clinical teams. An overview of the study is provided in Figure 1. Based on our research goal, we defined 2 exclusion criteria, individuals who had never been diagnosed with breast cancer in their lives or were under the age of 18 years. The questionnaire was designed in German. Besides basic demographic information, questions covered technology access, health app use, perceived usefulness of health apps, personal empowerment, health app acceptance, and data sharing. Further details about the survey instrument will be presented in the following chapters.

XSL•FO
RenderX

Figure 1. Visual abstract summarizing the objectives, methodology, and key findings of our survey study on health app usage among patients with breast cancer.



The survey was implemented in Unipark (Tivian XI GmbH). The questions included both open and closed questions; answer options were mainly on a 5-point Likert scale for closed questions. All items of the questionnaire were mandatory. All questions asking sensitive information provided a nonresponse option of "Prefer not to answer," "not applicable," or "neutral." Participants were able to review all answers before finishing the survey. The question order was not randomized; 9 questions were conditional, depending on answers to previous questions.

The minimum number of items was 29 and the maximum number was 38. These questions were distributed over 21 survey pages in total. The full questionnaire can be found in Multimedia Appendix 1. Before submitting the questionnaire, an optional feedback field was provided for any further comments or remarks. The survey included 6 open questions, where participants were instructed to provide short written answers in bullet points. Additionally, a text field for further comments and feedback was provided.

Technology and Health App Use

To gain an overview of the usage of technology in general and related to their personal health, all participants were asked about the use and usage frequency of smartphones, tablets, and wearables. The aim was to gain insights into the engagement and adoption of technology in general and mHealth in particular. This was complemented by questions regarding internet usage frequency and the confidence in using smartphones and mobile apps.

Related to personal health, participants were asked about recently used resources to improve their personal health (podcasts and audiobooks, social media, video calls, smartphone apps, printed materials, and internet). Specifically, we asked which health apps they have installed on their devices, how often they use them, what they use them for, and whether they felt informed about the risks and benefits associated with these apps at the time of this writing. The question was targeted specifically at all health-related apps and was not limited to specific breast cancer apps, as patients might use a variety of different apps to manage their condition (eg, general fitness apps or meditation apps).

Perceived Usefulness

According to the Technology Acceptance Model (TAM), the perceived usefulness of a technology is closely linked to their attitude toward it as well as their intention to use it [16]. If patients perceive mHealth apps as useful, they are more likely to engage with them regularly. As shown in a meta-analysis by Gan et al [17], greater engagement in mHealth apps was significantly associated with mental health improvements. Perceived usefulness was measured through a questionnaire developed by Voss et al [18] that has been used in previous work and captures the hedonic and utilitarian dimensions of attitudes toward a product category. With 10 different adjective pairs, it captures the perceived usefulness on a 5-point Likert scale, for example, effective-not effective and boring-exciting. Additionally, participants were asked about the perceived benefits and risks in 2 open questions.

Personal Empowerment

Patient empowerment is fundamental for effective chronic disease management, as it is closely linked to improved self-management, better adherence to treatment, and enhanced overall well-being. Empowered patients are more likely to engage in informed decision-making, feel confident in managing their condition, and experience greater autonomy in their care journey [19,20]. In this study, all measurement items were adapted from Lemire et al [21] and included items related to

XSL•FO
RenderX

the ability to apply gained knowledge to effectively make decisions related to personal health [22]. To further explore this dimension, participants were also asked whether they experience discrimination in their daily lives and whether they fear additional disadvantages or discrimination related to using mHealth apps.

Acceptance of Health Apps

Daily interaction with an mHealth app can induce psychological harm from health anxiety or cyberchondria, the excessive or compulsive seeking of online health-related information that exacerbates anxiety [23,24]. For patients with breast cancer, this phenomenon can be particularly pronounced, as apps often encourage users to continuously monitor their health, presenting daily reminders of their condition.

To investigate the acceptance toward health apps in breast cancer care, we used questions from Potdar et al [25] that have been used to investigate acceptability toward mHealth among patients with cancer in general. The questions targeted features for monitoring the health status, tracking information on the smartphone, and using a health app to improve their health-related knowledge, as these represent features often used in mHealth apps for breast cancer [26]. With an open question, we then asked participants about specific features they would expect or want in an mHealth app designed to support patients with breast cancer.

Data Sharing

According to the Privacy Calculus Theory, individuals weigh perceived benefits, such as improved therapies or valuable research contributions, against potential risks, like loss of privacy or misuse of data [27]. By targeting data-sharing scenarios under specific conditions—such as offering incentives, ensuring transparency in data use, or highlighting therapeutic advancements—we aimed to identify the factors that encourage or discourage data sharing. For this purpose, we used an adapted version of the questionnaire developed by Pletscher et al [28]. Questions targeted specifically at sharing data with research institutions under certain conditions, for example, if incentives are offered, therapies can be improved, or the purpose for the data use is known. Additionally, reasons why participants would want to share or not share their data were asked in an open question.

Survey Pretesting

The final questionnaire went through a test run with 4 independent assessors to identify any usability issues, verify that questions are clear and understandable, and reduce any possible biases in questions.

Recruitment Process

The survey was designed as an open, online-only survey; everyone with access to the link was able to participate on a voluntary basis. On the landing page of the survey, website visitors were informed about the study, the objectives, the estimated duration, the types of data collected within the survey, storage of the data, and the contact information of the investigators in charge. In order to start the survey, active consent for participation was required. After consent, 2 eligibility questions were asked; all participants had to be at least 18 years old and had to have a breast cancer diagnosis at some point in their lives, as the target population was patients with breast cancer.

A convenience sampling strategy was chosen. The survey was distributed through a combination of clinic-based recruitment and online channels. The survey was promoted by posters and flyers in the university hospitals in Erlangen and Würzburg (Bavaria, Germany); via personal contacts; as well as various mailing lists and patient associations. Online recruitment included the university website and social networks such as Instagram (Meta Platforms), Facebook (Meta Platforms), LinkedIn (Microsoft Corp), and nebenan.de (Good Hood GmbH). On Facebook, the survey was posted in patient groups related to breast cancer. On Instagram, various influencers were asked to distribute the flyer to their followers. Among them, PINK gegen Brustkrebs GmbH (Germany), the company behind the breast cancer app Pink! Coach promoted the survey through an advertisement. Social media recruitment was included to increase reach and accessibility, allowing participation from patients who might not regularly attend the recruiting clinics or who live outside the immediate hospital regions. The data collection was performed between November 2023 and June 2024.

Data Analysis

Quantitative data were analyzed using descriptive statistics (frequencies, percentages, means, and SDs). All analyses were conducted using Python (Python Software Foundation). All answers to open questions were analyzed with an inductive category formation of a qualitative content analysis according to Mayring and Fenzl [29]. The analysis was conducted following Mayring's systematic procedure of iterative category development; initial codes were developed and refined through repeated review. Representative quotes were selected to illustrate key themes. Survey results are reported in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) [30]; the full checklist is provided in Checklist 1.

Ethical Considerations

All data were collected anonymously. The dataset did not contain any personal data in accordance with the GDPR. Thus, this study was exempt from institutional review. The exemption was confirmed by the Institutional Review Board of FAU Erlangen-Nürnberg (23 - 333-ANF). Informed consent was obtained from all participants prior to participation in the survey; participants did not receive any compensation.

Results

Survey Statistics

In total, the landing page of the survey was visited 404 times, and 154 individuals agreed to participate in the survey and filled out the eligibility screening. 145 reportedly met the defined criteria and started the questionnaire, and 90 completed it. Therefore, the completion rate was 62%. Only completed surveys are included in the subsequent analysis. On average, participants spent 15.9 (SD 7.8) minutes completing the survey.

XSL•FO
RenderX

No atypical timestamps were detected; therefore, no participant was excluded due to the duration of the participation.

Participant Demographics

The mean age of participants is 49 (SD 10) years, with 88 out of 90 (98%) indicating female sex; 2 (2%) participants out of 90 preferred not to specify. Detailed sociodemographic information about the participants is shown in Table 1.

Table . Demographic information of all survey participants.

Demographics	Participants (N=90)
Age (years)	
Range	26-71
Mean (SD)	49 (10)
Age at breast cancer diagnosis (years)	
Range	25-69
Mean (SD)	49 (9)
Gender, n (%)	
Woman	88 (98)
Unspecified	2 (2)
Education, n (%)	
High school diploma	33 (37)
Vocational education	26 (29)
University degree	27 (30)
Doctoral degree	1 (1)
Unspecified, n (%)	3 (3)
Employment, n (%)	
Full-time	30 (33)
Part-time	33 (37)
Self-employed	1 (1)
Retired	13 (14)
Nonworking	7 (8)
Unspecified	6 (7)
Health insurance, n (%)	
Private	14 (16)
Statutory	75 (83)
Unspecified	1 (1)
Cancer therapy status, n (%)	
In therapy	48 (53)
Aftercare	40 (44)
Unspecified	2 (2)

Technology Usage

All participants possess a smartphone, and most use it daily (87/90, 97%). A majority (63/90, 70%) have a smartwatch or fitness tracker and use it daily or almost daily (52/90, 83%). Additionally, participants rated how confident they feel using smartphones and apps. A majority feels very confident (46/90, 51%) or rather confident (26/90, 29%); a few participants feel very insecure (4/90, 4%) or rather insecure (2/90, 2%) in smartphone and app use. The majority of participants (86/90, 86%) reported a daily use of the internet; all participants accessed the internet at least once a week.

mHealth Usage

Among all participants, 59/90 (66%) stated that they had used mobile apps to improve their personal health status within the last 12 months. The internet in general was used by 82 (91%),

XSL•FO
RenderX

social media by 72 (80%), and audiobooks or podcasts by 51 (57%).

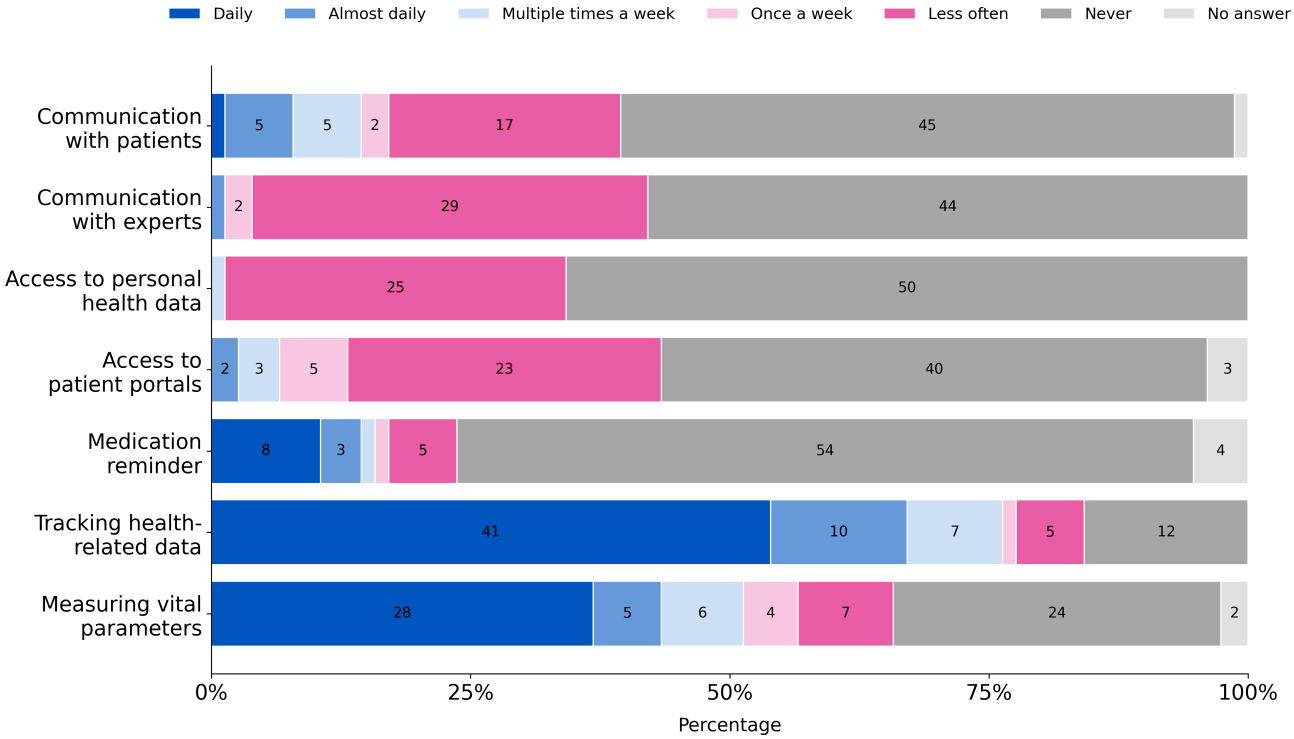
When asked about installed apps on their devices, 70 (78%) participants listed the apps. On average, participants had 2 apps related to their health installed; one participant listed 15 health-related apps.

The most frequently mentioned app was “Pink! Coach” (25/90, 28%), a medically certified app in Germany specifically designed to improve QoL and health literacy in patients with breast cancer [31]. Besides that, general fitness apps such as Fitbit (Fitbit Inc), Garmin Connect (Garmin Ltd), or Apple Health (Apple Inc) are frequently installed.

Regarding the frequency of using these apps, 59 out of 90 (66%) stated at least once a week, and 21 out of 90 (23%) mentioned daily use. Among all participants (n=90), 14 (16%) reported that they are not using health apps. A filter was applied so that only those who indicated app usage (n=76) received the subsequent questions regarding app usage purpose.

The purpose and frequency of using these apps for specific purposes are represented in Figure 2. Most participants use their apps to measure physiological parameters such as heart rate and to monitor other health-related data such as physical activity, steps, or diet. As shown in Figure 2, 67% (n=51) of app users are already monitoring health-related parameters daily or almost daily. In contrast, the use of an electronic health record, contact with health professionals or other patients, and access to patient portals were not common among the participants.

Figure 2. Answer distributions of the usage frequency to different features commonly included in mHealth apps by participants who use health apps (n=76).



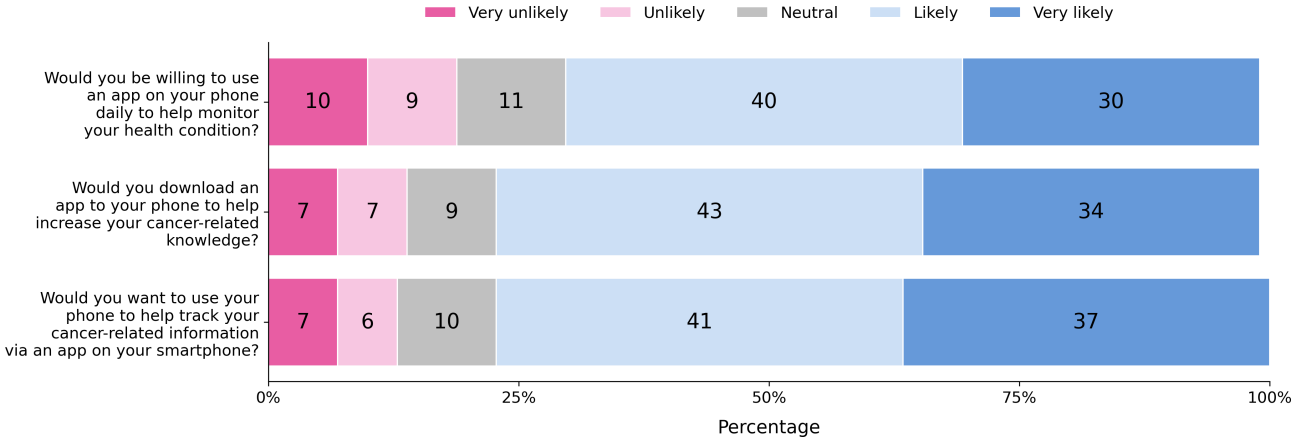
Among the app users, participants were asked whether they had been informed about the benefits and risks of using the health app before starting. A total of 30 (40%) out of 76 app users reported feeling uninformed about the risks associated with the app. In contrast, 36 out of 76 (47%) stated they had been adequately informed about their benefits.

Acceptance

Figure 3 shows the answers of all participants to the 3 questions related to the general acceptance of cancer-related mobile apps. Most participants are willing to use their smartphone for tracking cancer-related information through a mobile app and to use it as a health monitor on a daily basis. Furthermore, a majority of participants would be willing to download a mobile app to increase their knowledge about cancer.

XSL•FO
RenderX

Figure 3. Answers given to the acceptance-related survey items in percentages.

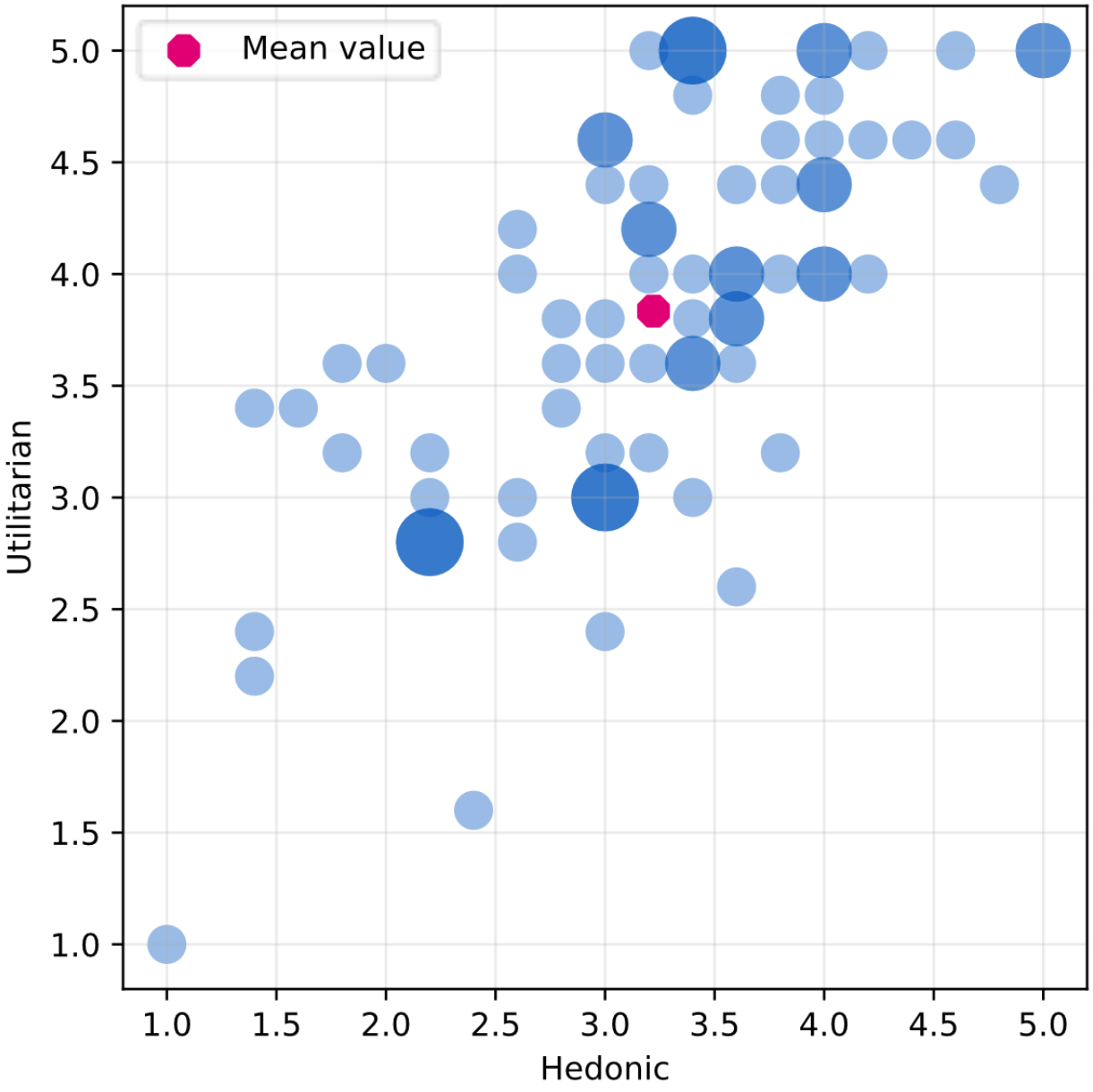


Perceived Usefulness

The utilitarian and hedonic dimensions of the attitude toward health apps of users are shown in Figure 4. Overall, the average on the hedonic dimension was 3.22 (SD 0.87) and on the utilitarian dimension 3.83 (SD 0.87). Thus, the used mHealth apps are perceived as both pleasurable and useful.

XSL•FO
RenderX

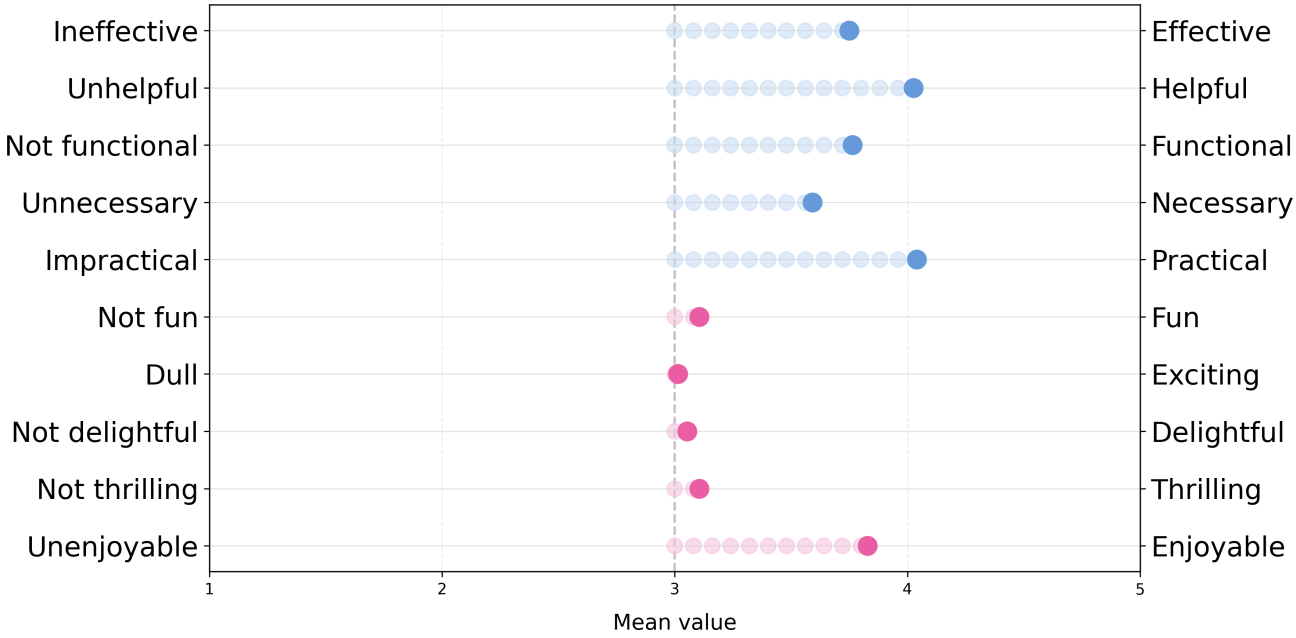
Figure 4. Scatterplot illustrating perceived usefulness scores for all participants (blue), where larger circles indicate multiple participants with the same response. The overall average is highlighted in pink.



In more detail, as shown in Figure 5, participants perceive their apps as effective, practical, and enjoyable.

XSL•FO
RenderX

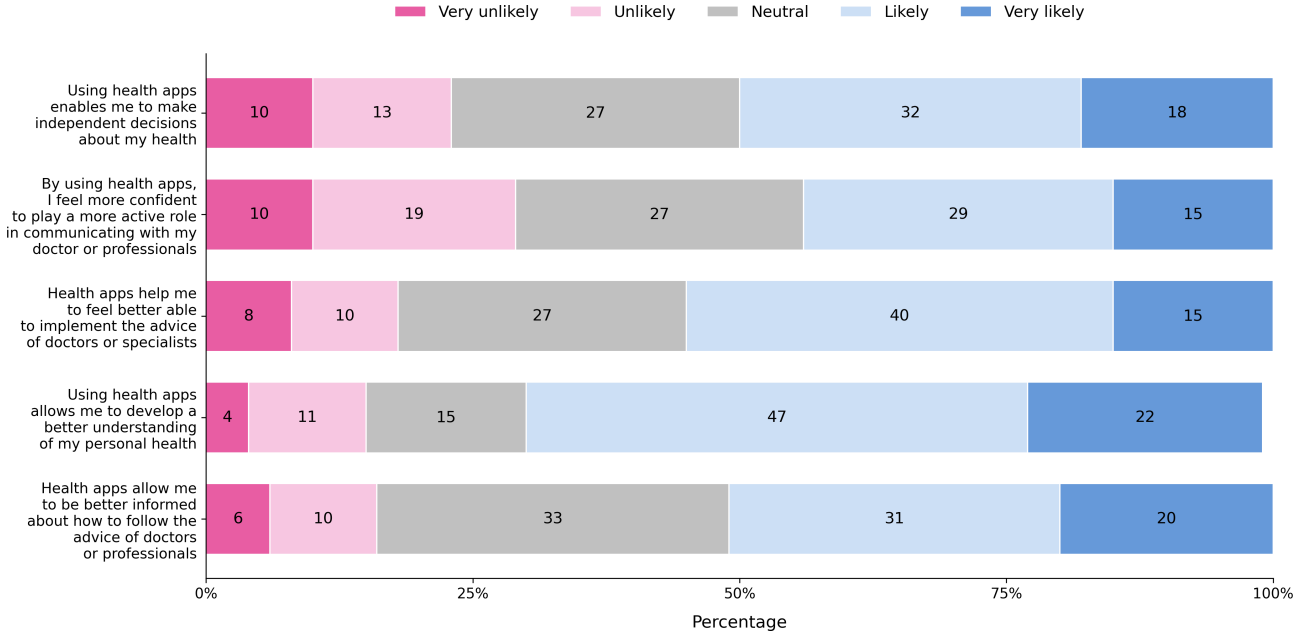
Figure 5. Mean values for all categories of the utilitarian and hedonic dimensions of the questionnaire for perceived usefulness.



Personal Empowerment

The findings, as illustrated in Figure 6, reveal varying levels of agreement among participants regarding the empowering potential of health apps. A majority of respondents indicated that health apps likely or very likely enable them to develop a better understanding of their personal health (62/90, 69%) and implement advice from doctors or specialists (49/90, 54%). Similarly, 39 out of 90 (43%) of participants felt that using health apps enhances their confidence in communicating with health care professionals, while 45 out of 90 (50%) agreed that these tools help them make independent decisions about their health. Overall, a minority of participants perceived health apps as very unlikely or unlikely to contribute to these empowerment domains. A total of 29% (26/90) think it is unlikely that using health apps could give them a more active role when talking with their doctors. With an overall mean for all items of 3.4 out of 5.0, the participants report a positive contribution of health apps to personal empowerment.

Figure 6. Participants' responses to 5 different aspects of personal empowerment: (1) making independent health decisions, (2) engaging more actively with health care professionals, (3) feeling more capable of following medical advice, (4) gaining a better understanding of personal health, and (5) staying informed on how to follow professional recommendations.



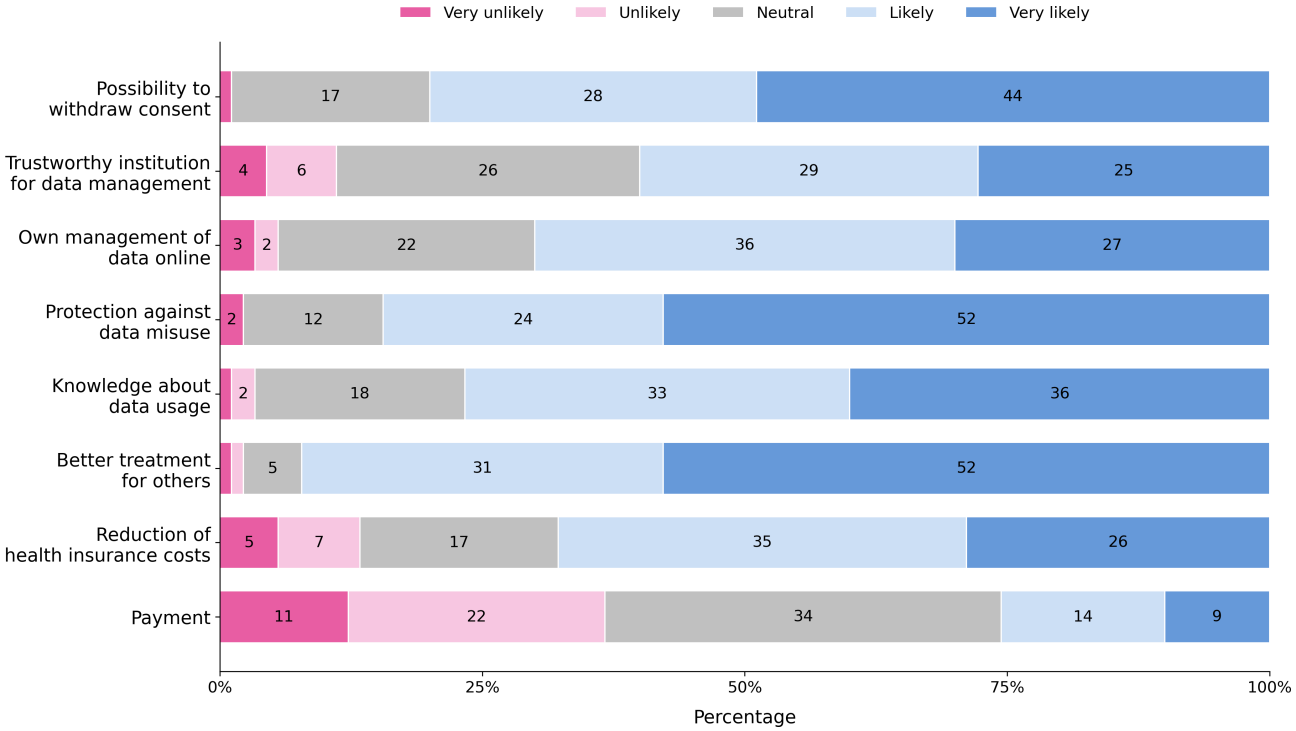
Data Sharing

The most relevant drivers of sharing personal health data for research purposes from the list provided to the participants, as shown in Figure 7, are (1) if people knew that by doing so, others would receive a better treatment (n=82 likely, or very likely); (2) if people were sure that data were protected from misuse (n=75 likely, or very likely); (3) if people could withdraw from the consent of data usage anytime (n=72, 80%); and (4) if it were known what research is done with the data (68/90,

XSL•FO
RenderX

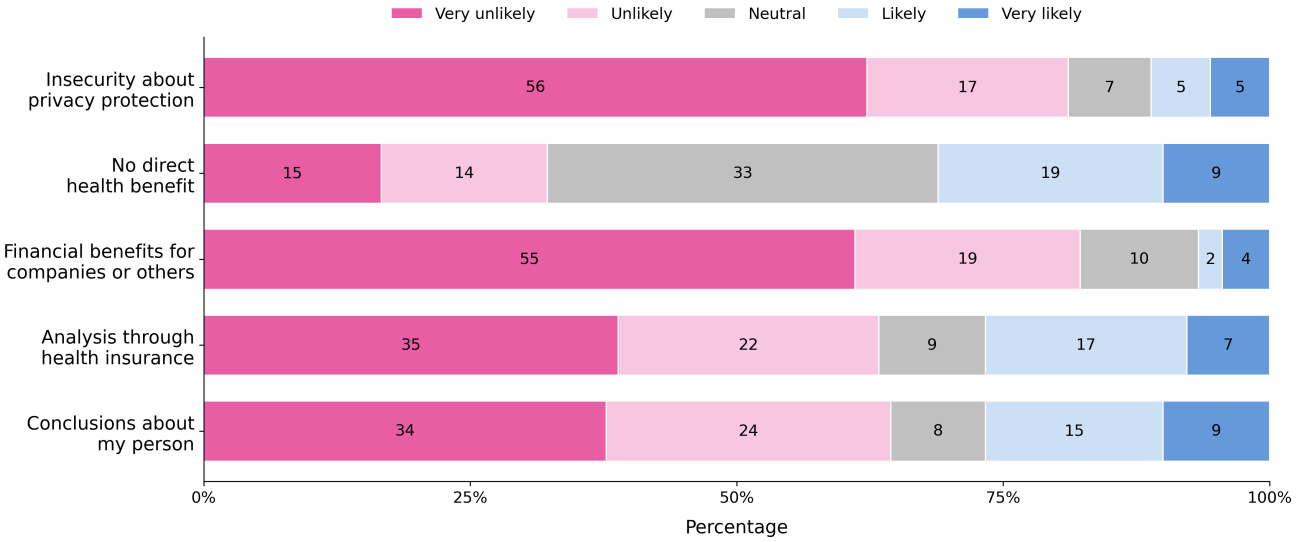
76%). On the other hand, money was perceived as less of an incentive (22/90, 24% likely or very likely).

Figure 7. Suggested incentives and the participants' attitudes toward sharing personal health data for research purposes.



Concerns about privacy protection and financial benefits for companies were the most significant barriers, with the majority of respondents indicating these were "Very likely" to hinder their data sharing. A lack of direct health benefit is perceived as less of a barrier. A detailed overview of the responses to potential barriers is shown in Figure 8.

Figure 8. Attitudes of participants toward 5 barriers to sharing personal health data.



Qualitative Data Analysis

Required App Features and Content

Participants were asked to list specific features they expect in an mHealth app supporting patients with breast cancer. Related to general usability and design features, 6 main categories were identified. Mostly mentioned were notifications and reminders (n=17), followed by clearness and simplicity (n=9), motivational features (n=7), adaptation to individual needs (n=6), and trustworthiness (n=4).

In terms of educational content, in addition to general cancer-related information (n=58), detailed advice and recommendations on diet (n=20), physical activity (n=19), and topics related to mental health and general well-being (n=17) were of particular interest to participants. In addition, more updates on recent research results and ongoing studies were frequently requested (n=16). Participants emphasized the importance of up-to-date information that evolves over time ("Up-to-dateness, continuous expansion of information" P79, translated), as well as concise explanations of medical terms ("Database [Explanation of terminology]" P69, translated).

XSL•FO
RenderX

Benefits and Risks of mHealth Apps

When asked about the benefits mHealth apps offer in the context of breast cancer, 31 specifically mentioned the provision of information, especially up-to-date information (n=2) and trustworthy information (n=2). Furthermore, participants see benefits in motivating users (n=20) to eat healthy or exercise. A large group of participants (n=22) highlighted the personal and mental support of these apps in dealing with disease, therapy, and side effects. Especially since they are available 24 hours a day (n=5). For example, one participant mentioned that these apps offer

> *personal support and helpful input, that can be selected individually, in addition to the often limited time available during consultation hours* [(P33, translated)]

Participants also mentioned that these apps can provide a sense of control and security (n=7). For example, they could help to

> *assess your personal lifestyle and identify opportunities for improvement (gives you the confidence to "do something" yourself)* [(P33, translated)]

In contrast, 23 participants did not see any benefits in health apps. As reasons, participants mentioned, for example, a lack of knowledge about available apps ("I would like more information about the available options," P36, translated) or that information is already available on the internet ("This knowledge can be obtained on the Internet on your own," P23). One participant stated that, in their experience, mHealth apps usually cover only a small aspect, lack usability, and are difficult to activate due to the need for a doctor's prescription (P90).

When asked about the risks of mHealth apps, the majority of participants (n=42) stated that they did not see any risks. Only 10 participants expressed concerns regarding data privacy and security. Furthermore, a few participants were concerned about potential misinformation (n=8) or negative mental health effects (n=6), for example, by providing wrong diagnoses or untrustworthy sources. Additionally, 7 participants mentioned a potential overload, for example, through providing too much information. As obstacles hindering participants in using available health apps, participants mentioned reasons such as insufficient quality, lack of usability, lack of time, or lack of cost coverage by health insurance.

Discussion

Overview

With this study, we investigated self-reported usage patterns, health app acceptance, and views on data sharing among patients with breast cancer in Germany. Our results provide insights into patients' perspectives on mHealth technology, including aspects they perceive as beneficial as well as challenges and barriers that still need to be overcome. Especially, aspects related to usability, effectiveness, and privacy concerns need further attention from the patient's viewpoint.

Participants reported a high prevalence of fitness tracker and smartwatch usage. This could have several reasons; on the one hand, this survey might have been especially interesting to technology-affine people. On the other hand, the popularity of mobile and wearable technology, particularly among older adults, is increasing [32]. Additionally, one of 6 patients with breast cancer in Germany is younger than 50 years old, which could also contribute to higher adoption of digital health tools in our sample [33,34].

In general, the majority of participants expressed openness to using mHealth apps for monitoring their personal health status and enhancing their health knowledge. Also, daily interaction with an app appears to be feasible and not a deterrent for most participants. This positive attitude aligns with the growing trend of integrating technology into health care practices, particularly for older adults [32].

Health Apps as a Tool for Patient Empowerment

The findings suggest that health apps are generally perceived as useful and have the potential to empower patients. Participants specifically highlighted the availability of these apps, offering support around the clock in addition to limited interaction time with health care providers. This availability was mentioned as a key factor in fostering a sense of control and confidence, both of which are central to patient empowerment [35]. Participants valued apps that could deliver timely and reliable health-related guidance, highlighting a broader shift toward digital health literacy as a central component of modern patient care.

When designed effectively, such apps can facilitate quick and reliable answers to users' questions, thereby fostering health literacy and eHealth literacy. This, in turn, has the potential to enhance self-management and improve QoL by supporting informed decision-making, healthy behaviors, and adherence to therapy. Notably, the extent to which these apps actually enhance self-management likely depends on factors such as usability, personalization, and integration with clinical care issues that remain underexplored in the digital health landscape at the time of this writing [36].

Effective empowerment requires not just access to information but also actionable, tailored guidance that aligns with individual health needs. To maximize their impact, future health apps should go beyond providing generic health content and instead offer dynamic, interactive features that support patient autonomy. This includes adaptive recommendations, such as tailored exercise or lifestyle guidance and nudges based on app usage patterns, but also the seamless integration with clinical care, for example, through secure messaging, automated report sharing, or appointment planning. Beyond these, user-driven customization options should be provided, such as adjustable dashboards, modular feature selection, and flexible notification settings. These features can help patients tailor the app experience to their individual needs, literacy levels, and desired depth of information and engagement. Further research is needed to identify which app functionalities most effectively contribute to different dimensions of empowerment and how these functionalities are perceived with different patient characteristics and care contexts.

XSL•FO
RenderX

Widespread Use of Health Apps and Their Limitations

Interest in health apps is high, with many participants using apps not specifically designed for breast cancer management. Fitness apps, in particular, were frequently mentioned, as they support a healthy lifestyle, a key factor for many participants in improving QoL. This suggests that patients are already adopting existing digital tools to support their health, even in the absence of condition-specific alternatives. Furthermore, these findings suggest that patients exhibit intrinsic motivation to enhance their health and actively contribute to their therapy. It highlights their desire to play an active role in their health care journey.

However, while general health apps offer broad functionality, they may not address the unique concerns of patients with breast cancer, such as symptom tracking, treatment side effects, or emotional well-being. This raises an important question. Do patients actually need breast cancer-specific apps, or are they sufficiently served by more general health applications? The widespread use of existing tools suggests that some needs may already be met; however, it remains unclear whether these apps are fully adequate for the challenges of managing a chronic illness. Future research should investigate whether condition-specific features—such as integration with oncology care, personalized symptom management, or peer support—would enhance engagement and health outcomes. Understanding these preferences could inform the development of more targeted digital health solutions that complement, rather than duplicate, existing app functionalities.

Balancing Data Privacy Concerns and Willingness to Share

Data privacy remains a central concern for many participants, with uncertainty surrounding how health apps handle sensitive personal information. This skepticism reflects broader public debates on digital health security and trust in technology providers [37]. However, our findings reveal that privacy concerns do not translate into outright rejection of data sharing. Instead, participants demonstrated a conditional willingness to share their health data, particularly when contributing to research with tangible benefits for future treatment advancements. This highlights a critical balance between risk perception and perceived utility, suggesting that concerns about data security are not absolute barriers but negotiable factors. These findings align with previous research indicating that patients with cancer are particularly open to data sharing when they perceive benefits for themselves or others from the research conducted [38].

Transparent communication regarding data usage, robust security measures, and opportunities for users to maintain control over their data are essential for improving trust. Beyond technical safeguards, actively involving patients in research processes—through co-design initiatives, participatory data governance models, or better communication of research findings—could further shift perceptions, positioning data sharing as an informed and mutually beneficial decision rather than a passive risk.

Structural Barriers

Despite their potential, these health apps often fail to fully meet users' needs. A key challenge was the fragmentation of app functionalities, with users often having to rely on multiple apps to cover different aspects of their health management. This lack of integration creates inefficiencies and may discourage long-term use. Additionally, structural issues—such as unclear reimbursement policies and prescription processes—further complicate the adoption of digital health solutions within formal health care systems. Additionally, our study did not account for individuals who have no structural or material access to mHealth at all, whether due to a lack of technical infrastructure (ie, no smartphone) or knowledge barriers (eg, illiteracy, lack of digital literacy).

Psychological Impact of Health Tracking

Another notable concern was the psychological impact of continuous health tracking. Some participants questioned whether constant app engagement might lead to an excessive preoccupation with their illness rather than providing reassurance. This highlights the need for a balanced approach in app design. Digital tools should support patient autonomy without reinforcing anxiety or over-monitoring. While tracking tools can facilitate behavior change, their effectiveness depends on sustained engagement [39,40]. If apps demand too much attention or become emotionally taxing, users may abandon them over time. Striking the right balance between engagement and cognitive load is crucial for long-term use. Practical design strategies could include adjustable frequencies of self-monitoring, optional data summaries, or modular features that can be enabled or disabled depending on individual needs and preferences.

Toward Integrated and Patient-Centered Solutions

Addressing these challenges requires a more holistic approach to app development. This approach to app development, characterized by fragmented and often redundant solutions, may not be sustainable [41]. Developers should shift their focus toward creating integrated, multifunctional platforms that address a broad spectrum of patient needs, as also identified in other related works [42,43]. Additionally, engaging stakeholders—including health care providers and patients—in the design process can help ensure that apps are practical, trustworthy, and aligned with real-world usage contexts [44]. Beyond design considerations, clear regulatory frameworks, standardized evaluation criteria, and improved access to reimbursement will be crucial in embedding digital health tools into routine care.

Limitations

Overall, this work provides insights into the actual usage and perspectives toward mHealth apps among patients with breast cancer in Germany. Nevertheless, this study is subject to several limitations. The sample size was relatively small, which may limit the generalizability of the findings. The study was conducted as an online survey and was mainly distributed online. Thus, only participants with internet access and at least a computer or smartphone were able to participate. In combination with the general topic of the survey, it is likely that people with

XSL•FO
RenderX

an affinity for technology in particular were addressed and that the group of patients without internet access is not represented in the collective. Due to the survey tool, which did not include cookies, internet protocol address checks, or other measures, repeated participation cannot be ruled out completely.

Since this study investigated self-reported data, actual usage patterns may differ from what participants stated in this questionnaire.

Furthermore, the survey was available only in German. As other countries have different possibilities of mHealth in breast cancer and are at different levels of digitizing health care, acceptance and usage might differ largely.

Furthermore, the participants were, on average, younger than the typical age at breast cancer diagnosis and had a higher level of education, which may limit the generalizability of the findings to the broader patient population. Finally, the small sample size and the fact that we were not able to recruit any non-female participants, even though specific organizations were contacted for that, highlight the need for interpreting the results carefully.

Conclusions

This study provides valuable insights into current usage behaviors and attitudes toward health apps from the perspective of patients with breast cancer. We highlight key aspects that could enhance adoption, engagement, empowerment, and effective data sharing for future mHealth applications and research. Despite the lack of condition-specific apps, patients are already using general health and fitness apps to improve their QoL, indicating a strong intrinsic motivation for self-management in our study population. There is an opportunity to build on existing app usage behaviors by tailoring future digital health tools to meet the unique needs of patients with cancer. However, significant barriers, particularly privacy concerns, must be addressed to enhance trust and encourage broader adoption.

The findings emphasize the importance of aligning health apps with patient needs and preferences to maximize their utility and adoption. For health care, integrating apps into routine care pathways can provide patients with supplementary tools for self-management, potentially reducing the burden on health care providers. Collaboration among developers, health care professionals, and patients will be critical to ensure that future apps effectively address real-world challenges.

Further research is needed to assess the long-term impact of health apps on patient empowerment, including sustained engagement, decision-making autonomy, and self-efficacy. Additionally, studies should examine their effects on QoL, including mental well-being, treatment-related distress, and daily functioning, as well as clinical outcomes such as symptom management, adherence to therapy, and disease progression. Finally, the potential of artificial intelligence, machine learning, and wearable sensors should be explored to address the requirements of patients, such as personalization or real-time health insights.

Acknowledgments

We sincerely thank all participants for their time and willingness to complete the survey, which made this study possible. Generative AI (ChatGPT, GPT-4, OpenAI) was used to assist with language editing and stylistic refinement of the manuscript; the authors retained full control over the content and are responsible for all aspects of the text.

Funding

This project is supported by the Bavarian State Ministry of Health and Care, project grant number PBN-MGP-2010-0004-DigiOnko and by the Federal Ministry for Economic Affairs and Climate Action under grant number 68GX21004F (TEAM-X). Furthermore, this work was partly funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1483 – Project-ID 442419336, EmpkinS.

Data Availability

The datasets generated and analyzed during this study are not publicly available, as participants did not provide consent for public data sharing. However, they are available from the corresponding author upon reasonable request.

Authors' Contributions

MF conceived the study, developed the study design, analyzed data, recruited participants, and wrote the manuscript draft. AM and TO assisted in developing the study design, distributed the survey, and revised the manuscript. MN and HL assisted in designing the questionnaire and revised the manuscript. PAF, HH, MWB, BME, and HL initiated and supervised the project. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full version of the questionnaire in German.
[PDF File, 146 KB - cancer_v12i1e77750_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
English translation of the questionnaire.
[PDF File, 71 KB - cancer_v12i1e77750_app2.pdf]

Checklist 1
Full Checklist for Reporting Results of Internet E-Surveys (CHERRIES).
[PDF File, 172 KB - cancer_v12i1e77750_app3.pdf]

References

1. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digit Med 2020;3(1):4. [doi: 10.1038/s41746-019-0206-x] [Medline: 31970289]
2. Luo X, Chen Y, Chen J, et al. Effectiveness of mobile health-based self-management interventions in breast cancer patients: a meta-analysis. Support Care Cancer 2022 Mar;30(3):2853-2876. [doi: 10.1007/s00520-021-06568-0] [Medline: 34561732]
3. Würstlein R, Drewes C, Harbeck N, Schinköthe T. Onkologie. eHealth – Wunsch oder Wirklichkeit in der Gynäkoonkologie? Geburtshilfe Frauenheilkd 2014 Jun 6;74(5):415-416 [FREE Full text] [doi: 10.1055/s-0033-1350981]
4. Huang J, Chan PS, Lok V, et al. Global incidence and mortality of breast cancer: a trend analysis. Aging (Milano) 2021 Feb 28;13(4):5748-5803 [FREE Full text] [doi: 10.18632/aging.202502]
5. Kim J, Harper A, McCormack V, et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries. Nat Med 2025 Apr;31(4):1154-1162. [doi: 10.1038/s41591-025-03502-3] [Medline: 39994475]
6. Hübner J, Katalinic A, Waldmann A, Kraywinkel K. Long-term incidence and mortality trends for breast cancer in Germany. Geburtshilfe Frauenheilkd 2020 Jun;80(6):611-618. [doi: 10.1055/a-1160-5569] [Medline: 32565551]
7. Debela DT, Muzazu SG, Heraro KD, et al. New approaches and procedures for cancer treatment: current perspectives. SAGE Open Med 2021;9:20503121211034366. [doi: 10.1177/20503121211034366] [Medline: 34408877]
8. Culbertson MG, Bennett K, Kelly CM, Sharp L, Cahir C. The psychosocial determinants of quality of life in breast cancer survivors: a scoping review. BMC Cancer 2020 Oct 2;20(1):948. [doi: 10.1186/s12885-020-07389-w] [Medline: 33008323]
9. Borstelmann NA, Gray TF, Gelber S, et al. Psychosocial issues and quality of life of parenting partners of young women with breast cancer. Support Care Cancer 2022 May;30(5):4265-4274. [doi: 10.1007/s00520-022-06852-7] [Medline: 35091846]
10. Aprilianto E, Lumadi SA, Handian FI. Family social support and the self-esteem of breast cancer patients undergoing neoadjuvant chemotherapy. J Public Health Res 2021 Apr 14;10(2):2234. [doi: 10.4081/jphr.2021.2234] [Medline: 33855418]
11. Galmarini CM. Lessons from Hippocrates: time to change the cancer paradigm. Int J Chronic Dis 2020;2020:4715426. [doi: 10.1155/2020/4715426] [Medline: 32566644]
12. Altmannshofer S, Flaucher M, Beierlein M, et al. A content-based review of mobile health applications for breast cancer prevention and education: characteristics, quality and functionality analysis. Digit Health 2024;10:20552076241234627. [doi: 10.1177/20552076241234627] [Medline: 38528967]
13. Nissen M, Huang SY, Jäger KM, et al. Smartphone pregnancy apps: systematic analysis of features, scientific guidance, commercialization, and user perception. BMC Pregnancy Childbirth 2024 Nov 25;24(1):782. [doi: 10.1186/s12884-024-06959-1] [Medline: 39587534]
14. Flaucher M, Zakreuskaya A, Nissen M, et al. Evaluating the effectiveness of mobile health in breast cancer care: a systematic review. Oncologist 2023 Oct 3;28(10):e847-e858. [doi: 10.1093/oncolo/oyad217] [Medline: 37536278]
15. Mit digitaler medizin gegen brustkrebs. digiOnko. 2021. URL: https://www.digionko-bayern.de/ [accessed 2025-12-08]
16. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
17. Gan DZQ, McGillivray L, Han J, Christensen H, Torok M. Effect of engagement with digital interventions on mental health outcomes: a systematic review and meta-analysis. Front Digit Health 2021;3:764079. [doi: 10.3389/fdgth.2021.764079] [Medline: 34806079]
18. Voss KE, Spangenberg ER, Grohmann B. Measuring the hedonic and utilitarian dimensions of consumer attitude. J Mark Res 2003 Aug;40(3):310-320. [doi: 10.1509/jmkr.40.3.310.19238]
19. Denysyuk HV, Pires IM, Garcia NM. A roadmap for empowering cardiovascular disease patients: a 5P-medicine approach and technological integration. PeerJ 2024;12:e17895. [doi: 10.7717/peerj.17895] [Medline: 39224824]
20. Madanian S, Nakarada-Kordic I, Reay S, Chetty T. Patients' perspectives on digital health tools. PEC Innov 2023 Dec;2:100171. [doi: 10.1016/j.pecinn.2023.100171] [Medline: 37384154]
21. Lemire M, Sicotte C, Paré G. Internet use and the logics of personal empowerment in health. Health Policy 2008 Oct;88(1):130-140. [doi: 10.1016/j.healthpol.2008.03.006] [Medline: 18436330]
22. Salgado T, Tavares J, Oliveira T. Drivers of mobile health acceptance and use from the patient perspective: survey study and quantitative model development. JMIR Mhealth Uhealth 2020 Jul 9;8(7):e17588. [doi: 10.2196/17588] [Medline: 32673249]

XSL•FO
RenderX

23. Morley J, Floridi L, Goldacre B. The poor performance of apps assessing skin cancer risk. BMJ 2020 Feb 10;368:m428. [doi: 10.1136/bmj.m428] [Medline: 32048610]
24. Nissen M, Perez CA, Jaeger KM, et al. Usability and perception of a wearable-integrated digital maternity record app in Germany: user study. JMIR Pediatr Parent 2023 Dec 15;6:e50765. [doi: 10.2196/50765] [Medline: 38109377]
25. Potdar R, Thomas A, DiMeglio M, et al. Access to internet, smartphone usage, and acceptability of mobile health technology among cancer patients. Support Care Cancer 2020 Nov;28(11):5455-5461. [doi: 10.1007/s00520-020-05393-1] [Medline: 32166381]
26. Jongerius C, Russo S, Mazzocco K, Pravettoni G. Research-tested mobile apps for breast cancer care: systematic review. JMIR Mhealth Uhealth 2019 Feb 11;7(2):e10930. [doi: 10.2196/10930] [Medline: 30741644]
27. Majumdar A, Bose I. Privacy calculus theory and its applicability for emerging technologies. In: Sugumaran V, Yoon V, Shaw MJ, editors. E-Life Web-Enabled Converg Commer Work Soc Life: Springer International Publishing; 2016:191-195. [doi: 10.1007/978-3-319-45408-5_20]
28. Pletscher F, Mändli Lerch K, Glinz D. Willingness to share anonymised routinely collected clinical health data in Switzerland: a cross-sectional survey. Swiss Med Wkly 2022 Jun 20;152(2324):w30182. [doi: 10.4414/smw.2022.w30182] [Medline: 35752970]
29. Mayring P, Fenzl T. Qualitative inhaltsanalyse. In: Baur N, Blasius J, editors. Handb Methoden Empirischen Sozialforschung: Springer Fachmedien Wiesbaden; 2019:633-648. [doi: 10.1007/978-3-658-21308-4_42]
30. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
31. PINK! coach. PINK Brustkrebs. URL: https://pink-brustkrebs.de/aktiv-gegen-brustkrebs/pink-coach/ [accessed 2025-01-02]
32. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023 Jul;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8] [Medline: 37464029]
33. Hendrick RE, Monticciolo DL, Biggs KW, Malak SF. Age distributions of breast cancer diagnosis and mortality by race and ethnicity in US women. Cancer 2021 Dec 1;127(23):4384-4392. [doi: 10.1002/cncr.33846] [Medline: 34427920]
34. Krebs in Deutschland Für 2019/2020: Robert Koch-Institut; 2023. [doi: 10.25646/11357]
35. Hickmann E, Richter P, Schlieter H. All together now - patient engagement, patient empowerment, and associated terms in personal healthcare. BMC Health Serv Res 2022 Sep 2;22(1):1116. [doi: 10.1186/s12913-022-08501-5] [Medline: 36056354]
36. Abernethy A, Adams L, Barrett M, et al. The promise of digital health: then, now, and the future. NAM Perspect 2022;2022(22). [doi: 10.31478/202206e] [Medline: 36177208]
37. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
38. Köngeter A, Schickhardt C, Jungkunz M, Bergbold S, Mehlis K, Winkler EC. Patients' willingness to provide their clinical data for research purposes and acceptance of different consent models: findings from a representative survey of patients with cancer. J Med Internet Res 2022 Aug 25;24(8):e37665. [doi: 10.2196/37665] [Medline: 36006690]
39. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016 May 16;18(5):e97. [doi: 10.2196/jmir.4883] [Medline: 27185295]
40. Flaucher M, Pruemer F, Jaeger KM, et al. Motivational factors for experienced users of mobile health applications in heart failure management. 2024 Oct 13 Presented at: NordiCHI Adjunct 2024; Oct 31, 2024; Uppsala Sweden p. 1-5 URL: https://dl.acm.org/doi/proceedings/10.1145/3677045 [doi: 10.1145/3677045.3685447]
41. Peiris D, Miranda JJ, Mohr DC. Going beyond killer apps: building a better mHealth evidence base. BMJ Glob Health 2018;3(1):e000676. [doi: 10.1136/bmjgh-2017-000676] [Medline: 29527353]
42. Philip BJ, Abdelrazek M, Barnett S, Bonti A, Grundy J. Toward a unified mHealth platform: a survey of current user challenges and expectations. IEEE Access 2023;11:19876-19891. [doi: 10.1109/ACCESS.2023.3249786]
43. Alharthi SA. Bridging the digital divide in health for older adults: a repeated cross-sectional study of mHealth in Saudi Arabia. IEEE Access 2025;13:63757-63773. [doi: 10.1109/ACCESS.2025.3558368]
44. Pross T, Karsten MM, Blohmer JU. From gaps to solutions: semi-structured interviews to identify care gaps in breast cancer care and how to solve them with digital solutions. Geburtshilfe Frauenheilkd 2024 Sep;84(9):845-854. [doi: 10.1055/a-2369-1489] [Medline: 39229629]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
mHealth: mobile health
QoL: quality of life

XSL•FO
RenderX

Edited by M Balcarras, N Cahill; submitted 27.May.2025; peer-reviewed by L Weinert, S Fan; revised version received 08.Dec.2025; accepted 23.Dec.2025; published 07.Apr.2026.

Please cite as:
Flaucher M, Ott T, Nissen M, Dabrock P, Huebner H, Beckmann MW, Fasching PA, Leutheuser H, Eskofier BM, Mocker A Acceptance, Perceived Usefulness, and Data Sharing in Mobile Health Apps Among Patients With Breast Cancer: Cross-Sectional Survey Study
JMIR Cancer 2026;12:e77750
URL: https://cancer.jmir.org/2026/1/e77750
doi:10.2196/77750

© Madeleine Flaucher, Tabea Ott, Michael Nissen, Peter Dabrock, Hanna Huebner, Matthias W Beckmann, Peter A Fasching, Heike Leutheuser, Bjoern M Eskofier, Alexander Mocker. Originally published in JMIR Cancer (https://cancer.jmir.org), 7.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Adoption of Mobile Health Apps Among Patients with Head and Neck Cancer After the COVID-19 Pandemic: Secondary Analysis of Cross-Sectional Survey

Christian Wilhelm[1*], MSc, MD; Matthias Scheich[1*], MD; Michael Winter[2], MSc, PhD; Johannes Allgaier[3,4], PhD; Johanna Rippstein[1], MC; Sylke Ruth Zeissig[4,5], Prof Dr; Rüdiger Pryss[3,4], Prof Dr; Carsten Vogel[3,4], MSc; Johannes Schobel[6], Prof Dr; Katrin Radeloff[7], Prof Dr; Joanna Karolonek[7], MC; Agmal Scherzad[1], Prof Dr; Stephan Hackenberg[1], Prof Dr; Rudolf Hagen[1], Prof Dr

[1]Department of Otorhinolaryngology, Plastic, Aesthetic and Reconstructive Head and Neck Surgery, Julius Maximilian University of Wuerzburg, Schneider-Strasse 11, Wuerzburg, Germany

[2]University of Würzburg, Würzburg, Germany

[3]Institute for Medical Data Sciences, University of Wuerzburg, Wuerzburg, Germany

[4]Institute for Clinical Epidemiology and Biometry, University of Wuerzburg, Wuerzburg, Germany

[5]Regional Centre Wuerzburg, Bavarian Cancer Registry, Bavarian Health and Food Safety Authority, Wuerzburg, Germany

[6]DigiHealth Institute, Neu-Ulm University of Applied Sciences, Neu-Ulm, Germany

[7]Department of Otorhinolaryngology, Plastic, Aesthetic and Reconstructive Head and Neck Surgery, Evangelisches Krankenhaus Oldenburg, Medical Campus of the Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

[*]these authors contributed equally

Corresponding Author:
Christian Wilhelm, MSc, MD
Department of Otorhinolaryngology, Plastic, Aesthetic and Reconstructive Head and Neck Surgery, Julius Maximilian University of Wuerzburg, Schneider-Strasse 11, Wuerzburg, Germany

Abstract

Background: Patients with head and neck cancer (HNC) face significant challenges in accessing coordinated care due to the complexity and multimodality of their treatment and the impact on vital functions. The COVID-19 pandemic has disrupted cancer care while accelerating the adoption of digital health solutions. Mobile health (mHealth) apps offer potential solutions for remote symptom monitoring, communication between patients and providers, and continuity of care. Nevertheless, their acceptance among patients with HNC remains limited due to age-related digital divides and concerns about accessibility.

Objective: The aim of this study was to investigate the impact of the COVID-19 pandemic on patients with HNC and to explore their attitudes toward mHealth apps as a supplement to cancer treatment.

Methods: A secondary analysis of a cross-sectional survey was conducted between January 2023 and May 2024 at Julius Maximilian University of Würzburg in Germany. A total of 355 patients with HNC were recruited and completed the structured "Cancer and COVID-19" survey via the Corona Health app or in paper form. The 25-question survey assessed sociodemographic information, the impact of the pandemic on diagnosis/treatment/aftercare, and interest in mHealth apps. Descriptive statistics and bivariate correlations were used for the analysis.

Results: The cohort comprised 261 (74%) men and 94 (26%) women with an average age of 67 (SD 11.5) years. Most participants (n=264, 74%) stated that the pandemic had no impact on their cancer treatment, although 20% (n=71) experienced disruptions, particularly in follow-up appointments and treatment monitoring. Only 10% (n=36) currently used health apps, but 57% (n=203) expressed a willingness to use mHealth technologies. Younger patients, patients with higher education, and participants who were more affected by the pandemic showed greater openness to digital health solutions. The most significant barriers included age, digital literacy, and perceived usefulness, while preferred app features included interaction with physicians (n=160, 45%) and data sharing with researchers (n=153, 43%).

Conclusions: Although the COVID-19 pandemic had only a limited direct impact on HNC care at the institution at hand, it revealed significant patient interest in mHealth apps. However, significant barriers remain, particularly among older adults with lower digital literacy. Future mHealth initiatives should focus on improving digital literacy, addressing privacy concerns, demonstrating clinical benefits, and developing personalized, accessible solutions to optimize cancer care for this vulnerable population.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e65192) doi:10.2196/65192

KEYWORDS

head and neck cancer; COVID-19 pandemic; mHealth; digital health solutions; cancer care; telemedicine; digital literacy; health care access; patient-centered care; mobile health

Introduction

Head and neck cancer (HNC) is a particularly difficult form of cancer due to its impact on vital functions, such as speaking, swallowing, and breathing, as well as its demanding, multimodal treatment pathways [1]. The treatment of these patients requires continuous medical, psychosocial, and logistical support throughout all phases of care. Ensuring continuity of care and patient retention is therefore a key objective in head and neck oncology [2].

Despite advances in treatment, many patients with HNC face barriers to accessing timely and coordinated care [3]. Advanced age, multiple comorbidities, and the physical and emotional burden of the disease can limit patients' autonomy and complicate long-term follow-up care [4]. In addition, patients often report unmet needs in terms of communication, symptom management, and psychological support—highlighting gaps in traditional care structures [5].

Digital health technologies, particularly mobile health (mHealth) apps, offer new tools to bridge these gaps [6]. By supporting remote symptom monitoring, improving communication between patients and providers, and enabling self-management, mHealth can improve both the quality of care and patient autonomy [7-10]. However, acceptance among patients with HNC remains limited, partly due to age-related digital divides and a lack of tailored, accessible solutions [11,12].

The COVID-19 pandemic posed significant challenges to oncology care, including appointment postponements and treatment interruptions [13,14]. Simultaneously, it accelerated the adoption of telemedicine and digital health tools in many health care systems. For patients with HNC, the pandemic provided a unique opportunity to try out digital solutions—not just as an emergency measure, but as a permanent component of cancer care [15].

This study examines the experiences of patients with HNC during the COVID-19 pandemic and their perception of mHealth apps. Using self-reported patient data from the Corona Health app, we analyze disruptions in cancer treatment, patterns of digital health service use, and the most important factors that promote or hinder the adoption of mHealth. The results may, in turn, promote the development of accessible, patient-centered digital interventions for people with HNC.

Methods

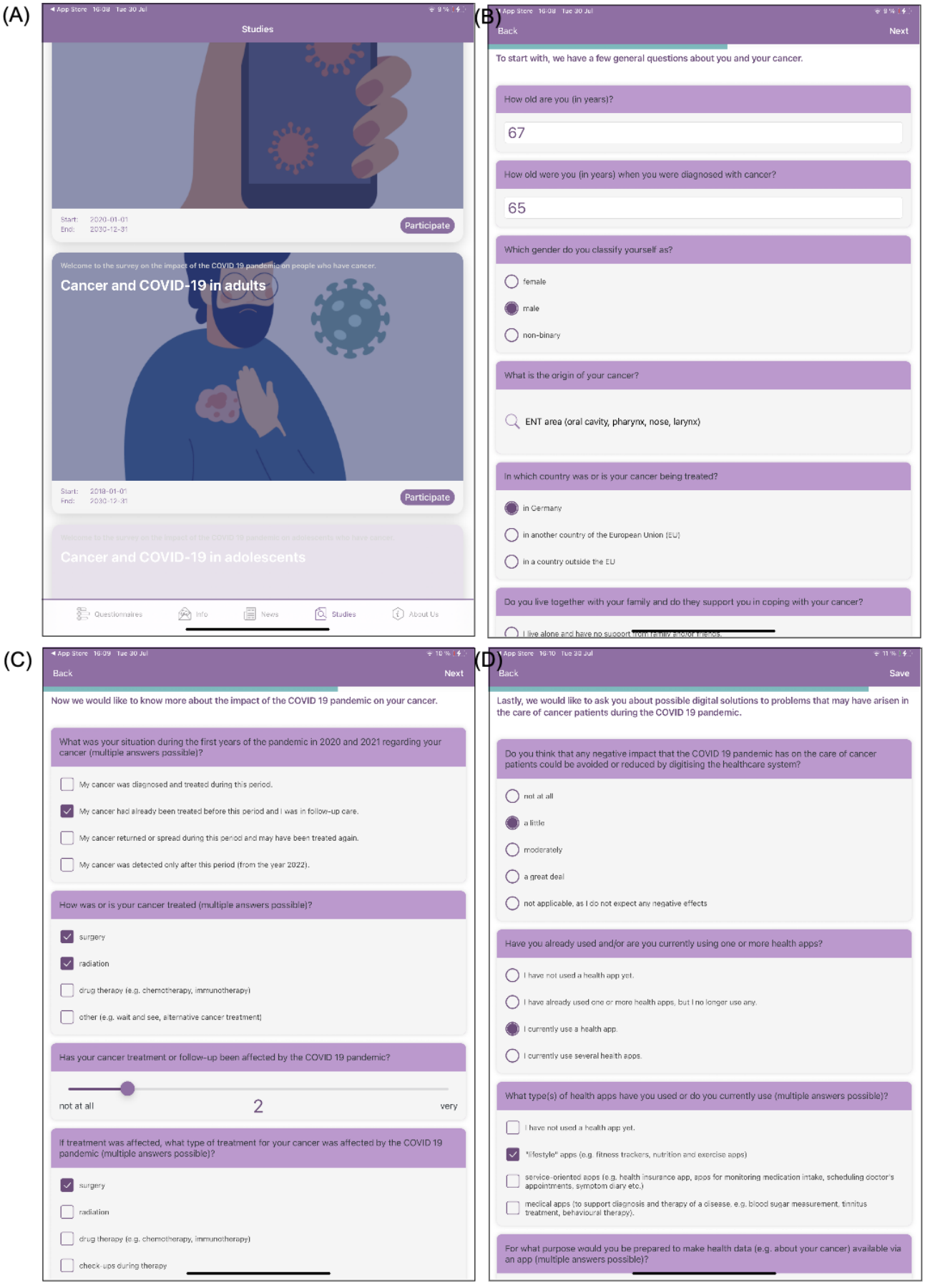
Study Design and Participants

This study is a secondary analysis of data collected via the Corona Health mobile app platform, a study- and sensor-based mHealth infrastructure developed to investigate various aspects of the COVID-19 pandemic and conduct multiple questionnaire-based studies [16]. Details on the technical implementation of the platform, data collection procedures, and characteristics of the initial cohort have been published by Beierle et al [16,17] (2021). For this analysis, we used cross-sectional data from patients with HNC, who participated through a dedicated study module integrated into the Corona Health app. Participants were recruited between January 2023 and May 2024 at the Department of Otorhinolaryngology, Plastic, Aesthetic, and Reconstructive Head and Neck Surgery at Julius Maximilian University of Würzburg, Germany. Recruitment took place primarily within the department's clinical infrastructure. The weekly tumor follow-up clinic for patients with HNC, which offers both initial consultations and follow-up appointments, was the main recruitment site, with approximately 35 to 40 patients participating per week. During waiting times, suitable patients were informed about the study and invited to participate. In addition, inpatients from 2 wards specializing in the diagnostic and therapeutic treatment of HNC were approached to participate. Patients with a confirmed HNC diagnosis were included in the study, whereas patients with cognitive impairments or language barriers that prevented participation were excluded.

Participants who were unable or unwilling to use the app completed a paper version of the same survey (see Figure 1), the responses to which were later digitized and entered into the database. The survey comprised 25 specially designed questions (Q 1 - 25) on participants' sociodemographic characteristics (Part A); the impact of the pandemic on diagnosis, treatment, and follow-up care (Part B); and their interest and willingness to use mHealth apps (Part C).

XSL•FO
RenderX

Figure 1. iOS in-app screenshots. (A) Overview of studies with the cancer and COVID-19 survey. (B) Part A of the survey: sociodemographic information. (C) Part B of the survey: the impact of the pandemic on cancer. (D) Part C of the survey: mobile health app use.



XSL•FO
RenderX

The response formats included single choice (Q 1 - 9, Q 20 - 22, Q 24), multiple choice (Q 10 - 11, Q 13, F 22 - 23, F 25), and matrix/rating scale questions (F 12, F 14 - 19) with 1 to 8 points to reduce central tendency bias. For analysis, scores were divided into 4 categories (1, 2-4, 5-7, 8) and labeled according to context (eg, never, rarely, sometimes, and often for F 16 - 17; no, low, moderate, and high for F 12, F 14 - 15). Trained health care professionals assisted participants in completing the survey.

The survey was developed by experts in epidemiology and ear, nose, and throat medicine as well as head and neck surgery at the University Hospital of Würzburg, Germany. Before large-scale use, it was tested in 3 oncology outpatient clinics with approximately 30 patients, who evaluated the clarity and relevance of the individual questions. Based on their feedback, questions were refined, added, or removed to improve comprehensibility and validity. This iterative process resulted in a robust, patient-centered instrument that enabled a meaningful assessment of the experiences and perspectives of patients with HNC during the pandemic. The study was also open to anyone who could download the app for free from the Google and Apple app stores.

Data Analysis

Survey responses were analyzed to identify patterns and significant findings related to the impact of the COVID-19 pandemic on HNC treatment and attitudes toward mHealth apps. Descriptive statistics were used to summarize demographic characteristics and survey items.

The relationships between demographic variables and survey variables were examined using Pearson correlation for continuous data and Spearman ρ for ordinal data. Statistical significance was set at $P \leq .05$, with interpretation emphasizing the strength and direction of correlations. Coefficients of approximately |0.10 - 0.29|, |0.30 - 0.49|, and ≥|0.50| were interpreted as weak, moderate, and strong associations, respectively [18]. Positive coefficients indicated that the variables increased together, whereas negative coefficients indicated inverse relationships.

The study was exploratory in nature and aimed to characterize the experiences of patients with HNC during the pandemic and their attitudes toward the use of mHealth. Due to the lack of previous data on mHealth acceptance in this population, formal performance calculations were not possible. The final sample of 355 participants (response rate: 60%) exceeded the typical minimum requirements for correlation analyses (30 - 50 participants per variable) and provided sufficient power (>80%) to detect medium effect sizes ($r \geq 0.15$) at $\alpha = .05$. Subgroup analyses among mHealth users (34/355, 10%) were found to be undersized, and the corresponding results should be interpreted as preliminary and validated in larger samples.

Ethical Considerations

The Corona Health app used was developed and implemented in accordance with medical device regulations and received the appropriate certification [19]. The University of Würzburg ethics committee approved the original data collection for this study (130/20-me), with explicit coverage for both initial collection and future secondary analyses of the data. The primary study describing the Corona Health platform and its data collection across multiple integrated studies has been published by Beierle et al [16]. Therefore, no additional institutional review board approval was required for this secondary analysis. Under local regulations and institutional guidelines [20], the secondary use of existing anonymized data does not require separate ethical review. All procedures complied with institutional ethical standards for research involving human subjects, and participant data were completely anonymized to ensure privacy protection. Participation was voluntary and no financial compensation was provided. Finally, this cross-sectional study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for reporting cross-sectional studies (Checklist 1).

Results

The results are divided into 3 main sections according to the survey structure: (1) demographic data and clinical characteristics of participants, (2) impact of the pandemic on cancer care, and (3) attitudes toward mHealth apps.

Part A: Participant Sociodemographics

A total of 355 individuals with head and neck malignant neoplasms participated in the study, including 261 (74%) males and 94 (26%) females, with a median age of 67 years (IQR 18 - 92 y; Q 1 and Q 3) at the time of the survey.

The median age at cancer diagnosis was 65 (IQR 55-70) years (Q 3). Time since diagnosis was calculated by subtracting age at diagnosis (Q 3) from the current age (Q 2). In addition, 80% (n=283) of the participants received their initial diagnosis within the past 5 years, with the proportion decreasing from 36% (n=102) diagnosed within the past year to 8% diagnosed 5 years ago.

All individuals received their treatment in Germany (Q 5). Additionally, 18% (n=65) of the participants had a history of other cancers, mainly gastrointestinal cancer (n=18, 5%), malignant melanoma (n=10, 3%), and blood or lymphatic cancer (n=10, 3%).

Half of the participants (n=179, 50%) had 8 to 9 years of school education, 158 (45%) had a higher educational level, and 18 (5%) had a lower educational level.

More than half of the participants (n=192, 54%) had completed 3 to 4 years of professional training or higher education. About a quarter (n=92, 26%) had a longer qualification, while a fifth (n=71, 20%) had a shorter qualification.

Approximately 3 quarters (n=261, 74%) of the participants live with and receive support from their families. One in 5 (n=72, 20%) participants lives alone but receives support from family and/or friends. A small number of individuals either live alone without support (n=16, 5%) or live with their families but receive little or no support (n=6, 2%).

At the time of the survey, approximately 1 out of 4 (n=90, 25%) participants was employed, while two-thirds (n=239, 68%) were unemployed. Of those who were unemployed, the majority were either doing housework and/or raising children, retired, or

XSL•FO
RenderX

preretired (n=233, 66%), with a small minority actively seeking work (n=6, 2%).

In summary, a comprehensive summary of all results for Part A can be found in Table 1.

XSL•FO
RenderX

Table . Participant demographics and clinical characteristics (see Table S1 in Multimedia Appendix 1 for the detailed overview of the questions).

Variable and category	Values, n (%)
Age by gender	
20-29	• Female: 4 (1.1) • Male: 0 (0)
30 - 39	• Female: 7 (2) • Male: 2 (0.6)
40 - 49	• Female: 4 (1.1) • Male: 8 (2.3)
50-59	• Female: 15 (4.2) • Male: 40 (11.3)
60 - 69	• Female: 34 (9.6) • Male: 106 (29.9)
70 - 79	• Female: 21 (5.9) • Male: 77 (21.7)
80 - 89	• Female: 8 (2.3) • Male: 27 (7.6)
90 - 99	• Female: 1 (0.3) • Male: 1 (0.3)
Total	• Female: 94 (26) • Male: 261 (74)
Time between survey and diagnosis (y)	
0	102 (28.7)
1	59 (16.6)
2	43 (12.1)
3	34 (9.6)
4	23 (6.5)
5	22 (6.2)
>5	72 (20.3)
Cancer entities (ie, origin of cancer)	
Renal	1 (0.3)
Urologic	3 (0.8)
Gynecologic	4 (1.1)
Lung	6 (1.7)
Blood/lymphatic	10 (2.8)
Melanoma	10 (2.8)
Other cancer	13 (3.7)
Gastrointestinal tract	18 (5.1)
HNC[a] only	290 (81.7)
School education (y)	
≤7	18 (5.1)
8-9	179 (50.4)
10	73 (20.6)
11-12	42 (11.8)

XSL•FO
RenderX

Variable and category	Values, n (%)
≥13	43 (12.1)
Job qualification (y)	
0	32 (9)
1-2	39 (11)
3-4	192 (54.1)
5-6	47 (13.2)
≥7	45 (12.7)
Living situation	
Alone, no support	16 (4.5)
Alone, supported by family/friends	72 (20.3)
With family, supported	261 (73.5)
With family, little/no support	6 (1.7)
Employment status	
Retired	233 (65.6)
Seeking work	6 (1.7)
Civil servant	79 (22.3)
Self-employed	11 (3.1)
Other	26 (7.3)

[a]HNC: head and neck cancer.

Part B: Impact of the COVID-19 Pandemic

Among 355 participants, 118 (33%) were diagnosed with cancer before the pandemic (2019 and earlier), 108 (30%) during the pandemic (2020 and 2021), and 126 (35%) of them after the pandemic (2022 and later). Of all respondents, only 13 stated that they had been treated due to a recurrence of their cancer.

Out of 355 participants, 290 (82%) underwent surgery as part of their treatment plan. Radiation therapy was included for 222 (63%) participants, and drug therapy (eg, chemotherapy and immunotherapy) was part of the treatment for 126 (35%) participants. Additionally, 11 (3%) participants received other therapies (eg, watchful waiting and alternative cancer treatments). Combination treatments were common: 170 (48%) participants had both surgery and radiation, 112 (32%) had both radiation and drug therapy, 111 (31%) had surgery only, 14 (4%) had radiation only, 5 (1%) had drug therapy only, and 7 (2%) received solely other therapies.

About 3 out of 4 participants reported no impact of the pandemic on their therapy and aftercare (n=264, 74%), while most of those who experienced disruptions considered them to be minor (n=57, 16%).

Out of 355 participants, 285 (80%) did not report any disruption in their therapy or aftercare. Among the remaining 70 (20%) participants, the most documented disruptions were in follow-up examinations (34, 10%) and checkups during therapy (21, 6%). Fewer participants experienced disruptions in radiation therapy (15, 4%), surgery (14, 4%), drug therapy (10, 3%), and supportive therapy (9, 3%).

More than half of the participants reported no fear of SARS-CoV-2 infection during (n=204, 57%) or after the pandemic (n=205, 58%). Approximately 1 in 4 participants experienced little fear during (n=94, 26%) and slightly more after the pandemic (n=115, 32%). Moderate fear declined from 11% (n=39) during the pandemic to 7% (n=24) after the pandemic, as did very strong fear (n=18, 5% during; n=11, 3% after).

Most participants did not experience cancellations of appointments by the clinic due to the pandemic (n=299, 84%), and even fewer patients canceled appointments themselves (n=309, 87%). About 1 in 10 participants had appointments rarely canceled by the clinic (n=47, 13%) and rarely canceled appointments themselves (n=40, 11%). Almost no participants faced regular appointment cancellations.

Around two-thirds (n=223, 63%) of the participants had no psychological effects due to the pandemic at all, 101 (28%) participants had minor effects, and only 26 (7%) and 5 (1%) of them were affected moderately or highly.

Similarly, around two-thirds of the participants had no fear of negative effects of the pandemic on their own cancer therapy (n=234, 66%), 102 (29%) participants had minor fear, while 16 (5%) and 3 (1%) documented moderate or major fear, respectively.

In summary, a comprehensive summary of all results for Part B can be found in Table 2.

XSL•FO
RenderX

Table . Treatment characteristics and COVID-19 pandemic impact (see Table S1 in Multimedia Appendix 1 for detailed overview of the questions).

Variable and category	Values, n (%)
Date of cancer diagnosis	
2019 and earlier	118 (33.5)
2020 - 2021	108 (30.7)
2022 and later	126 (35.8)
Type of therapy	
Surgery	290 (81.7)
Radiation	222 (62.5)
Drug therapy	126 (35.5)
Other	11 (3.1)
Surgery only	111 (31.3)
Radiation only	14 (3.9)
Drug therapy only	5 (1.4)
Other only	7 (2)
Surgery+radiation	170 (47.9)
Radiation+drug therapy	112 (31.5)
Impact on therapy and aftercare	
None	264 (74)
Minor	57 (16)
Moderate	28 (8)
Major	6 (2)
Affected type of therapy (n=70/355)	
Supportive therapy	9 (2.5)
Drug therapy	10 (2.8)
Surgery	14 (3.9)
Radiation	15 (4.2)
Therapy checkups	21 (5.9)
Follow-up	34 (9.6)
Not affected	285 (80.3)
Fear of SARS-CoV-2 infection	
During the pandemic	
No	204 (57.7)
Minor	94 (26.5)
Moderate	39 (11)
Major	18 (5.1)
After the pandemic	
No	205 (57.7)
Minor	115 (32.4)
Moderate	24 (6.8)
Major	11 (3.1)
Cancellation of appointments	
Clinic	
Never	299 (84.2)

XSL•FO
RenderX

Variable and category	Values, n (%)
Rarely	47 (13.2)
Sometimes	5 (1.4)
Frequently	4 (1.1)
Patient	
Never	309 (87)
Rarely	49 (11.3)
Sometimes	6 (1.7)
Frequently	0 (0)
Psychic stress due to pandemic	
None	223 (62.8)
Minor	101 (28.5)
Moderate	26 (7.3)
Major	5 (1.4)
Fear of negative effects on cancer therapy	
None	234 (65.9)
Minor	102 (28.7)
Moderate	16 (4.5)
Major	3 (0.8)

Part C: mHealth App Use

About 1 out of 4 (n=83, 23%) participants expected no positive effect of digital health solutions on cancer care, whereas 1 out of 3 (n=114, 32%) expected minor effects. Only 19% (n=67) and 10% (n=34) of the participants anticipated moderate or major effects. Additionally, 16% (n=57) of the participants saw no need for digitalization in cancer care at all.

About 1 out of 4 participants (n=83, 23%) expected no positive effect of digital health solutions on cancer care, whereas 1 out of 3 (n=114, 32%) expected minor effects.

Only 10% (n=34) of participants were currently using health apps, with 8% (n=27) using only one health app and 2% (n=7) using more than one health app. Additionally, 6% (n=23) had formerly used a health app, whereas the majority of the participants (n=298, 84%) had never used a health app.

Out of the 57 (16%) app users, about half had been using lifestyle apps (eg, fitness trackers, nutrition, and exercise apps; n=28, 49%), and every fourth user had been using a service-oriented health app (eg, health insurance apps, apps for monitoring medication intake, scheduling doctor's appointments, and symptom diaries, etc; n=14, 25%) or medical apps (eg, apps supporting diagnosis and therapy of diseases, such as blood sugar measurement, tinnitus treatment, and behavioral therapy; n=15, 26%).

About two-thirds of the participants expressed willingness to share their health data using a mHealth app. The most frequent possible recipients identified were researchers (n=135, 43%), therapists (n=134, 38%), and the participants themselves (n=112, 32%). Only a minority of participants indicated willingness to share their data for self-help purposes (n=58, 16%) or publicly (n=44, 12%).

Out of 355 participants, 151 (43%) indicated that they would never use a health app to record symptoms of their cancer. The responses varied among those willing to use such an app: 103 (29%) participants preferred to use it symptom-dependent rather than at a fixed frequency. Specifically, 38 (11%) participants would use it once, whereas 63 (18%) participants would use it frequently, with breakdowns as follows: every 3 months (n=15, 4%), monthly (n=15, 4%), weekly (n=15, 4%), multiple times weekly (n=8, 2%), daily (n=7, 2%), and multiple times daily (n=3, 1%).

Asked about their desired interaction within a cancer app, 145 out of 355 (41%) indicated that they would not use it. Among potential users, preferences for interaction varied: 158 (45%) participants expressed interest in interacting with physicians, whereas 98 (28%) participants wanted to interact with researchers. An even smaller proportion preferred interaction with nurses (n=41, 12%) or engaging in self-care activities (n=41, 12%). Only 15 (4%) participants stated they did not desire any interaction within the app.

In summary, a comprehensive summary of all results for Part C can be found in Table 3.

XSL•FO
RenderX

Table . Digital health attitudes and app usage patterns (see Table S1 in Multimedia Appendix 1 for the detailed overview of the questions).

Variable and category	Values, n (%)
Effect of digital health on cancer care	
No effect	83 (23.4)
Minor	114 (32.1)
Moderate	67 (18.9
Major	34 (9.6)
Not necessary	57 (16.1)
Health app usage	
Nonuser	298 (83.9)
Ex-user	23 (6.5)
Single-app user	27 (7.6)
Multiapp user	7 (2)
Type of health app used (n=57/355)	
No app	297 (83.9)
Lifestyle	28 (7.9)
Service	14 (4)
Medical	15 (4.2)
Willingness to share health data	
Public	44 (12.4)
Self-help groups	58 (16.3)
For personal use	112 (31.5)
Not at all	120 (33.8)
Therapists	134 (37.7)
Researchers	153 (43.1)
Intended frequency of app use (n=63/355)	
Multiple times daily	3 (0.8)
Daily	7 (2)
Multiple times weekly	8 (2.3)
3× monthly	15 (4.2)
Monthly	15 (4.2)
Weekly	15 (4.2
Once	38 (10.7)
Symptom-dependent	103 (29)
Never	151 (42.5)
Desired interaction within the cancer app	
Nobody	15 (4.2)
Self-care	41 (11.5)
Nurses	41 (11.5)
Researchers	98 (27.6)
No app use	145 (40.8)
Physicians	158 (44.5)

XSL•FO
RenderX

Relationships Between Patient Characteristics and Study Outcomes

Furthermore, we explored how demographic variables correlated with survey responses. Depending on the measurement level of the variables, either Pearson (for continuous variables) or Spearman (for ordinal variables) correlations were applied, revealing several significant relationships: older age was negatively associated with app usage (P<.001, r=−0.20), psychological stress (P=.009, r=−0.14), and the expected effect of digital solutions on cancer care (P=.39, r=−0.11). In addition, we found the following:

- Male participants showed a negative correlation with fear of SARS-CoV-2 infection during (P=.03, r=−0.11) and after the pandemic (P=.03, r=−0.11).
- Higher education positively correlated with app usage (P<.001, r=0.17) and higher professional qualification with expected effect of digital solutions (P=.049, r=0.10). Moreover, a higher educational level was associated with more significant concern about pandemic-related impacts on personal cancer outcomes (P=.05, r=0.15).
- Perceived disruptions of cancer care due to the pandemic, fear of personal repercussions, and cancellation of appointments by the clinic (P<.001, r=0.39) or by the patient (P<.001, r=0.34) were all positively correlated with fear of SARS-CoV-2 infection and psychological stress (all P<.001, r=0.39; P<.001, r=0.59).
- The perceived impact of the pandemic on participants' cancer care was positively correlated with app usage (P=.03, r=0.12).

To understand relationships between patient characteristics and study outcomes, we examined bivariate correlations between key variables. Table 4 presents all statistically significant correlations, organized to show how demographic factors (age, gender, education, and qualification), pandemic experiences (impact, fear, and cancellations), and mHealth attitudes (digital solutions, app usage, and frequency) relate to one another.

XSL•FO
RenderX

Table . Overview of all statistically significant Pearson and Spearman correlations[a].

Variable 1 and variable 2	*P* value	*r*
Age		
• Gender	<.001	0.18[b]
• Education	<.001	−0.25[b]
• Employment	<.001	−0.64[c]
• Psychic stress	.009	−0.14[b]
• App usage	<.001	−0.20[b]
• Digital solution	.04	−0.11[b]
• Frequency	<.001	−0.24[b]
Gender		
• Fear before	.04	−0.11[b]
• Fear now	.04	−0.11[b]
Education		
• Qualification	<.001	0.33[d]
• Negative effects	.005	0.15[b]
• Digital solution	.07	0.10
• App usage	<.001	0.17[b]
• Frequency	<.001	0.24[b]
Qualification		
• Cancellation (clinic)	<.001	0.18
• Cancellation (patient)	.007	0.15
• Digital solution	.049	0.10
• App usage	.10	0.09
Impact		
• Fear before	<.001	0.39
• Fear now	<.001	0.31
• Cancellation (clinic)	<.001	0.43
• Cancellation (patient)	<.001	0.34
• Psychic stress	<.001	0.39
• Negative effects	<.001	0.45
• Digital solution	.10	0.09
• App usage	.03	0.12
Negative effects		

XSL•FO
RenderX

Variable 1 and variable 2		*P* value	*r*
	• Fear before	<.001	0.45
	• Fear now	<.001	0.55
	• Cancellation (clinic)	<.001	0.44
	• Cancellation (patient)	<.001	0.35
	• Psychic stress	<.001	0.59
	• Digital solution	.03	0.11
	• Frequency	.003	0.16

[a]Bivariate correlations were calculated between demographic characteristics, pandemic experiences, and mHealth attitudes. Depending on the measurement level of the variables, either Pearson (for continuous variables) or Spearman (for ordinal variables) correlations were used. Correlation coefficients (*r* or ρ) indicate the strength and direction of the relationship: weak correlations=0.10≤*r*≤0.29, moderate correlations=0.30≤*r*≤0.49, else strong correlation. Positive coefficients indicate that variables increase together, whereas negative coefficients indicate that one variable increases as the other decreases.

[b]Weak correlation.

[c]Strong correlation.

[d]Moderate correlation.

Note that it must be emphasized that correlations are not a robust measure of association and do not allow causal inference. Moreover, the analyses did not account for potential confounding variables.

Discussion

Overview of Findings

This study provides new insights into the experiences of patients with HNC in terms of diagnosis, treatment, and follow-up care, as well as their attitudes toward mHealth apps. Although most patients reported no major disruptions to their cancer treatment, a significant proportion (20% - 26%) experienced interruptions—particularly in follow-up appointments and treatment-related checkups. This subgroup highlights the importance of maintaining continuity of care for all patients, even during health care crises.

Correlation analyses revealed several notable patterns. Older participants were less likely to use digital health solutions, whereas higher levels of education were associated with greater acceptance of mHealth. In addition, patients who had experienced pandemic-related disruptions to care reported both increased anxiety and greater interest in digital alternatives. The observed correlations between pandemic-related anxiety, disruptions in care, and psychological distress (r=0.27 - 0.49) suggest that these factors form an interconnected cluster of pandemic-related health care anxieties.

Pandemic Impact on Cancer Care

The COVID-19 pandemic has posed unprecedented challenges to health care systems worldwide and necessitated rapid adjustments in cancer care [21]. At our institution, treatment pathways remained largely stable; however, 20% - 26% of patients reported interruptions in care, which is consistent with international observations of treatment delays, limited access to support services, and changed follow-up appointments [22,23]. Previous studies on head and neck oncology also document pandemic-related delays and increased psychological distress [23], while short-term survival outcomes appear to have remained largely unaffected [24]. These findings suggest that rapid institutional adjustments have effectively mitigated serious clinical consequences despite significant logistical and emotional challenges.

Potential of Digital Health Solutions

Despite current low usage (10%), patients showed significant interest in mHealth solutions for symptom monitoring, communication with health care providers, and data sharing. Greater acceptance was associated with younger age, higher education level, and pandemic-related disruptions to care, consistent with general trends in telemedicine adoption during COVID-19 [25], although regulatory barriers remain in many countries [26].

mHealth offers opportunities for remote symptom monitoring, timely interventions, and personalized patient education. Feasibility in patients with HNC has been demonstrated in pilot studies [27] and randomized controlled trials [28]. However, large-scale implementation continues to be limited by early discontinuation, challenges with scalability, and limited integration into routine clinical workflows [29].

Barriers to Adoption

The findings reveal several barriers to the adoption of mHealth that are consistent with previous research findings. Age, digital literacy, and perceived usefulness remain important determinants of acceptance [30]. Although most participants stated that they had minimal concerns about data security within the university, previous studies show that data protection outside clinical research environments remains a major barrier [31,32]. This is particularly relevant in Germany, where data protection concerns are often cited as a significant barrier to the adoption of digital health services [32].

XSL•FO
RenderX

Implications for Practice

Future strategies to enhance mHealth adoption in HNC care should include the following:

1. Improving digital literacy—targeted training programs for older adults and individuals with lower educational levels [33].
2. Addressing privacy concerns—transparent communication of robust data security measures [31].
3. Demonstrating utility—evidence-based, user-friendly, and personalized app functionalities to enhance trust and sustained use [34].

Previous studies suggest that despite ongoing concerns about implementation, oncologists are generally open to integrating mHealth technologies into patient care [35]. By removing barriers on both the patient and provider sides, mHealth could make the transition from a complementary tool to an integral part of multidisciplinary cancer treatment.

Although the COVID-19 pandemic was unprecedented, the challenges it revealed—such as scheduling constraints, communication gaps between visits, the need for remote monitoring, and increased patient anxiety about accessing health care—are not unique to pandemics. Similar disruptions also occur during natural disasters, staff shortages, transportation barriers, and routine capacity constraints. Our findings that 57% (n=202) of patients were willing to share health data and 45% (n=160) preferred to communicate with their doctors using an app provide actionable insights for strengthening continuity of care under various disruptive conditions.

The rapid digital transformation triggered by COVID-19 has permanently changed the expectations of patients and providers. The pandemic has accelerated a decade of digital health development in a matter of months, revealing both barriers and enablers that will continue to be relevant as health systems drive this transformation forward. The demographic predictors identified in our study—age, education, and prior experience with health care—represent enduring factors that are likely to influence mHealth adoption regardless of future circumstances.

Strengths and Limitations

This study focused on a high-need population—patients with HNC—who often face complex treatments and intensive follow-up care requirements. By examining both pandemic-related disruptions to care and the introduction of digital health solutions within the same cohort, it offers a holistic perspective on health care resilience.

However, there are some limitations to consider. The study design, conducted at a single center in Würzburg, limits generalizability, and the relatively small, specialized cohort may not reflect the broader cancer population. Nonresponse bias cannot be ruled out, as the participation of nonrespondents was not tracked. The use of self-reported data may have led to bias due to recall or social desirability, although the survey design aimed to minimize these effects. The low prevalence of mHealth users (34/355, 10%) limited subgroup analyses and restricts the generalization of user behavior. Finally, correlation analyses do not allow conclusions about causality and did not control for potential confounding factors. Future multicenter studies with larger, more diverse samples and causal modeling approaches are needed to validate and extend these findings.

Conclusion

This study highlights both the challenges and opportunities faced by patients with HNC during and after the COVID-19 pandemic. While most patients continued to have access to essential care, a subset experienced significant disruptions, underscoring the need for resilient care pathways. The pandemic has also sparked interest in mHealth apps; however, barriers remain in terms of age, digital literacy, and perceived usefulness. Targeted education, robust data protection, and clear evidence of clinical benefit will be critical to integrating mHealth into cancer care and improving outcomes and quality of life for patients with HNC.

Acknowledgments

Assistance in writing and editing this manuscript was provided by ChatGPT, a large language model developed by OpenAI. The authors reviewed and verified the accuracy of all content generated with the assistance of this tool.

Funding

This publication was funded by the German Federal Ministry of Education and Research (BMBF) Network of University Medicine 2.0: "NUM 2.0," grant number 01KX2121, project: [Collateral Effects of Pandemics – CollPan].

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

CW and MS conceptualized and designed the study and jointly wrote and reviewed the manuscript. JR conducted the survey. JA, CV, MW, JS, and RP developed and managed the Corona Health app and were responsible for data processing. MW, JS, and RP also contributed to manuscript review. SRZ assisted in the development of the questionnaire and reviewed the manuscript. KR and JK supported the study's implementation and coordinated its continuation at an additional site. AS, SH, and RH oversaw study implementation and participated in the manuscript review.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table S1. Complete cancer and COVID-19 survey within the Corona Health app.
[DOCX File, 19 KB - cancer_v12i1e65192_app1.docx]

Checklist 1
STROBE checklist.
[DOCX File, 33 KB - cancer_v12i1e65192_app2.docx]

References

1. Wirth M, Schmidl B, Wollenberg B, Pigorsch S. Editorial: challenges and their implications for the clinical practice of head and neck cancer. Front Oncol 2023;13:1131639. [doi: 10.3389/fonc.2023.1131639] [Medline: 36741021]
2. Machiels JP, René Leemans C, Golusinski W, et al. Squamous cell carcinoma of the oral cavity, larynx, oropharynx and hypopharynx: EHNS-ESMO-ESTRO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2020 Nov;31(11):1462-1475. [doi: 10.1016/j.annonc.2020.07.011] [Medline: 33239190]
3. Nieminen M, Aro K, Mäkitie A, et al. Challenges in diagnosing head and neck cancer in primary health care. Ann Med 2021 Dec;53(1):26-33. [doi: 10.1080/07853890.2020.1802060] [Medline: 32746635]
4. Szturz P, Van Laer C, Simon C, Van Gestel D, Bourhis J, Vermorken JB. Follow-up of head and neck cancer survivors: tipping the balance of intensity. Front Oncol 2020;10:688. [doi: 10.3389/fonc.2020.00688] [Medline: 32435619]
5. Reich M, Leemans CR, Vermorken JB, et al. Best practices in the management of the psycho-oncologic aspects of head and neck cancer patients: recommendations from the European Head and Neck Cancer Society Make Sense Campaign. Ann Oncol 2014 Nov;25(11):2115-2124. [doi: 10.1093/annonc/mdu105] [Medline: 24608199]
6. Giansanti D. The digital health: from the experience of the COVID-19 pandemic onwards. Life (Basel) 2022 Jan 6;12(1):78. [doi: 10.3390/life12010078] [Medline: 35054470]
7. Gomes MS, Bonan PRF, Ferreira VYN, et al. Development of a mobile application for oral cancer screening. Technol Health Care 2017;25(2):187-195. [doi: 10.3233/THC-161259] [Medline: 27689559]
8. Meng L, Zhao Y, Xu X. WeChat app in the follow up of thyroid cancer patients after thyroidectomy during the COVID-19 pandemic. Br J Surg 2020 Oct;107(11):e533. [doi: 10.1002/bjs.12009] [Medline: 32871024]
9. Chen ZJ, Liang WT, Liu Q, et al. Use of a remote oncology pharmacy service platform for patients with cancer during the COVID-19 pandemic: implementation and user acceptance evaluation. J Med Internet Res 2021 Jan 21;23(1):e24619. [doi: 10.2196/24619] [Medline: 33395398]
10. Teckie S, Solomon J, Kadapa K, et al. A mobile patient-facing app for tracking patient-reported outcomes in head and neck cancer survivors: single-arm feasibility study. JMIR Form Res 2021 Mar 19;5(3):e24667. [doi: 10.2196/24667] [Medline: 33739291]
11. Kessel KA, Vogel MM, Kessel C, et al. Mobile health in oncology: a patient survey about app-assisted cancer care. JMIR Mhealth Uhealth 2017 Jun 14;5(6):e81. [doi: 10.2196/mhealth.7689] [Medline: 28615159]
12. Duman-Lubberding S, van Uden-Kraan CF, Jansen F, et al. Feasibility of an eHealth application “OncoKompas” to improve personalized survivorship cancer care. Support Care Cancer 2016 May;24(5):2163-2171. [doi: 10.1007/s00520-015-3004-2] [Medline: 26563178]
13. Richards M, Anderson M, Carter P, Ebert BL, Mossialos E. The impact of the COVID-19 pandemic on cancer care. Nat Cancer 2020 Jun;1(6):565-567. [doi: 10.1038/s43018-020-0074-y] [Medline: 35121972]
14. Boniface D, Tapia-Rico G. Oncology during the COVID-19 pandemic: a lockdown perspective. Curr Oncol Rep 2022 Oct;24(10):1219-1235. [doi: 10.1007/s11912-022-01301-4] [Medline: 35622263]
15. Prasad A, Brewster R, Newman JG, Rajasekaran K. Optimizing your telemedicine visit during the COVID-19 pandemic: practice guidelines for patients with head and neck cancer. Head Neck 2020 Jun;42(6):1317-1321. [doi: 10.1002/hed.26197] [Medline: 32343458]
16. Beierle F, Schobel J, Vogel C, et al. Corona Health-a study- and sensor-based mobile app platform exploring aspects of the COVID-19 pandemic. Int J Environ Res Public Health 2021 Jul 10;18(14):7395. [doi: 10.3390/ijerph18147395] [Medline: 34299846]
17. Vogel C, Pryss R, Schobel J, Schlee W, Beierle F. Developing apps for researching the COVID-19 pandemic with the TrackYourHealth platform. 2021 Presented at: 2021 IEEE/ACM 8th International Conference on Mobile Software Engineering and Systems (MobileSoft); May 17-19, 2021. [doi: 10.1109/MobileSoft52590.2021.00015]
18. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg 2018;126(5):1763-1768. [doi: 10.1213/ANE.0000000000002864]
19. Holfelder M, Mulansky L, Schlee W, et al. Medical device regulation efforts for mHealth apps during the COVID-19 pandemic—an experience report of Corona Check and Corona Health. J 2021;4(2):206-222. [doi: 10.3390/j4020017]

XSL•FO
RenderX

20. Winter M, Vogel C, Schobel J, et al. Longitudinal mental health data collected via the Corona Health smartphone app during COVID-19. Sci Data 2026 Mar 11;13(1):392. [doi: 10.1038/s41597-026-07015-7] [Medline: 41813698]
21. Lai AG, Pasea L, Banerjee A, et al. Estimated impact of the COVID-19 pandemic on cancer services and excess 1-year mortality in people with cancer and multimorbidity: near real-time data on cancer care, cancer deaths and a population-based cohort study. BMJ Open 2020 Nov 17;10(11):e043828. [doi: 10.1136/bmjopen-2020-043828] [Medline: 33203640]
22. Kaufman HW, Chen Z, Niles J, Fesko Y. Changes in the number of US patients with newly identified cancer before and during the coronavirus disease 2019 (COVID-19) pandemic. JAMA Netw Open 2020 Aug 3;3(8):e2017267. [doi: 10.1001/jamanetworkopen.2020.17267] [Medline: 32749465]
23. Wilhelm C, Radeloff K, Scherzad A, Scheich M, Hagen R. COVID-19: collateral damage in head and neck oncology and preventive measures for future pandemics. Laryngorhinootologie 2023 Feb;102(2):104-110. [doi: 10.1055/a-2007-2002] [Medline: 36750111]
24. Balchander D, Shorbaji K, Cabrera CI, et al. Prognostic significance of time trends in treatment of head and neck squamous cell carcinoma. Am J Otolaryngol 2023;44(6):103966. [doi: 10.1016/j.amjoto.2023.103966] [Medline: 37481899]
25. Elhakeem I, Iqbal P, Nashwan AJ, et al. Patients' experience and satisfaction using telemedicine for outpatient services in a tertiary cancer center in Qatar during COVID-19: a cross-sectional study. Health Sci Rep 2022 Nov;5(6):e883. [doi: 10.1002/hsr2.883] [Medline: 36320657]
26. Lu DJ, Girgis M, David JM, Chung EM, Atkins KM, Kamrava M. Evaluation of mobile health applications to track patient-reported outcomes for oncology patients: a systematic review. Adv Radiat Oncol 2021;6(1):100576. [doi: 10.1016/j.adro.2020.09.016] [Medline: 33073061]
27. Zini EM, Lanzola G, Quaglini S, Bossi P, Licitra L, Resteghini C. A pilot study of a smartphone-based monitoring intervention on head and neck cancer patients undergoing concurrent chemo-radiotherapy. Int J Med Inform 2019 Sep;129:404-412. [doi: 10.1016/j.ijmedinf.2019.06.004] [Medline: 31445284]
28. Sprave T, Pfaffenlehner M, Stoian R, et al. App-controlled treatment monitoring and support for patients with head and neck cancer undergoing radiotherapy: results from a prospective randomized controlled trial. J Med Internet Res 2023 Oct 19;25:e46189. [doi: 10.2196/46189] [Medline: 37856185]
29. Ardito V, Golubev G, Ciani O, Tarricone R. Evaluating barriers and facilitators to the uptake of mhealth apps in cancer care using the consolidated framework for implementation research: scoping literature review. JMIR Cancer 2023 Mar 30;9:e42092. [doi: 10.2196/42092] [Medline: 36995750]
30. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
31. Benjumea J, Ropero J, Rivera-Romero O, Dorronzoro-Zubiete E, Carrasco A. Privacy assessment in mobile health apps: scoping review. JMIR Mhealth Uhealth 2020 Jul 2;8(7):e18868. [doi: 10.2196/18868] [Medline: 32459640]
32. Giebel GD, Speckemeier C, Abels C, et al. Problems and barriers related to the use of digital health applications: scoping review. J Med Internet Res 2023 May 12;25:e43808. [doi: 10.2196/43808] [Medline: 37171838]
33. Wang X, Luan W. Research progress on digital health literacy of older adults: a scoping review. Front Public Health 2022;10:906089. [doi: 10.3389/fpubh.2022.906089]
34. Vaffis S, Whaley S, Axon DR, et al. Features of cancer mHealth apps and evidence for patient preferences: scoping literature review. JMIR Cancer 2023 Apr 28;9:e37330. [doi: 10.2196/37330] [Medline: 37115587]
35. Berkowitz CM, Zullig LL, Koontz BF, Smith SK. Prescribing an app? Oncology providers' views on mobile health apps for cancer care. JCO Clin Cancer Inform 2017 Nov;1(1):1-7. [doi: 10.1200/CCI.17.00107] [Medline: 30657404]

Abbreviations

HNC: head and neck cancer
mHealth: mobile health
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by N Cahill; submitted 07.Aug.2024; peer-reviewed by M Haug, T Mitsunaga; accepted 28.Oct.2025; published 15.Apr.2026.

Please cite as:
Wilhelm C, Scheich M, Winter M, Allgaier J, Rippstein J, Zeissig SR, Pryss R, Vogel C, Schobel J, Radeloff K, Karolonek J, Scherzad A, Hackenberg S, Hagen R
Exploring the Adoption of Mobile Health Apps Among Patients with Head and Neck Cancer After the COVID-19 Pandemic: Secondary Analysis of Cross-Sectional Survey
JMIR Cancer 2026;12:e65192
URL: https://cancer.jmir.org/2026/1/e65192
doi:10.2196/65192

XSL•FO
RenderX

© Christian Wilhelm, Matthias Scheich, Michael Winter, Johannes Allgaier, Johanna Rippstein, Sylke Ruth Zeissig, Rüdiger Pryss, Carsten Vogel, Johannes Schobel, Katrin Radeloff, Joanna Karolonek, Agmal Scherzad, Stephan Hackenberg, Rudolf Hagen. Originally published in JMIR Cancer (https://cancer.jmir.org), 15.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of the ALIBIRD mHealth Platform for Care of Patients With Lung Cancer: Prospective Pilot Study

Jose M Iniesta-Chamorro[1,2*], MSc; Maria Sereno[3,4*], MD, PhD; Beatriz Garrido-Rubiales[3,4], BPharm; Marta Villarino Sanz[3,4], PhD; Isabel Espinosa-Salinas[5], PhD; Sandra Falagán Martínez[4], MD; Gustavo Rubio Romero[4], MD; Juan Moreno-Rubio[3,4], PhD; Beatriz Tabarés[4], MD, PhD; Yolanda Martín[4], BSc; Irene Hernández de Córdoba[4], MD; Maria Morales Parga[4], MD; Natalia Molinero[6], PhD; M Victoria Moreno-Arribas[6], PhD; Guillermo Reglero Rada[6,7], PhD; Ana Ramírez de Molina[5,7], PhD; Enrique J Gómez[1,2,8], PhD; Enrique Casado[3,4], MD, PhD

[1]Biomedical Engineering and Telemedicine Centre, ETSI Telecomunicación, Center for Biomedical Technology, Universidad Politécnica de Madrid, Madrid, Spain

[2]Instituto de Investigación Hospital 12 de Octubre (imas12), Hospital Universitario 12 de Octubre, Madrid, Spain

[3]Clinical Oncology Group, IMDEA Nutrition, CEI UAM + CSIC, Madrid, Spain

[4]Medical Oncology Department, Infanta Sofía University Hospital, Infanta Sofía and Henares Hospitals Foundation for Biomedical Research and Innovation (FIIB HUIS HHEN), Paseo de Europa, 34, San Sebastián de los Reyes, Madrid, Spain

[5]GENYAL Platform, IMDEA Nutrition, CEI UAM + CSIC, Madrid, Spain

[6]Institute of Food Science Research (CIAL), Consejo Superior de Investigaciones Científicas (CSIC)–Universidad Autónoma de Madrid (UAM), Madrid, Spain

[7]Molecular Oncology Group, IMDEA Nutrition, CEI UAM + CSIC, Madrid, Spain

[8]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), Instituto de Salud Carlos III, Madrid, Spain

[*]these authors contributed equally

Corresponding Author:
Maria Sereno, MD, PhD
Clinical Oncology Group, IMDEA Nutrition, CEI UAM + CSIC, Madrid, Spain

Abstract

Background: Mobile health (mHealth) represents a promising instrument for optimizing symptom management and important lifestyle strategies that enhance self-care and the quality of health care for patients with cancer. The ALIBIRD mHealth platform is a digital health solution specifically designed for the telemonitoring of oncology patients, fostering patient empowerment and supporting clinical decision-making.

Objective: The primary objective of this study was to evaluate the patient experience with the ALIBIRD platform. In addition, the study aimed to assess clinical outcomes, particularly in symptom management, nutritional status, and lifestyle, using patient-reported outcome measures (PROMs).

Methods: The evaluation was conducted over a 30-week period in patients with advanced lung cancer receiving active treatment. Outcome variables included usability, patient experience, symptom burden, lifestyle behaviors (diet, physical activity, and sleep), nutritional status, PROMs, and system-generated clinical alerts. Through the mobile app, patients reported symptoms and completed integrated REDCap (Research Electronic Data Capture) questionnaires assessing lifestyle behaviors and PROMs, while receiving personalized recommendations informed by nutrigenetic and gut microbiota assessments. Daily activity and sleep data were automatically captured using the Fitbit Inspire wearable. Clinicians remotely monitored patient data using a web-based dashboard and performed clinical actions when required, including phone calls, therapeutic adjustments, referrals, and appointment rescheduling. Statistical analysis included descriptive summaries and pre-post comparisons of clinical and patient-reported outcomes.

Results: Out of 20 patients recruited for the study, 14 completed the intervention. The System Usability Scale yielded a score of 90, indicating high usability. Among the 14 completers, adherence to scheduled questionnaires ranged from 94% to 100% for several instruments, and wearable-based monitoring ranged from 66% to 96% across visits. Overall, the ALIBIRD platform collected and processed 3589 patient-reported outcomes related to physical activity, 3468 related to sleep, 679 on-demand symptom entries, and 1524 completed questionnaires. Clinically, 143 alerts were resolved within an average of 2.05 days, resulting in 2 referrals to emergency rooms and 2 early detections of disease progressions. Furthermore, more than 2100 personalized recommendations contributed to a 21% (3/14 patients) increase in adherence to the Mediterranean diet and a 14% (2/14 patients) increase in moderate physical activity.

XSL•FO
RenderX

Conclusions: The evaluation of the ALIBIRD implementation yielded promising results in that it facilitated the adoption of healthier lifestyle habits while enhancing health self-management among oncology patients. The ALIBIRD mHealth platform emerges as an effective digital health tool that enables closer monitoring of patients and thereby more informed clinical decision-making.

Trial Registration: ClinicalTrials.gov NCT05770869; https://clinicaltrials.gov/study/NCT05770869

(JMIR Cancer 2026;12:e69525) doi:10.2196/69525

KEYWORDS

mHealth; digital health; usability; patient monitoring; PROs; PROMs; lung cancer; precision nutrition; patient-reported outcome; patient-reported outcome measure; mobile health

Introduction

Clinical Context and State of the Art

Cancer represents a significant health problem that, according to statistics from the GLOBOCAN 2022 database, has an incidence of nearly 2.5 million cases and causes 1.8 million deaths worldwide [1]. Among the different types, lung cancer (LC) is the most commonly diagnosed cancer globally (12.4% of total cases) and is also the leading cause of cancer-related deaths (18.7% of total cancer deaths). Effective management of disease and treatment-related symptoms, alongside the provision of supportive care, is essential for improving the quality of life (QoL) of patients with LC. Nevertheless, health care providers frequently underestimate the severity of symptoms during oncology consultations, a phenomenon attributed to demanding schedules, substantial workloads, and patients' hesitance to articulate their concerns [2,3]. Oncologists often prioritize the identification and management of severe adverse events, thereby overlooking the comprehensive spectrum of symptoms that significantly impact patients' QoL [4,5].

Consequently, patients with advanced lung cancer (aLC) experience multiple unmet needs, including complex symptom management, nutritional deterioration, and emotional distress, which are often insufficiently addressed by traditional care models. These models struggle to provide frequent, personalized follow-up, resulting in suboptimal QoL and potentially avoidable clinical events. Recent literature highlights the need for innovative, patient-centered strategies and hybrid care models that combine in-person and remote monitoring to deliver timely, tailored interventions [6,7]. Digital health solutions, particularly mobile health (mHealth) platforms, offer a promising approach by enabling continuous telemonitoring and supporting a multidisciplinary, multidimensional framework that integrates symptoms, lifestyle, nutrigenetics, and microbiota. Within this context, patient-reported outcomes (PROs) and their validated measures (patient-reported outcome measures [PROMs]) have become key instruments for capturing patients' health status [8,9]. Their electronic implementation (electronic PROs) allows patients to self-report symptoms, lifestyle, and treatment-related information, enhancing self-care and adherence while enabling proactive clinical monitoring. This, in turn, contributes to more efficient health care delivery by reducing unplanned visits, emergency room utilization, and hospital admissions [10,11].

Various lifestyle factors, such as a healthy diet and physical activity, significantly influence cancer treatment and prognosis. In aLC, the high prevalence of malnutrition and risk of cachexia underscore the need for nutritional support and lifestyle interventions [12,13]. Personalized nutrition and lifestyle recommendations have emerged as promising strategies, allowing the design of tailored dietary plans based on clinical and genetic profiles [14]. Nutrigenetics can positively impact cancer treatment by genotyping processes such as lactose, gluten, alcohol, and tobacco metabolism, as well as physical activity tolerance and predisposition to conditions such as constipation or insomnia, thereby informing tailored supportive treatments [15]. Furthermore, the intestinal microbiome is implicated in cancer development, progression, and treatment response, with dysbiosis frequently observed in various malignancies [16,17].

eHealth and mHealth interventions have been increasingly implemented to support follow-up and improve QoL in patients with cancer [18,19]. These approaches build upon broader evidence from information and communication technologies in health care, which have been successfully applied across various medical fields. Such interventions have been associated with reductions in unplanned hospital visits and earlier detection of symptom deterioration [20,21], improvements in health-related quality of life and symptom self-management [22], and high patient adherence and acceptability in clinical use [23]. Evidence from oncology-specific applications reinforces this trend: studies have shown that patients are willing to adopt digital tools when recommended by clinicians and demonstrate sustained engagement and high acceptability in clinical use [7,23]. Complementary evidence from recent mHealth evaluations indicates that these interventions are generally well accepted, with high usability and sustained patient engagement in oncology settings [24-26].

Comparative studies further suggest that mHealth-based follow-up can achieve outcomes equivalent or superior to conventional care in aspects such as physical function, symptom control, and patient satisfaction, supporting the development of hybrid care models that combine in-person visits with telematic monitoring [19,27,28]. Moreover, emerging evidence points to the potential cost-effectiveness of these interventions, which may optimize resource allocation in oncology care [29]. These findings reinforce the potential of mHealth to support integrated, multidisciplinary care models capable of addressing the complex needs of patients with advanced disease, ultimately enhancing patient empowerment and enabling better clinical decision-making.

XSL•FO
RenderX

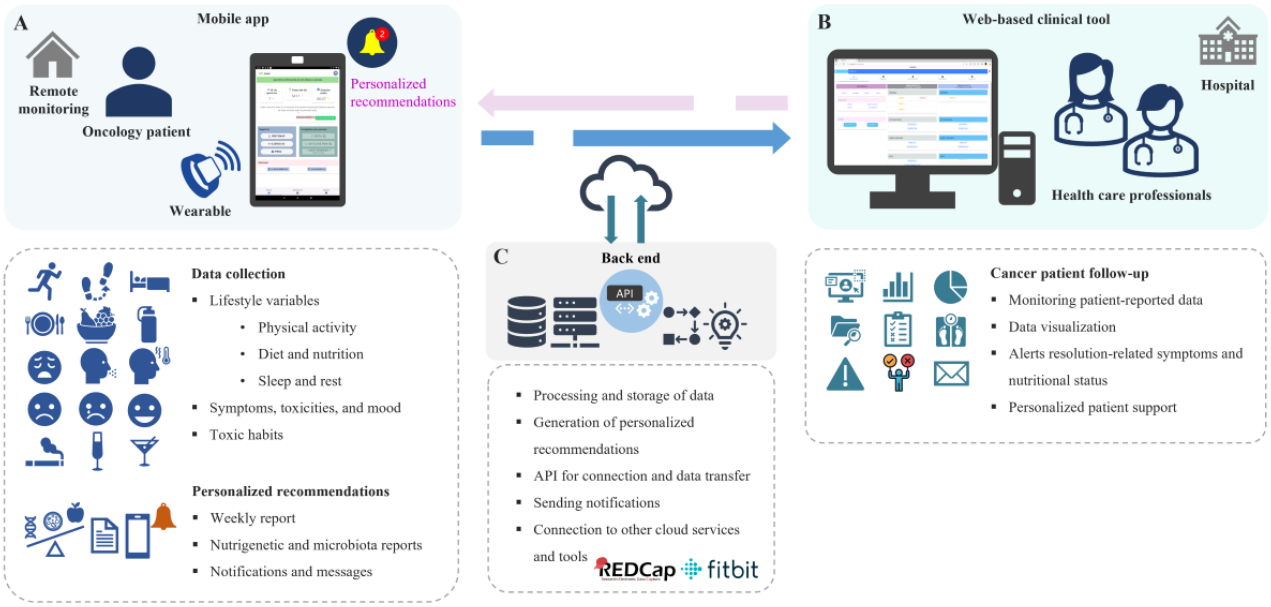
Description of the ALIBIRD mHealth Platform

The ALIBIRD mHealth platform was developed within the ALIBIRD2020-CM program (S2018/BAA-4343)—a multidisciplinary consortium of 5 institutions from the Community of Madrid, Spain—aimed at designing and validating precision nutrition strategies for patients with cancer [30]. The development followed a user-centered and participatory design framework [31,32], actively involving patients, clinicians, and researchers throughout all stages [33,34]. Agile methodologies were applied [35,36], using iterative cycles of design, prototyping, and validation to ensure usability, clinical relevance, and alignment with user needs. The platform was developed and internally validated over approximately 18 months prior to the clinical study, following this iterative, user-centered process, for its subsequent evaluation in a pilot intervention.

The platform comprises three main components: (1) a mobile app for patients, (2) a web-based tool for health care professionals, and (3) a back-end infrastructure that includes the application programming interfaces, databases, notification services, wearable and REDCap (Research Electronic Data Capture) [37] integration, and a rule-based engine.

The mobile app enables patients to report PROs and PROMs (symptoms, diet, and lifestyle) and to synchronize data from Fitbit devices to monitor physical activity and sleep. It provides personalized recommendations, nutrigenetic and gut microbiota reports, weekly progress summaries, and educational resources to promote self-management. The web-based tool allows health care professionals to remotely monitor patients' symptoms, nutritional status, and lifestyle patterns, providing real-time access to alerts and reports. The back-end processes PROs, lifestyle data, and other clinical inputs to generate automated reminders, positive reinforcement messages, weekly summaries, and clinical alerts. It also incorporates baseline nutrigenetic analyses and gut microbiota profiling as structured inputs that support individualized nutritional guidance and lifestyle advice. These analyses are stored within the platform and contribute to tailoring recommendations delivered through the app and to the content available for clinicians during follow-up. Figure 1 illustrates the overall architecture and interactions of the ALIBIRD platform, including the patient mobile app (Figure 1A), the web-based clinical tool for health care professionals (Figure 1B), back-end services, and third-party platforms such as REDCap and Fitbit (Figure 1C).

Figure 1. Overview of the ALIBIRD mobile health (mHealth) platform and the interactions and functionalities of its main components. The diagram illustrates (A) the patient mobile app, which collects lifestyle variables (physical activity, diet and nutrition, sleep and rest, and toxic habits), and patient-reported outcomes, including symptoms, toxicities, and mood, and delivers personalized recommendations (medication reminders, weekly reports, nutrigenetic and gut microbiota–based advice, and mobile notifications and messages); (B) the web-based clinical tool for health care professionals, which supports follow-up of patients with cancer through visualization of patient-reported data, resolution of symptom and nutritional alerts, and personalized patient contact; and (C) the back-end infrastructure, which enables secure processing and storage of data, generation of personalized recommendations and alerts, application programming interfaces for data connection and transfer, automated notifications, and integration with external cloud services and tools (eg, REDCap and Fitbit). API; application programming interface; REDCap: Research Electronic Data Capture.



Additional details of the rule engine and its logic, along with illustrative screenshots of the patient mobile app and the web-based clinical tool, are provided in Multimedia Appendix 1(Figures S1 and S2).

Personalized Recommendations and Alerts

The ALIBIRD platform includes a rule-based algorithm (rule engine) designed to generate automated recommendations and clinical alerts, informed by clinical guidelines, validated instruments, and expert consensus. The engine systematically analyzes PROs, lifestyle data, nutrigenetic information, gut microbiota findings, and other clinical inputs stored in the platform's database. Within this framework, nutrigenetic and microbiota-derived insights are integrated directly into the decision logic to refine nutritional and lifestyle recommendations according to each patient's biological profile.

XSL•FO
RenderX

Thresholds for dietary recommendations and nutritional risk (including weight management and malnutrition indicators) were derived from the Mediterranean Diet Adherence Screener (MEDAS), the Mediterranean Diet Serving Score, and the ESPEN practical guideline: Clinical Nutrition in Cancer [38-42]. Symptom severity was classified according to the Common Terminology Criteria for Adverse Events (CTCAE v5) [43], while recommendations related to physical activity and sleep patterns were aligned with widely accepted international standards, such as those from World Health Organization.

Four types of outputs are generated: (1) reminders to support adherence, (2) positive reinforcement messages to promote healthy habits, (3) weekly progress summaries, and (4) clinical alerts to notify health care professionals about adverse events or critical changes (eg, severe symptoms, reduced food intake, or significant weight loss). When an alert is triggered, clinicians can take timely actions, such as scheduling an appointment, adjusting treatment, or referring the patient to a specialist.

The development of the algorithm was informed by expert input from oncologists, nutritionists, and nursing professionals, with decision rules and thresholds being progressively refined to ensure clinically relevant and accurate outputs. Additional details of the rule engine, including monitored dimensions and a schematic overview of data processing and outputs, are provided in Multimedia Appendix 2 (Table S1 and Figure S1).

The primary objective of this study was to evaluate the usability and patient satisfaction with the mHealth ALIBIRD platform in patients with aLC. The secondary objective was to explore its clinical impact, including early detection of relevant events, resolution of alerts, and the promotion of healthier lifestyle habits.

Methods

Study Design

A single-arm, prospective, interventional study was conducted involving patients with advanced thoracic tumors, mainly non–small cell lung cancer (NSCLC), at the Infanta Sofía University Hospital in Madrid, Spain. The total study duration was 18 months, including a preparatory phase for study setup and patient recruitment, followed by a 30-week intervention period per patient. Patient recruitment occurred between November 2021 and January 2022, with consecutive enrollment as patients met eligibility criteria, resulting in a staggered but overlapping participation across the cohort.

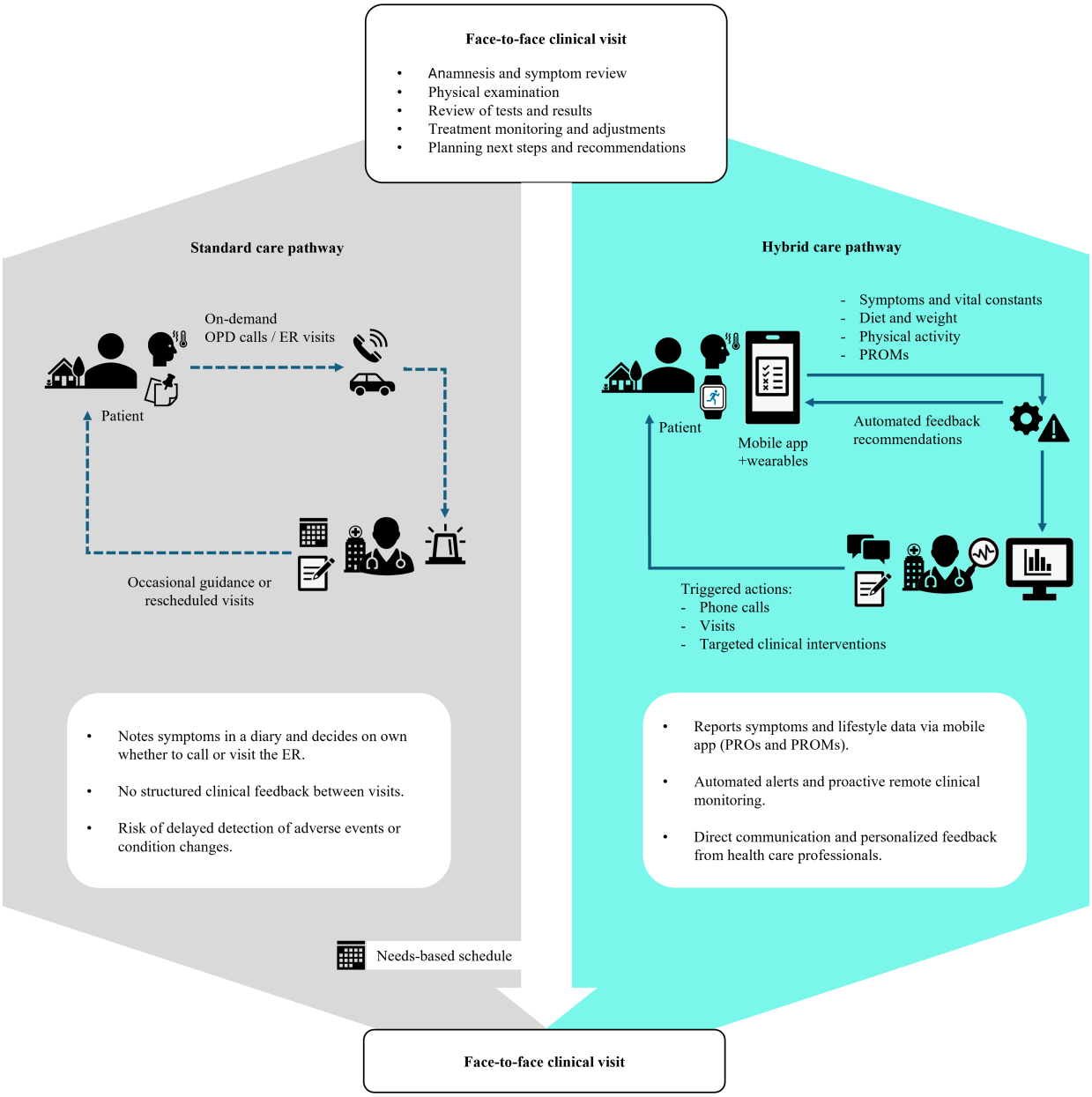
Integration Into Clinical Workflow

In oncology, conventional patient follow-up is typically based on scheduled in-person visits spaced every few weeks, with limited patient-clinician interaction in between. In our study, we implemented a hybrid (mHealth-based) care pathway, integrating the ALIBIRD mobile app and clinical dashboard with usual care to support continuous self-management and remote monitoring.

A multidisciplinary team of 5 clinicians (3 oncologists, 1 nutritionist, and 1 oncology nurse), supported by a data manager and a biomedical engineer, participated in patient monitoring and system oversight. Clinicians performed standard clinical visits (V1-V10) and used the dashboard to review patient-reported data, respond to alerts, and personalize interventions. The data manager conducted regular quality checks on data completeness, while the biomedical engineer ensured technical support and system performance (Figure 2).

XSL•FO
RenderX

Figure 2. Integrating patient-reported outcomes and supportive oncology. The diagram compares the standard care pathway with the hybrid mHealth-supported pathway implemented in the study using the ALIBIRD platform. In the standard pathway, routine visits are scheduled and structured; however, clinical follow-up between visits relies on patient-initiated contact, leading to unstructured feedback and a higher risk of delayed recognition of adverse events. In contrast, the hybrid pathway incorporates continuous digital monitoring through the mobile app and wearable devices, enabling patients to report symptoms, lifestyle data, and patient-reported outcome measures. Automated algorithms generate alerts and personalized recommendations, prompting timely clinical actions such as phone calls, rescheduled visits, or targeted interventions. Continuous lines indicate structured digital follow-up, while dashed lines represent occasional or patient-initiated contact. ER: emergency department; OPD: outpatient department; PROMs: patient-reported outcome measures; PROs: patient-reported outcomes.



Study Protocol

The primary objective was to evaluate the usability and patient satisfaction with the ALIBIRD platform in patients with aLC. The secondary objective was to explore the clinical impact of the system, including early detection of relevant events, resolution of alerts, and the promotion of healthier lifestyle habits. The intervention period lasted 30 weeks per patient and included 10 scheduled face-to-face visits (V1-V10), complemented by continuous remote monitoring through the ALIBIRD mobile app and a wearable device.

Eligible participants were consecutively identified by the study oncologists during routine outpatient visits for NSCLC. Once a patient met the inclusion criteria—histologically or cytologically confirmed NSCLC, aged 18 years or older, Eastern Cooperative Oncology Group performance status ≤2, ongoing systemic treatment, regular internet access, and sufficient digital literacy—and none of the exclusion criteria applied (eg, symptomatic brain metastases, neurological or psychiatric disorders, or inability to comply with study procedures), the oncologist presented the study and its objectives. Each candidate

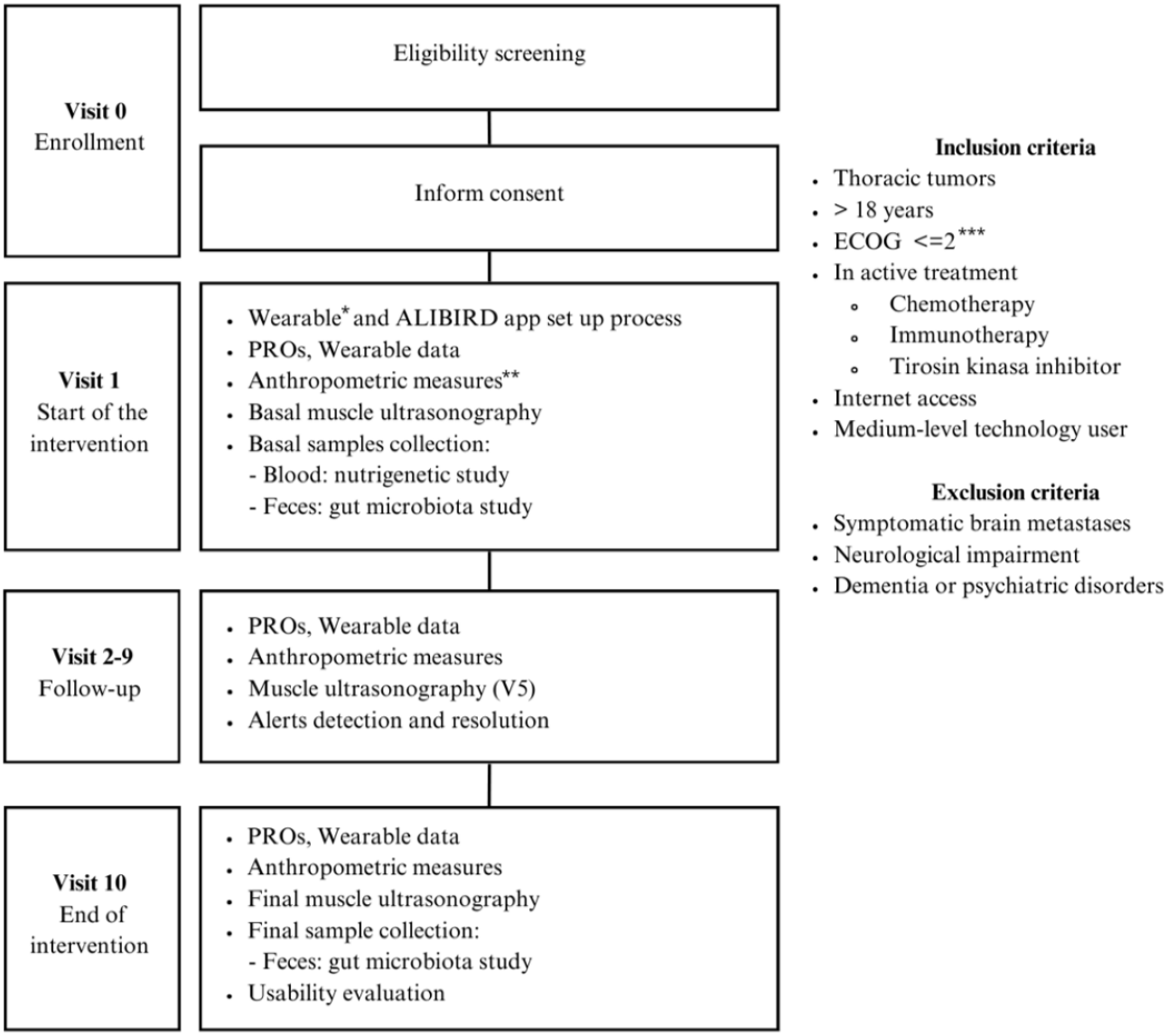
XSL•FO
RenderX

received an information sheet and provided written informed consent prior to enrollment.

Following enrollment, a biomedical engineer conducted individualized training sessions covering the purpose, scope, and functionalities of the ALIBIRD mobile app, as well as the proper use of the wearable device and the type and frequency of data to be recorded. Each participant received a Fitbit Inspire model wearable provided by the study team. Fitbit Inspire is an accelerometer-based device for monitoring daily steps, physical activity patterns, and sleep duration in adult populations. Continuous technical support was available throughout the study to address any issues or questions.

During the study, patients reported PROMs and lifestyle-related variables through the mobile app, while the wearable device automatically captured daily activity and sleep data. At each scheduled visit, oncologists conducted clinical assessments, and the nutritionist performed nutritional evaluations, including anthropometric measurements (weight, height, BMI, and body composition determined by bioimpedance and muscle ultrasonography). A research nurse was responsible for collecting biological samples—blood for nutrigenetic analyses and stool for gut microbiota characterization. A data manager supervised data quality and consistency across all data sources to ensure compliance with the study protocol and overall data integrity. Figure 3 illustrates the study flow, main intervention components, and inclusion and exclusion criteria.

Figure 3. Schematic representation of the study protocol for the ALIBIRD prospective pilot intervention in patients with advanced lung cancer. The diagram summarizes enrollment procedures, inclusion and exclusion criteria, and the schedule of assessments across the 30-week study, including baseline setup of the mobile app and wearable device, patient-reported outcomes, wearable-derived activity and sleep data, anthropometric measurements, muscle ultrasonography examinations, and collection of samples for nutrigenetic and gut microbiota analyses. Follow-up visits (V2-V9) included continuous remote monitoring and alert resolution, while the final visit (**V10**) incorporated end-of-intervention assessments and usability evaluation. *Wearable device model: Fitbit Inspire. **Anthropometric measures include weight, BMI, and body composition. ***Eastern Cooperative Oncology Group is a widely used performance status scale in oncology. ECOG: Eastern Cooperative Oncology Group; PROs: patient-reported outcomes.



All PROMs and patient-reported experience measures (PREMs), including the System Usability Scale (SUS), were implemented in REDCap and administered electronically through the ALIBIRD mobile app. Notifications were automatically sent

XSL•FO
RenderX

via the app to remind participants when questionnaires were due, facilitating timely completion and high response rates. An overview of all data sources and assessment tools used in the study is provided in Table 1.

XSL•FO
RenderX

Table . Instruments and measures used for on-demand, continuous, and scheduled data capture during the 30-week ALIBIRD mHealth intervention in patients with advanced lung cancer. The table summarizes all variables and corresponding measurement instruments, including patient-reported outcomes, patient-reported outcome measures, wearable-derived data, nutritional assessments, sleep and physical activity questionnaires, toxicity reporting tools, and usability and experience measures, along with their specific capture frequency.

Acquisition instruments	Capture frequency
Symptoms	
Symptom diary (type and severity):	On demand
Symptoms were selected based on the most frequent events in aLC[a] and classified into 3 severity levels: G1 (mild), G2 (moderate), and G3 (severe)	
European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire—Lung Cancer 13 [44]:	V1-V10
Standard instrument for measuring the quality of life of patients with lung cancer; covers 13 typical symptoms and side effects from treatment. Higher scores indicate a high level of symptomatology	
Nutritional status	
Anthropometric measures:	V1-V10
Included weight, height, BMI, waist and hip circumference, and bioimpedance analysis	
Malnutrition diagnosis: The Global Leadership Initiative on Malnutrition criteria [45]:	V1 and V10
Guidelines to diagnose malnutrition: at least 1 phenotypic and 1 etiologic criterion required	
Muscle ultrasonography:	V1, V5, and V10
Ultrasonography of rectus femoris for sarcopenia assessment: area, y-axis (relaxation/contraction), x-axis	
Diet	
Mediterranean Diet Adherence Screener [42]:	V1 and V10
14-item questionnaire to assess adherence to the Mediterranean diet	
Dietary intake and appetite log:	Weekly
Ad hoc questionnaire including daily number of meals, self-reported appetite, and food frequency for the last week (cereals, dairy, fruits, vegetables, oils, liquids, meats, fish, processed meats, etc)	
Physical activity	
International Physical Activity Questionnaire short form [46]:	V1-V10
Standardized tool to classify activity levels into mild, moderate, and vigorous	
Fitbit: daily steps and physical activity sessions	Continuous capture
Sleep	
Pittsburgh Sleep Quality Index [47]:	V1, V5, and V10
Self-report questionnaire assessing sleep quality (eg, waking during the night and need for medication) and classifies it into 2 groups: poor or good quality	
Fitbit: Sleep sessions	Continuous capture
Toxicity	
Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events [48]:	V2-V10
System developed by the National Cancer Institute to capture symptomatic adverse events in cancer trials	
Quality of life	

XSL•FO
RenderX

Acquisition instruments	Capture frequency
European Quality of Life-5 Dimensions 5 Levels [49]: *Generic self-assessed health-related quality-of-life measure across 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression); includes a visual analogue scale. Higher index scores indicate worse health*	V1-V10
Emotional status	
Hospital Anxiety and Depression Scale (HADS) [50]: *Self-report rating scale with 14 items on a 4-point Likert scale (0 - 3). It measures anxiety (HADS-A) and depression (HADS-D) through 7 items each, with subscale scores ranging 0 - 21 and a total score summing all items*	V1, V5, and V10
Perceived energy and emotional status: *Two ad hoc self-report questions: (1) overall mood during the past week (good, fair, and poor) and (2) perceived energy level (high, medium, and low)*	Weekly
Usability and satisfaction	
System Usability Scale [51]: *Widely used 10-item questionnaire; provides a 0 - 100 usability score (higher is better)*	V10
Patient-reported experience measures: *3-item measure assessing motivation, satisfaction, and willingness to recommend the app*	V10

[a]aLC: advanced lung cancer.

The sample size of 20 patients was selected for convenience, considering the limited population of patients with NSCLC at our center and the sample sizes reported in similar studies [28,52-56].

Evaluation Methods

Evaluation of the ALIBIRD platform included the analysis of adherence and data completeness, usability and patient experience, and the clinical impact of the intervention. System use and adherence were assessed through the number and consistency of patient-reported entries via the mobile app, as well as data automatically collected by the wearable device. Usability was evaluated using the SUS [51,57], a 10-item questionnaire rated on a 5-point Likert scale, widely used for subjective assessment of system usability. Scores range from 0 to 100, with values above 68 considered above average and indicative of acceptable usability. In addition to patient evaluation, usability was also assessed among health care professionals using the same SUS questionnaire, administered electronically through REDCap via a web-based survey link at the end of the study.

Patient experience was measured through a 3-item PREM, which assessed motivation to adhere to healthy recommendations, overall satisfaction with the app and follow-up, and willingness to recommend the platform. The exact questions are summarized in Textbox 1.

Both the SUS and the PREMs were administered to patients at the end of the intervention (V10).

Clinical impact was explored by analyzing both alerts generated and the corresponding clinical actions, as well as PROs and lifestyle data, including symptoms and toxicities, QoL, diet adherence, physical activity, and sleep parameters.

Textbox 1. Questions about patient experience with the mobile app: 2 of the questions used a 5-point Likert scale, while the third question used a Yes/No format to assess patients' willingness to recommend the app.

- "So far, the app has motivated me to follow guidelines to better control possible symptoms and adopt healthy lifestyle recommendations." (Strongly disagree/Disagree/Neutral/Agree/Strongly agree)
- "How would you rate your level of satisfaction with the app and the follow-up you receive in the study?" (Completely dissatisfied/Dissatisfied/Somewhat satisfied/Satisfied/Completely satisfied)
- "Would you recommend an app with similar features and follow-up to someone you know who is in your situation?" (Yes/No)

Nutrigenetics and Gut Microbiota

Saliva samples were collected at baseline (V1) for nutrigenetic analysis, including genes related to metabolism, nutritional traits, and digestive predispositions. Results were based on single nucleotide polymorphisms (SNPs) and composite scores derived from multiple SNPs, with each SNP weighted according to the strength of scientific evidence (high=1, medium=0.5, and low=0.25). Favorable, intermediate, or unfavorable genotypes contributed to an overall score.

XSL•FO
RenderX

Fecal samples were collected at baseline (V1) and at the end of the intervention (V10) to assess gut microbiota composition and dysbiosis status. Key parameters were derived from the relative abundances of microbial groups associated with positive or negative health effects and the production of beneficial metabolites, such as short-chain fatty acids (Table S1 in Multimedia Appendix 3) [58-63].

The results of both analyses were used to generate personalized reports for patients and health care professionals, which were translated into tailored nutritional recommendations delivered through the ALIBIRD mobile app. Detailed protocols for DNA extraction, genotyping, and SNP scoring (nutrigenetics), as well as 16S rRNA sequencing, taxonomic assignment, and dysbiosis assessment (gut microbiota), are provided in Multimedia Appendix 3.

Data Analysis

Data were analyzed using descriptive statistics (frequencies, proportions, and means, SD), Shapiro-Wilk normality testing, and paired sample tests (*t* test or Wilcoxon) or McNemar test for categorical comparisons (V1 vs V10). The SUS score was calculated following the standard procedure (sum of item scores multiplied by 2.5 to obtain a 0‐100 scale), and PREM responses were summarized as frequencies and percentages. Data were exported from the system database and REDCap, preprocessed in Excel, and analyzed with IBM SPSS Statistics (version 29; IBM Corp).

Ethical Considerations

The study was conducted in accordance with the principles of the Declaration of Helsinki (Fortaleza, Brazil, 2013, and subsequent revisions) and the International Council for Harmonisation Good Clinical Practice guidelines. The protocol and all related documentation were reviewed and approved by the Research Ethics Committee for Medicines of the Hospital Universitario La Paz, Madrid, Spain (approval code: HULP PI-4735). The study also complied with the Spanish Royal Decree 957/2020 of November 3, regulating observational studies with human-use medicines, and with the applicable regional regulations.

All participants received verbal and written information about the study's objectives, procedures, and potential risks and benefits. Written informed consent was obtained from each participant prior to inclusion, in compliance with the ethical principles of autonomy and informed decision-making. Participants were informed of their right to withdraw at any time without providing justification and without any impact on their medical care.

The protection, processing, and analysis of study data were carried out in accordance with the General Data Protection Regulation (EU) 2016/679 and the Spanish Organic Law 3/2018 of December 5 on the Protection of Personal Data and Guarantee of Digital Rights, as well as Law 41/2002 of November 14 on Patient Autonomy and Rights and Obligations Regarding Clinical Information and Documentation. All data were pseudonymized and stored within a secure research data infrastructure encompassing both the REDCap and the ALIBIRD platform databases. These systems were hosted on secure servers and included encrypted data transfer, user authentication, and activity logging. Access to identifiable data was restricted to authorized study personnel only. No financial or material compensation was provided to participants for their participation in this study.

Results

Principal Results

The analysis focused on patient adherence, usability, and clinical impact of the ALIBIRD mHealth platform. Results are organized to highlight patient-reported events and the corresponding clinical actions, as well as lifestyle and PROs in participants who completed the intervention, with exploratory nutrigenomic and microbiota data.

Baseline Participant Characteristics

Of the 20 patients with aLC enrolled, 14 completed the full 30-week intervention, which included attending all scheduled visits (V1-V10) and consistent submission of self-reported data through the app. Four patients did not complete the study due to death and 2 due to clinical and emotional deterioration. Baseline characteristics for all initially enrolled participants are summarized in Table 2; participants who discontinued the study were censored for longitudinal analyses.

XSL•FO
RenderX

Table . Baseline characteristics of participants enrolled in the ALIBIRD prospective pilot intervention for advanced lung cancer (N=20)[a].

Characteristic	Values
Age (years), mean (SD)	
43-77	61.05 (9.70)
Sex, n (%)	
Female	12 (60)
Male	8 (40)
User technological level, n (%)	
Low (caregiver support)	6 (30)
Medium	9 (45)
High	5 (25)
Tobacco habit, n (%)	
Former smoker	10 (50)
No (never)	9 (45)
Yes	1 (5)
Tumor location, n (%)	
Lung	17 (85)
Pleura	1 (5)
Thymus	2 (10)
Stage of cancer, n (%)	
IIIB/C	2 (10)
IV	18 (90)
Eastern Cooperative Oncology Group[b], n (%)	
0	11 (55)
1	8 (40)
2	1 (5)
Systemic treatment, n (%)	
CT[c]	3 (15)
IT[d]	1 (5)
CT-IT	7 (37)
TKI[e]	8 (40)
Radical therapy, n (%)	
CT-RT[f]	1 (5)

[a]Data include sociodemographic, clinical, and treatment-related variables. Continuous variables are expressed as mean (SD) and categorical variables as n (%).

[b]A performance status scale used in oncology.

[c]CT: chemotherapy.

[d]IT: immunotherapy.

[e]TKI: tyrosine kinase inhibitors.

[f]RT: radiotherapy.

Usage Metrics, Data Completeness, and Usability Assessment

This section presents a comprehensive analysis of system usage, data collection patterns, and perceived usability, integrating insights from both patients and health care professionals. Table 3 summarizes the number of data entries provided by participants throughout the study period, offering insights into patient compliance and engagement levels.

XSL•FO
RenderX

Table . Summary of patient-generated data entries and wearable-derived records during the ALIBIRD mHealth intervention (N=20). The table reports the total number of automatic wearable recordings (physical activity and sleep), on-demand symptom and weight entries, and scheduled questionnaire submissions collected across all participants.

Data entries	Total records, n
Automated records[a]	
Physical activity	3589
Sleep	3468
On-demand manual records[b]	
Symptoms	679
Weight	123
Manual scheduled records[c]	
Dietary intake and appetite log	487
European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire—Lung Cancer 13	170
Mediterranean Diet Adherence Screener	34
International Physical Activity Questionnaire short form	159
Pittsburgh Sleep Quality Index	53
Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events	145
European Quality of Life-5 Dimensions 5 Levels	169
Hospital Anxiety and Depression Scale	51

[a]Automated records: entries automatically captured by wearable devices, such as physical activity and sleep data.

[b]On-demand manual records: entries made by patients whenever they experience symptoms or wish to log specific data, such as weight.

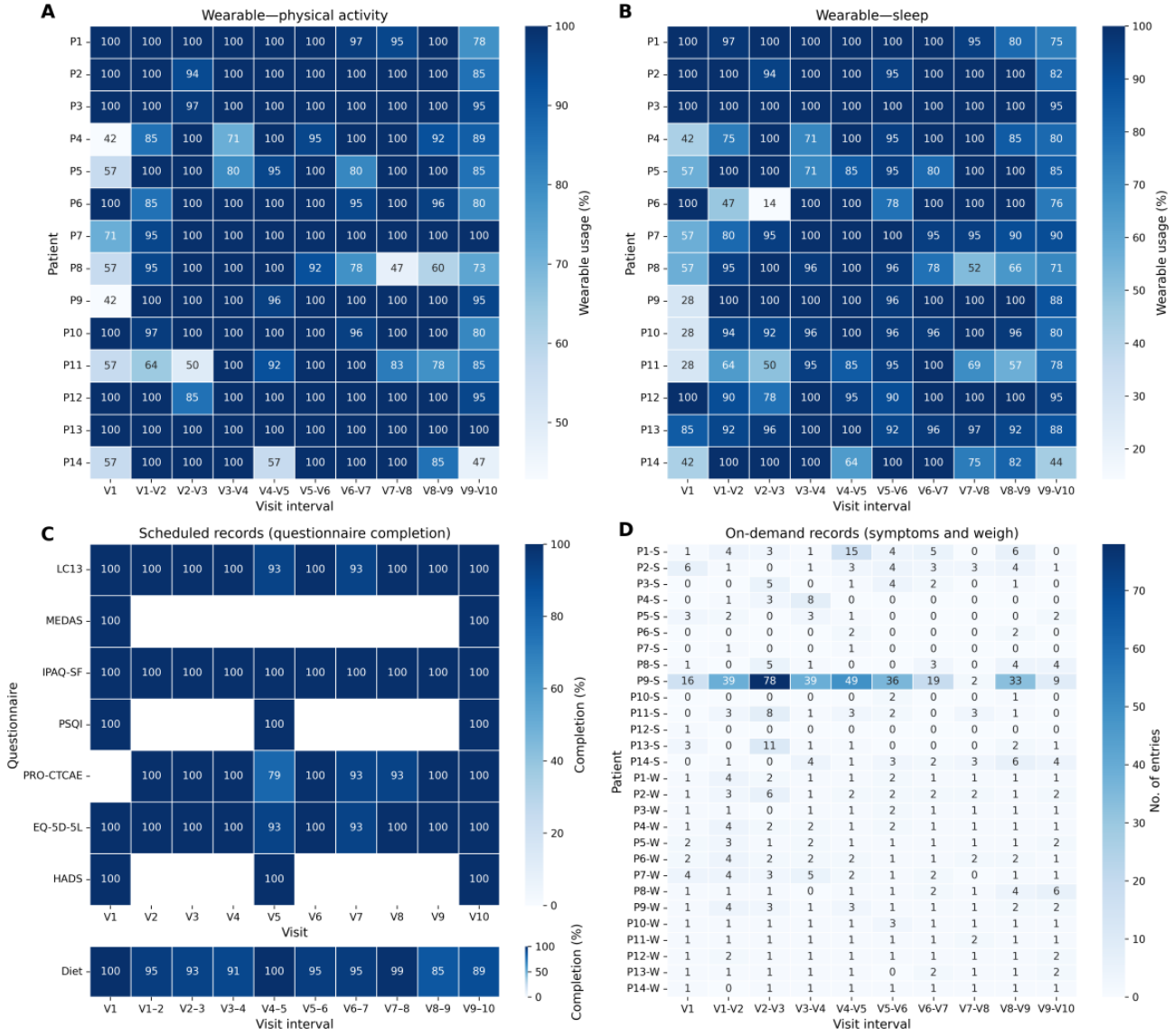
[c]Scheduled records: entries that patients complete at predefined intervals according to the study schedule, typically in the form of questionnaires.

Across the 14 participants who completed the intervention, adherence to scheduled questionnaires remained high throughout the 30-week follow-up. Completion rates ranged from 94% for dietary records to 100% for several instruments, with visit-level averages between 95% and 100%. Wearable-based passive monitoring showed adherence between 66% and 96% across visits, with sleep tracking representing the lower range and physical activity data showing slightly higher and more stable capture. Symptom reporting was heterogeneous, with peaks during the early-mid follow-up and variable contributions among patients, while weight entries were more evenly distributed, with at least 1 measurement recorded by most participants between consecutive visits. Figure 4 illustrates individual-level data collection across visits, including wearable-derived physical activity (Figure 4A), wearable-derived sleep data (Figure 4B), scheduled questionnaire completion (Figure 4C), and on-demand records for symptoms and weight (Figure 4D). This visualization highlights variations in engagement over time and among participants, helping to identify adherence patterns and potential gaps in data reporting.

In addition to evaluating what data were collected, it is also important to understand how patients interacted with the application interface. App usage data revealed that, on average, participants spent the largest proportion of their app interactions on the home screen (33%), followed by the weekly tasks (19%—which include actions related to diet and other lifestyle variables such as sleep, mood, energy, and gastrointestinal rhythm), symptom reporting (15%), and the notifications or message box (10%—which include automated notifications and clinic-patient exchanges). These percentages represent the mean proportion of interactions per patient, calculated relative to each individual's total app interactions and then averaged across participants.

XSL•FO
RenderX

Figure 4. Individual-level data collection across visits by data type and recording modality (N=14). (**A**) Wearable—physical activity, (**B**) wearable—sleep, (**C**) scheduled records (questionnaire completion), and (**D**) on-demand records (symptoms and weight). Manual scheduled records include questionnaires with different expected schedules: QLQ-LC13, MEDAS, IPAQ-SF, and EQ-5D-5L were expected at each visit (V1-V10); PSQI and HADS were scheduled only at visits V1, V5, and V10; and PRO-CTCAE was expected from visit 2 to 10 (V2-V10). The dietary log was completed on a weekly basis, and adherence is represented using the intervals between visits. Percentages of adherence for dietary log and wearable monitoring are expressed relative to each patient's expected entries, as visit intervals and monitoring periods varied across participants. EQ-5D-5L: European Quality of Life-5 Dimensions 5 Levels; HADS: Hospital Anxiety and Depression Scale; IPAQ-SF: International Physical Activity Questionnaire short form; LC13: Lung Cancer 13; MEDAS: Mediterranean Diet Adherence Screener; PRO-CTCAE: Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events; PSQI: Pittsburgh Sleep Quality Index.



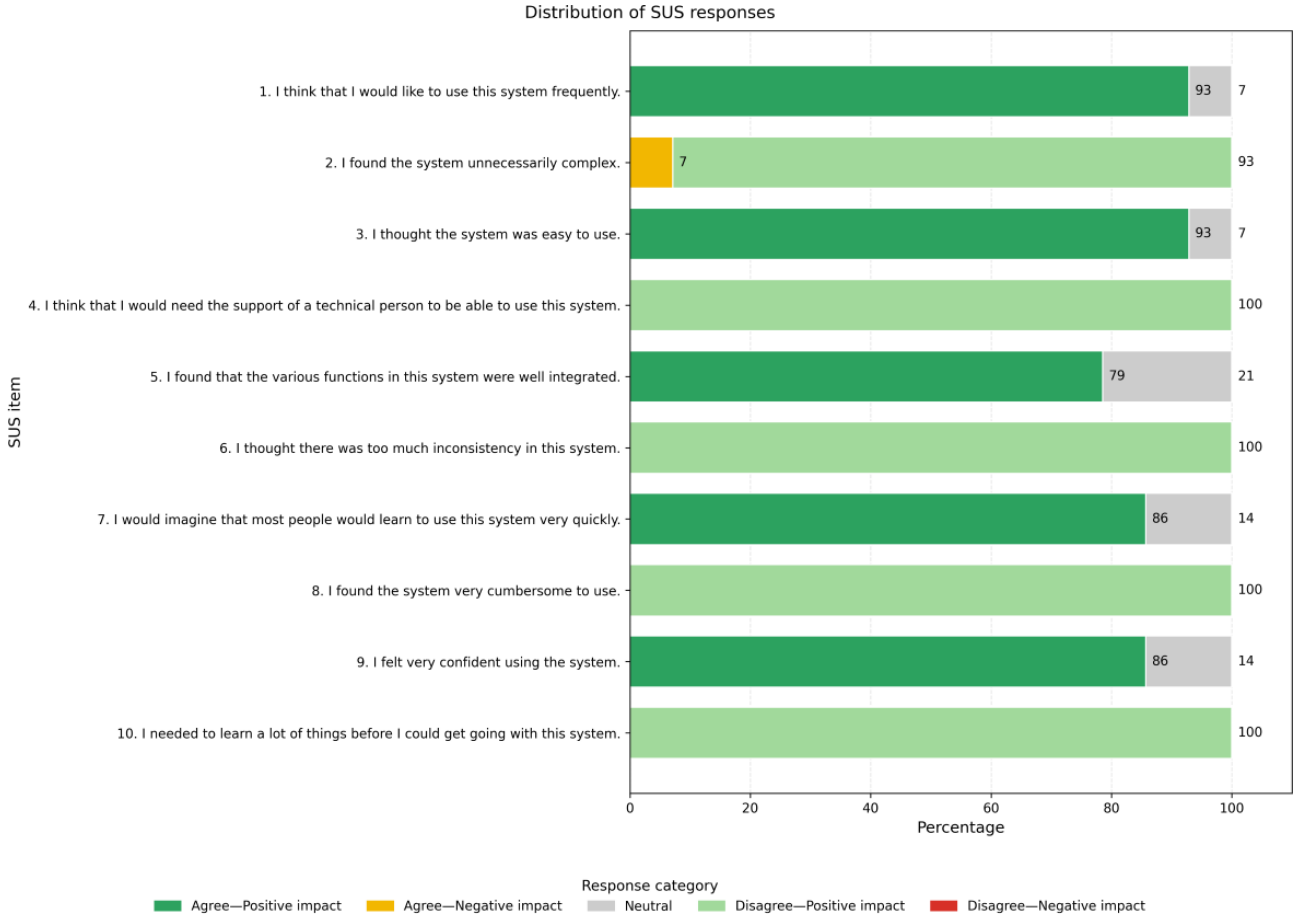
The average SUS score among patients who completed the study (n=14) was 90, indicating a high level of usability for the mobile app. This value is well above the commonly accepted benchmark of 68, which is considered the threshold for above-average systems. Figure 5 displays the distribution of responses for each item of the SUS questionnaire, offering a visual overview of user sentiment across all items. The consistently positive responses suggest a strong perception of ease of use and user satisfaction among participants.

PREMs provided valuable insights into the participants' perception of the ALIBIRD platform. Regarding the question "So far, the app has motivated me to follow guidelines to better control possible symptoms and adopt healthy lifestyle recommendations," the majority of participants responded positively, with 9 out of 14 (64%) indicating "Strongly agree" and 3 out of 14 (21%) indicating "Agree," while 2 out of 14 (14%) remained neutral. For the question "How would you rate your level of satisfaction with the app and the follow-up you receive in the study?," 50% of the participants reported being "Completely satisfied" and the other 7 out of 14 (50%) being "Satisfied." Finally, in response to "Would you recommend an app with similar features and follow-up to someone you know who is in your situation?," all participants (14/14, 100%) responded "Yes."

In addition, usability was also assessed among health care professionals using the clinical web-based interface. The mean SUS score among professionals (n=5) was 86, reflecting a positive perception of the platform's usability in clinical practice. The evaluation involved 5 members of the clinical team, including oncologists, a nutritionist, and a nurse.

XSL•FO
RenderX

Figure 5. Distribution of System Usability Scale responses categorized by agreement and their impact on usability (N=14). The bar chart shows the distribution of responses for each of the 10 System Usability Scale items. Items 1, 3, 5, 7, and 9 are positively phrased, where Agree/Strongly agree responses indicate positive usability and Disagree/Strongly disagree responses indicate negative usability. Items 2, 4, 6, 8, and 10 are negatively phrased, where Disagree/Strongly disagree responses indicate positive usability and Agree/Strongly agree responses indicate negative usability. Neutral responses correspond to the midpoint of the Likert scale (Neither agree nor disagree). The "Disagree—Negative Impact" category is included for completeness, although no responses were observed in this category. SUS: System Usability Scale.



Evaluation of Outcome Data

In this section, we analyze the PRO and PROM results to assess the overall clinical impact of ALIBIRD mHealth platform, focusing on both clinically actionable events and patient-reported measures collected during follow-up. Outcomes include the detection and management of symptom- and nutrition-related alerts that prompted clinical actions, as well as longitudinal changes in lifestyle behaviors, symptoms, and quality-of-life indicators reported by patients. Together, these results provide an integrated view of the platform's clinical impact and its role in supporting patient monitoring and care during the intervention.

Clinical Actions Triggered by Reported Events

Clinical actions were driven by events captured and processed via the ALIBIRD mHealth platform, including patient-reported symptoms and nutritional indicators, which triggered automated alerts and subsequent interventions by the health care team. During the study, the most frequently reported symptoms among patients (n=20) were asthenia (164/679, 24.2%) and dyspnea (132/679, 19.4%). Overall, most symptom reports were classified as mild (grade 1: 367/679, 54.0%), followed by moderate (grade 2: 238/679, 35.1%) and severe (grade 3: 74/679, 10.9%).

These patient-reported symptoms, combined with nutritional monitoring, triggered the majority of the 2100 automated recommendations and 143 clinical alerts generated by the platform for the health care team. Symptom-related alerts (75/143, 52.4%) were most often due to constipation (28/75, 37%), pain (19/75, 25%), and dyspnea (9/75, 12%), followed by fever (9/75, 12%) and less frequent events such as hemoptysis (4/75, 5%), diarrhea, drowsiness, and vomiting, each (2/75, 3%). Nutritional alerts represented 47.6% (68/143), mainly related to reduced food intake (52/68, 76%), fewer than 3 meals per day (9/68, 13%), and weight loss greater than 5% (7/68, 10%).

Health care professionals managed these alerts using the ALIBIRD web-based tool to resolve alerts, which led to 42 phone calls and 20 in-app messages to patients. The mean average resolution time for an alert was 2.05 (1.72 SD) days. Clinical responses included 52 therapeutic adjustments, 10 expedited medical visits, 15 referrals to other specialties, 2 emergency referrals, and 4 imaging reevaluations, with 2 cases confirming clinical-radiological progression.

Lifestyle and PROs

Among patients who completed the intervention, several lifestyle and patient-reported outcome measures were monitored over

XSL•FO
RenderX

time. Adherence to the Mediterranean diet (Mediterranean Diet Adherence Screener) increased from 50% (7/14) at baseline (V1) to 71% (10/14) by the end of the intervention (V10), while BMI remained relatively stable across all visits, with no clinically meaningful changes.

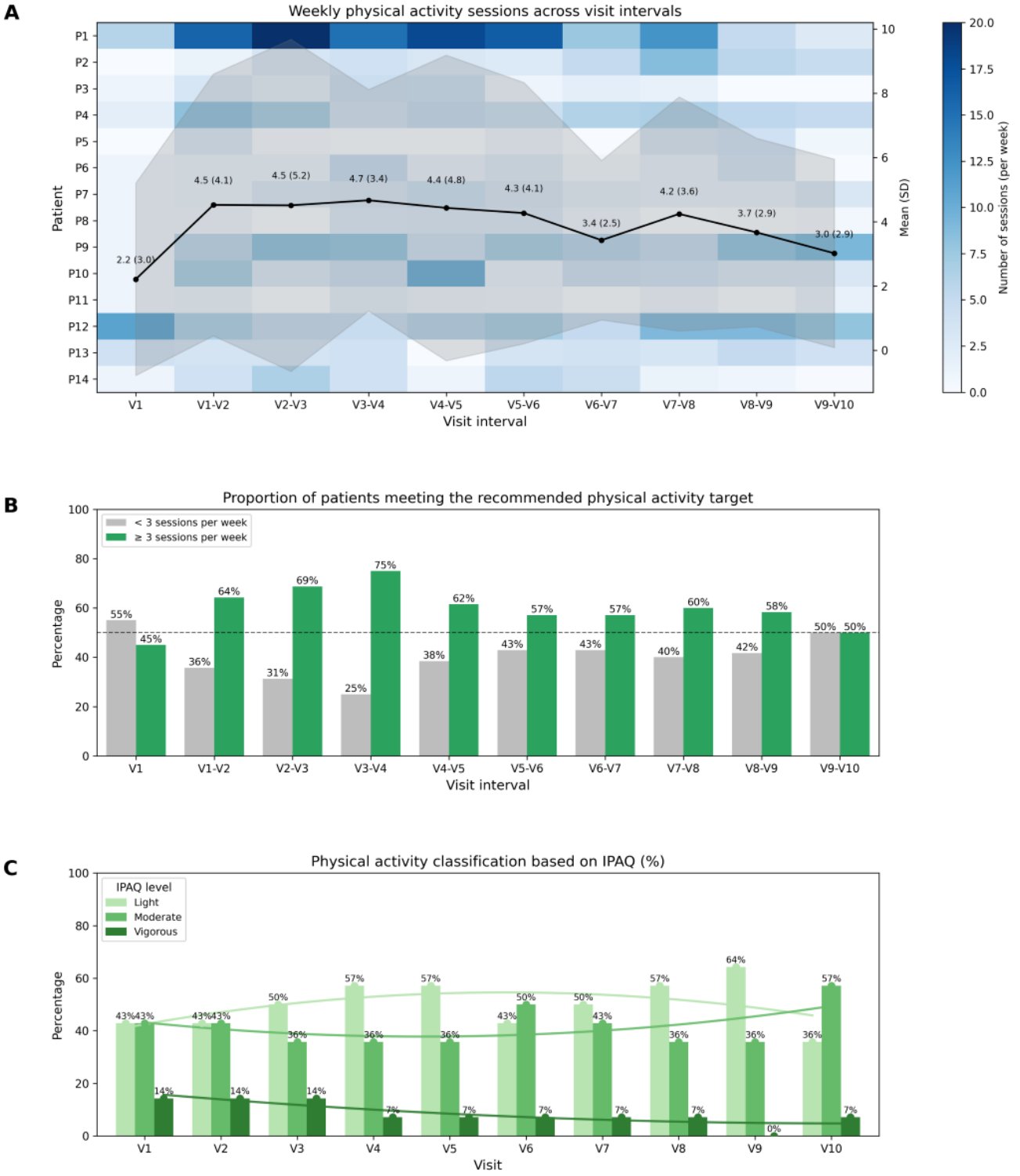
Physical activity, monitored via the ALIBIRD mobile app and Fitbit devices, showed that participants performed a mean average of 3.90 (SD 3.68) sessions per week and 6087 daily steps (SD 2675). Figure 6 illustrates the evolution of these metrics across the study period, including the number of weekly physical activity sessions, adherence to the recommended target of 3 or more sessions per week, and the distribution of activity intensity levels (light, moderate, and vigorous) according to the International Physical Activity Questionnaire short form questionnaire. Figure 6 summarizes physical activity outcomes across the study period, including weekly exercise frequency (Figure 6A), the proportion of patients meeting the recommended target of ≥3 sessions per week (Figure 6B), and activity intensity levels based on the International Physical Activity Questionnaire short form results (Figure 6C). Step count data collected through Fitbit indicated predominantly light levels of physical activity, with relatively stable averages across visits.

Sleep monitoring revealed an average duration of 7.61 (SD 1.34) hours per day according to Fitbit data, consistent with recommended sleep duration. However, Pittsburgh Sleep Quality Index results indicated persistently poor subjective sleep quality among most participants. The proportion of patients classified as having poor sleep quality was 71.4% (10/14) at baseline, slightly improved at midintervention (V5) to 64.3% (9/14), and worsened again by the end of the study (V10) to 78.6% (11/14).

Regarding emotional health, Hospital Anxiety and Depression Scale—A indicated a slight increase in clinically relevant anxiety from 0% (0/14) to 7% (1/14), while Hospital Anxiety and Depression Scale—D revealed that 14% (2/14) of patients developed depressive symptoms by V10. In relation to patient QoL and specific symptoms, the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire—Lung Cancer 13 questionnaire generated a total of 655 symptom reports throughout the study. The most frequently reported symptoms were total pain (242/655, 36.9% of reports), cough (115/655, 17.6%), and total dyspnea (113/655, 17.3%). In addition, the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events questionnaire, which captured 5025 patient-reported toxicity items over the 10 study visits, highlighted the most frequent symptom groups as follows: sleep and wake patterns (962/5025, 19.1% of reports), pain (847/5025, 16.9%), and neurological and perception disorders (774/5025, 15.4%). QoL, assessed by the European Quality of Life-5 Dimensions 5 Levels questionnaire, remained stable throughout the study. The mean score at baseline was 0.86 (SD 0.10) and showed only minor variations across visits, with a slight increase at midintervention (V4: mean 0.92, SD 0.09) and a final mean score of 0.88 (SD 0.15) at V10. Overall, participants maintained good perceived health status during the intervention, with no clinically relevant decline or improvement observed. Table 4 provides a summary of the pre- and postintervention comparison between baseline (V1) and the end of the study (V10) for participants who completed the intervention.

XSL•FO
RenderX

Figure 6. Physical activity outcomes across the study period (N=14). (**A**) Heatmap showing the average number of weekly physical activity sessions recorded via Fitbit across visit intervals, with overlaid mean (SD). (**B**) Percentage of patients who met the recommended physical activity target of ≥3 sessions per week compared with those who did not, aggregated by visit interval. (**C**) Distribution of physical activity intensity levels (light, moderate, and vigorous) based on International Physical Activity Questionnaire short form responses across study visits. IPAQ: International Physical Activity Questionnaire.



XSL•FO
RenderX

Table . Pre- and postintervention comparison of dietary, nutritional, physical activity, and sleep parameters between baseline (V1) and end of study (V10) among participants who completed the 30-week ALIBIRD mHealth intervention (N=14). Statistical significance was assessed using paired tests or McNemar test where applicable[a].

Variables	Preintervention V1	Postintervention V10	*P* value
Diet			
Mediterranean Diet Adherence Screener, % of adherence (n)	50 (7/14)	71 (10/14)	.12
Food intake reduction, percentage of patients (n)			
No	43 (6/14)	43 (6/14)	—[b]
Mild-moderate	57 (8/14)	50 (7/14)	—[b]
Severe	0 (0/14)	7 (1/14)	—[b]
Dietary guideline compliance,[c] % of patients, (n)			
Meat, eggs, fish, legumes, and dairy	21 (3/14)	57 (8/14)	.06
Fruits and vegetables	0 (0/14)	57 (8/14)	—[b]
Liquids	0 (0/14)	100 (14/14)	—[b]
Cold meat	43 (6/14)	71 (10/14)	.22
Pastries	21 (3/14)	50 (7/14)	.13
Sugary drinks and sweets	79 (11/14)	71 (10/14)	1
Alcohol	64 (9/14)	71 (10/14)	1
Nutritional status			
Global Leadership Initiative on Malnutrition Criteria, % of patients (n)			
Normonutrition	79 (11/14)	79 (11/14)	.26
Moderate malnutrition	14 (2/14)	36 (5/14)	
Severe malnutrition	7 (1/14)	14 (2/14)	
BMI, mean (SD)	26.14 (2.93)	26.13 (3.08)	.99
Muscle ultrasonography, mean (SD)			
Area, cm^2	4.43 (1.40)	4.50 (0.91)	.80
y-axis in relaxation, cm	1.41 (0.39)	1.43 (0.21)	.81
y-axis in contraction, cm	1.57 (0.48)	1.52 (0.42)	.77
x-axis, cm	3.92 (0.72)	3.76 (0.50)	.44
Bioimpedance (ohms), mean (SD)	526.93 (7.68)	508.86 (79.9)	.42
Physical activity, mean (SD)			
Weekly sessions	2.21 (2.88)	3.02 (2.82)	.34
Daily steps	3945 (2015)	4975 (2409)	.19
International Physical Activity Questionnaire short form, % of patients (n)			
Walking	43 (6/14)	36 (5/14)	.53
Moderate	43 (6/14)	57 (8/14)	
Vigorous	14 (2/14)	7 (1/14)	

XSL•FO
RenderX

Variables	Preintervention	Postintervention	*P* value
	V1	V10	
Sleep			
Duration (hours)	6.93 (1.62)	7.07 (1.18)	.75
Pittsburgh Sleep Quality Index, % of patients (n)			
Poor quality	71 (10/14)	79 (11/14)	.53
Good quality	29 (4/14)	21 (3/14)	

[a]Percentages refer to patients; absolute numbers are indicated within parentheses (n/N).

[b]The *P* value could not be calculated. In some analyses, the McNemar test could not be performed due to the lack of variability in the data. Specifically, when one of the categorical variables involved is constant, meaning all responses fall within a single category, statistical tests that rely on variation between groups, such as McNemar test, are unable to calculate an associated *P* value or significance measure.

[c]Dietary recommendations: meat + eggs: minimum 4-6 servings per week; fish + shellfish: minimum 3-4 servings per week; legumes: minimum 2-4 servings per week; dairy products: minimum 2 servings per day; fruit + vegetables: minimum 5 servings per day; liquids (water, sugar-free isotonic drinks, infusions, coffee, juices, etc): minimum 4 servings (glasses) or 1 liter per day; processed meats: maximum 1-2 servings per week; pastries: maximum 1-2 servings per week; sugary drinks and sweets: maximum 1-2 servings per week; and alcohol: do not consume.

Exploratory Nutrigenetic and Gut Microbiota Profiling

As components of the personalized intervention, baseline nutrigenetic analysis and gut microbiota profiling were incorporated into the platform to inform individualized nutritional recommendations. In this study, these measures were not evaluated as clinical outcomes; rather, descriptive summaries of the baseline profiles (and V10 profiles for microbiota) are presented to contextualize patient heterogeneity.

Genetic analysis was performed on all initial participants (n=20), targeting 20 variants associated with lifestyle habits, nutrition, and cancer. The analysis revealed that 60% of the population carried a CLOCK gene variant associated with protection against sleepiness, favorable breakfast timing, and enhanced intestinal motility, which may support appetite and dietary intake during cancer treatment. Most participants did not present metabolic alterations for gluten sensitivity (16/20, 80%) or lactose intolerance (15/20, 75%), although 5 patients exhibited lack of lactase persistence and were recommended a lactose-free diet. Only 25% (5/20) of the sample displayed a favorable genetic profile for systemic inflammation, which may contribute to better clinical outcomes and reduce circulating proinflammatory factors. Detailed results of the nutrigenetic analysis are provided in Table S1 in Multimedia Appendix 4.

Gut microbiota analyses were conducted in the subset of participants who completed the intervention and had valid samples available at both baseline and end of intervention (n=13). At baseline (V1), 38.5% (5/13) of the participants exhibited a gut microbiota profile similar to the healthy population (dysbiosis grade <30%), whereas 61.5% (8/13) showed a clear state of dysbiosis (dysbiosis grade >30%). After completing the intervention (V10), 15.4% (2/13) of the participants showed improvement in dysbiosis grade, 15.3% (2/13) maintained their baseline status, 23.1% (3/13) experienced a slight deterioration, and 46.2% (6/13) presented a marked deterioration.

Discussion

Principal Findings

This pilot study demonstrates the feasibility, usability, and acceptability of the ALIBIRD mHealth platform as a digital health solution to support the care of patients with aLC. The ALIBIRD apps achieved a mean SUS score of 90 of 100 among patients and 86 of 100 among health care professionals, reflecting excellent usability across both user groups. These results were complemented by the PREMs, which showed that 100% of the participants were satisfied and willing to recommend the platform, and 85% of the participants expressed high motivation to follow the recommendations received. Together with the high completion rates of scheduled questionnaires (94% - 100%) and sustained wearable adherence (81% - 96% after the initial week), these findings indicate strong engagement and adherence over the 30-week follow-up.

Patients actively used the mobile app and wearable device to report symptoms, record lifestyle data, and track physical activity and sleep, while clinicians remotely reviewed this information to guide interventions. A total of 679 symptoms and 143 clinical alerts were processed, with a mean resolution time of 2.05 days. Notably, 2 alerts led to early identification of unsuspected disease progression, underscoring the potential value of continuous digital monitoring in supporting timely clinical decision-making. Importantly, lifestyle-related indicators showed favorable trends during the intervention: adherence to the Mediterranean diet increased from 50% to 71%, physical activity levels improved modestly, and sleep duration averaged 7.6 hours per night. Although subjective sleep quality (Pittsburgh Sleep Quality Index) remained poor for most participants, these findings highlight the potential of the platform to promote awareness and gradual behavioral changes in diet and activity.

QoL, as measured by European Quality of Life-5 Dimensions 5 Levels, remained stable throughout the study, suggesting that continuous monitoring and early interventions may help preserve well-being in this patient population. This stability is particularly

XSL•FO
RenderX

relevant, given the typically unfavorable clinical course of aLC. Taken together, these results confirm the feasibility and usability of ALIBIRD for sustained multimodal engagement in long-term digital oncology interventions.

Comparison With Prior Work

The introduction of PROMs into routine clinical practice has become increasingly important in cancer care, as these tools provide a structured way to assess patients' health status and monitor disease evolution. However, their routine implementation still faces practical barriers, including workload burden, the need for training and support, associated costs, and reluctance among some clinicians. The digital collection of PROMs (ePROMs) has emerged as a strategy to overcome many of these limitations by enabling more efficient patient monitoring and supporting timely clinical decision-making. Evidence from previous trials also indicates that ePROM-based follow-up can contribute to improved clinical outcomes and has the potential to optimize health care resource use [10,11]. In this context, digital health platforms such as ALIBIRD can help overcome routine implementation barriers by facilitating automated data collection and generating clinically actionable information.

The usability, adherence, and engagement results observed in this study compare favorably with previous long-term mHealth interventions in oncology and chronic disease management [25,64]. Anders et al [24] described SUS scores of 85 - 89 in patients with breast cancer, Kessel et al [26] reported 83% willingness to continue app use, and Wu et al [25] documented 61% - 77% daily engagement in long-term self-management apps. ALIBIRD's higher usability scores and sustained adherence likely reflect its integrative design, combining active self-reporting, passive monitoring, personalized recommendations, and timely clinician-patient communication.

In addition, the high levels of satisfaction and motivation reflected by the PREMs align with prior findings that emphasize the role of positive user experience in sustaining engagement and promoting self-care behaviors in digital health interventions [65,66]. Our results also support evidence that digital health tools can enhance self-management, enable early symptom detection, and optimize patient-centered care [25,28].

In this study, frequent alerts related to constipation, pain, dyspnea, and nutritional decline were effectively addressed through remote interactions, demonstrating how digital platforms can extend clinical follow-up beyond hospital visits. The improvements in dietary adherence and weight stabilization observed among participants are particularly relevant, as they suggest that app-based guidance can mitigate the nutritional challenges and risk of malnutrition typically observed in patients with aLC [67-69]. This may help explain the stable quality-of-life scores observed in our cohort, contrasting with the decline usually associated with disease progression.

A distinctive aspect of ALIBIRD is its integration of precision-nutrition elements, including nutrigenetic and microbiota analyses. This approach aligns with emerging evidence on the role of gut health and genetic predispositions in modulating treatment response and nutritional outcomes in cancer [70-77]. The prevalence of dysbiosis and inflammation-related variants in this cohort reinforces the importance of combining molecular data with behavioral and lifestyle tracking to support personalized, adaptive interventions.

Strengths and Limitations

Implementing PRO-based monitoring between clinic visits offers clear advantages in terms of accessibility and closer patient follow-up. At the same time, it also introduces several operational challenges, such as ensuring reliable and robust platform performance, maintaining completion adherence, and providing clinicians with timely information to support clinical decision-making. Within this context, this study provides early evidence supporting the feasibility of deploying an integrated, multidomain digital health platform in oncology care for patients with aLC. The combined monitoring and response to clinical symptoms, lifestyle behaviors, and personalized nutrition represent a holistic approach rarely evaluated in real-world settings. Another strength is the dual usability assessment among patients and health care professionals, which offers insight into cross-stakeholder acceptance—an essential component for successful digital implementation in clinical workflows.

However, several limitations should be acknowledged. The small sample size (n=20) and the single-arm design limit statistical power and external generalizability. Of the 20 participants initially enrolled, 6 discontinued early—4 due to rapid clinical deterioration related to disease progression and 2 due to worsening emotional well-being—illustrating the clinical fragility of this population rather than limitations of the digital health solution itself. Importantly, despite the high clinical vulnerability commonly observed in aLC, most patients who remained in the study were able to sustain meaningful engagement with the platform. This may have been facilitated by the frequency of in-person visits, family or caregiver support, and the involvement of both the clinical team and the biomedical engineering support. These combined elements may have helped reinforce motivation and support sustained engagement with the digital health solution.

Other populations with less frequent clinical contact or fewer support resources may require additional strategies—such as adaptive reminders, motivational feedback, or gamification—to maintain long-term engagement. In addition, the economic implications of implementing such digital health interventions were not assessed in this study, and future research should examine cost-effectiveness and scalability in routine practice.

Finally, while this study confirmed feasibility, usability, and high user satisfaction, several findings suggest the potential for meaningful clinical benefits. The rapid resolution of a substantial number of clinical alerts, the remote identification of unsuspected tumor progression in 2 patients, and the preservation of QoL and healthy lifestyle behaviors over time indicate that the intervention may have supported early clinical decision-making and proactive patient management. Nonetheless, as a pilot study with a limited sample size, it was not powered to evaluate clinical outcomes or survival effects, which should be investigated in larger controlled trials.

XSL•FO
RenderX

Conclusions

Future studies should aim to validate these findings in larger, multicenter randomized trials and assess the impact of the ALIBIRD mHealth platform on clinical outcomes, nutritional status, and QoL over longer periods. Integration of artificial intelligence models could further enhance the platform's predictive capacity, enabling personalized alerts, adaptive feedback, and early detection of adverse events.

From a clinical perspective, ALIBIRD demonstrates the potential of combining digital monitoring, precision nutrition, and PROs to enable more proactive and individualized oncology care. Its holistic, multidisciplinary design supports continuous data collection and facilitates the incorporation of lifestyle, molecular, and behavioral variables into personalized interventions.

Moving forward, early-stage work is currently focused on 2 key areas. The first involves adapting the platform for use in other tumor types and in clinical scenarios with lower symptom burden, such as survivorship care or follow-up of patients in remission, enabling customization to both tumor-specific and individual patient needs. The second area relates to the design of a new care model within the oncology department to manage the continuous flow of patient-generated information and support earlier clinical responses. This will require establishing coordinated workflows involving health care professionals and biomedical engineering teams to ensure efficient data management and timely intervention.

In summary, this pilot study highlights the high usability, sustained adherence, and positive patient experience achieved with the ALIBIRD mHealth platform. These results underscore its feasibility for integration into routine oncology care and lay the groundwork for future research aimed at scaling up digital health, patient-centered strategies in cancer management.

Acknowledgments

The authors would like to express their gratitude to the patients who participated in the clinical study, to all health care professionals and researchers of the ALIBIRD Program, and to the students of the biomedical engineering degree at the UPM who collaborated in the project. The authors also wish to acknowledge the work and support of Professor Tiziana Fornari-Reali and Raisa González in coordinating and managing the ALIBIRD2020-CM program. Generative artificial intelligence (ChatGPT and OpenAI) was used to assist in the language editing of this manuscript, specifically to improve grammar, phrasing, and readability of the English text. All scientific content, interpretation of results, and data analyses were conceived, written, and verified by the authors. The authors take full responsibility for the integrity and accuracy of the content.

Funding

This publication is funded by the Community of Madrid, TEC-2024/BIO-281 IALIBIRD-CM (ORDER 5696/2024, B.O.C.M. N°. 307, 12/26/2024). The underlying research was also supported by the ALIBIRD2020-CM program (S2018/BAA-4343): "Therapeutic precision nutrition formulations for cancer," and by NutriSION-CM (Y2020/BIO-6350), funded by the Research and Innovation General Directorate of the Community of Madrid and cofinanced with European Union Structural Funds. Additional support was provided by grant PID2022-138295OB-I00 funded by MICIU/AEI/10.13039/501100011033 and by the European Union, the REACT-EU Program (FACINGLCOVID-CM project, Community of Madrid, and the European Regional Development Fund [ERDF]), and EU Structural Funds.

Data Availability

The data supporting the findings of this study are not publicly available due to institutional and consortium restrictions on data sharing. The datasets include pseudonymized clinical and digital health information collected within the ALIBIRD research program. Access to deidentified data may be considered upon reasonable request to the corresponding author in compliance with applicable ethical approvals and data protection regulations.

Authors' Contributions

Conceptualization: JMI-C, MS, JM-R, MVM-A, ARM, EJG, EC
Data curation: JMI-C, MS, BG-R, BT
Formal analysis: JMI-C, MS, BG-R
Funding acquisition: MVM-A, GRRa, ARM, EJG, EC
Investigation: JMI-C, MS, BG-R, MVS, IE-S, SFM, GRR, BT, YM, IHC, MMP, NM, MVM-A
Methodology: JMI-C, MS, MVS, JM-R, MVM-A, ARM, EJG, EC
Project administration: JM-R, MVM-A, GRRa, ARM, EJG, EC
Resources: IE-S, JM-R, MVM-A, GRRa, ARM, EJG, EC
Software: JMI-C, MS, BG-R, MVS, IE-S, SFM, GRR, BT, EJG, EC
Supervision: MVM-A, GRRa, ARM, EJG, EC
Writing – original draft: JMI-C, MS, BG-R
Writing – review & editing: JMI-C, MS, BG-R, MVS, IE-S, SFM, GRR, JM-R, NM, MVM-A, ARM, EJG, EC

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Technical description and development of the ALIBIRD mHealth platform.
[PDF File, 320 KB - cancer_v12i1e69525_app1.pdf]

Multimedia Appendix 2
Rule engine and patient-monitoring dimensions of the ALIBIRD platform.
[PDF File, 215 KB - cancer_v12i1e69525_app2.pdf]

Multimedia Appendix 3
Full methods for nutrigenetics and gut microbiota.
[PDF File, 297 KB - cancer_v12i1e69525_app3.pdf]

Multimedia Appendix 4
Detailed nutrigenetic results.
[PDF File, 148 KB - cancer_v12i1e69525_app4.pdf]

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Laugsand EA, Sprangers MAG, Bjordal K, Skorpen F, Kaasa S, Klepstad P. Health care providers underestimate symptom intensities of cancer patients: a multicenter European study. Health Qual Life Outcomes 2010 Sep 21;8(1):104. [doi: 10.1186/1477-7525-8-104] [Medline: 20858248]
3. Basch E, Iasonos A, McDonough T, et al. Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study. Lancet Oncol 2006 Nov;7(11):903-909. [doi: 10.1016/S1470-2045(06)70910-X] [Medline: 17081915]
4. Fellowes D, Fallowfield LJ, Saunders CM, Houghton J. Tolerability of hormone therapies for breast cancer: how informative are documented symptom profiles in medical notes for “well-tolerated” treatments? Breast Cancer Res Treat 2001 Mar;66(1):73-81. [doi: 10.1023/a:1010684903199] [Medline: 11368413]
5. Beaver CC, Magnan MA. Managing chemotherapy side effects: achieving reliable and equitable outcomes. Clin J Oncol Nurs 2016 Dec 1;20(6):589-591. [doi: 10.1188/16.CJON.589-591] [Medline: 27857268]
6. Parikh RB, Basen-Enquist KM, Bradley C, et al. Digital health applications in oncology: an opportunity to seize. J Natl Cancer Inst 2022 Oct 6;114(10):1338-1339. [doi: 10.1093/jnci/djac108] [Medline: 35640986]
7. Collado-Borrell R, Escudero-Vilaplana V, Calles A, et al. Oncology patient interest in the use of new technologies to manage their disease: cross-sectional survey. J Med Internet Res 2018 Oct 23;20(10):e11006. [doi: 10.2196/11006] [Medline: 30355554]
8. Dai Y, Cook OY, Yeganeh L, Huang C, Ding J, Johnson CE. Patient-reported barriers and facilitators to seeking and accessing support in gynecologic and breast cancer survivors with sexual problems: a systematic review of qualitative and quantitative studies. J Sex Med 2020 Jul;17(7):1326-1358. [doi: 10.1016/j.jsxm.2020.03.004] [Medline: 32331967]
9. Nguyen H, Butow P, Dhillon H, Sundaresan P. A review of the barriers to using Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs) in routine cancer care. J Med Radiat Sci 2021 Jun;68(2):186-195. [doi: 10.1002/jmrs.421] [Medline: 32815314]
10. Basch E, Deal AM, Dueck AC, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017 Jul 11;318(2):197-198. [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
11. Basch E, Schrag D, Jansen J, et al. Digital symptom monitoring with patient-reported outcomes in community oncology practices: A U.S. national cluster randomized trial. JCO 2021 Dec 20;39(36_suppl):349527-349527. [doi: 10.1200/JCO.2021.39.36_suppl.349527]
12. Ligibel J. Lifestyle factors in cancer survivorship. J Clin Oncol 2012 Oct 20;30(30):3697-3704. [doi: 10.1200/JCO.2012.42.0638] [Medline: 23008316]
13. Bolte FJ, McTavish S, Wakefield N, et al. Association of sarcopenia with survival in advanced NSCLC patients receiving concurrent immunotherapy and chemotherapy. Front Oncol 2022;12:986236. [doi: 10.3389/fonc.2022.986236] [Medline: 36212442]

XSL•FO
RenderX

14. Ordovas JM, Ferguson LR, Tai ES, Mathers JC. Personalised nutrition and health. BMJ 2018 Jun 13;361:bmj.k2173. [doi: 10.1136/bmj.k2173] [Medline: 29898881]
15. Nasir A, Bullo MMH, Ahmed Z, et al. Nutrigenomics: epigenetics and cancer prevention: a comprehensive review. Crit Rev Food Sci Nutr 2020;60(8):1375-1387. [doi: 10.1080/10408398.2019.1571480] [Medline: 30729798]
16. Manor O, Dai CL, Kornilov SA, et al. Health and disease markers correlate with gut microbiome composition across thousands of people. Nat Commun 2020 Oct 15;11(1):5206. [doi: 10.1038/s41467-020-18871-1] [Medline: 33060586]
17. Cullin N, Azevedo Antunes C, Straussman R, Stein-Thoeringer CK, Elinav E. Microbiome and cancer. Cancer Cell 2021 Oct 11;39(10):1317-1341. [doi: 10.1016/j.ccell.2021.08.006] [Medline: 34506740]
18. Sereno M, Iniesta-Chamorro JM, Garrido-Rubiales B, Gomez EJ, Casado Sáenz E. M-Health in lung cancer: a literature review. SAGE Open Med 2023;11:20503121231172011. [doi: 10.1177/20503121231172011] [Medline: 37205155]
19. Langius-Eklöf A, Crafoord MT, Christiansen M, Fjell M, Sundberg K. Effects of an interactive mHealth innovation for early detection of patient-reported symptom distress with focus on participatory care: protocol for a study based on prospective, randomised, controlled trials in patients with prostate and breast cancer. BMC Cancer 2017 Jul 4;17(1):466. [doi: 10.1186/s12885-017-3450-y] [Medline: 28676102]
20. Moreno-Blanco D, Solana-Sánchez J, Sánchez-González P, et al. Technologies for monitoring lifestyle habits related to brain health: a systematic review. Sensors (Basel) 2019 Sep 26;19(19):4183. [doi: 10.3390/s19194183] [Medline: 31561599]
21. Gárate FJ, Chausa P, Whetham J, et al. EmERGE mHealth platform: implementation and technical evaluation of a digital supported pathway of care for medically stable HIV. Int J Environ Res Public Health 2021 Mar 18;18(6):3156. [doi: 10.3390/ijerph18063156] [Medline: 33803821]
22. Cros MR, Hernando ME, García-Sáez G. Digital health and telehealth for pregnancy. In: Diabetes Digital Health and Telehealth: Elsevier; 2022:187-199. [doi: 10.1016/B978-0-323-90557-2.00021-2]
23. Lau Y, Wong SH, Chee DGH, et al. Technology-delivered personalized nutrition intervention on dietary outcomes among adults with overweight and obesity: a systematic review, meta-analysis, and meta-regression. Obes Rev 2024 May;25(5):e13699. [doi: 10.1111/obr.13699] [Medline: 38296771]
24. Anders C, Moorthy P, Svensson L, et al. Usability and user experience of an mHealth app for therapy support of patients with breast cancer: mixed methods study using eye tracking. JMIR Hum Factors 2024 Mar 5;11:e50926. [doi: 10.2196/50926] [Medline: 38441959]
25. Wu W, Graziano T, Salner A, et al. Acceptability, effectiveness, and roles of mHealth applications in supporting cancer pain self-management: integrative review. JMIR Mhealth Uhealth 2024;12:e53652. [doi: 10.2196/53652]
26. Kessel KA, Vogel MM, Alles A, Dobiasch S, Fischer H, Combs SE. Mobile app delivery of the EORTC QLQ-C30 questionnaire to assess health-related quality of life in oncological patients: usability study. JMIR Mhealth Uhealth 2018 Feb 20;6(2):e45. [doi: 10.2196/mhealth.9486] [Medline: 29463489]
27. Langius-Eklöf A, Craftman Å, Gellerstedt L, et al. Digital and nurse-led support intervention in primary care during the first year after curative intent treatment for breast or prostate cancer: study protocol of two cluster randomised controlled pilot trials. BMJ Open 2025 Feb 22;15(2):e090848. [doi: 10.1136/bmjopen-2024-090848] [Medline: 39986997]
28. Gustavell T, Sundberg K, Langius-Eklöf A. Using an interactive app for symptom reporting and management following pancreatic cancer surgery to facilitate person-centered care: descriptive study. JMIR Mhealth Uhealth 2020 Jun 17;8(6):e17855. [doi: 10.2196/17855] [Medline: 32554375]
29. Crafoord MT, Ekstrand J, Sundberg K, Nilsson MI, Fjell M, Langius-Eklöf A. Mobile electronic patient-reported outcomes and interactive support during breast and prostate cancer treatment: health economic evaluation from two randomized controlled trials. JMIR Cancer 2025 Mar 11;11:e53539. [doi: 10.2196/53539] [Medline: 40067346]
30. Programa ALIBIRD2020-CM, S2018/BAA-4343—fórmulas terapéuticas de nutrición de precisión para luchar contra el cáncer [ALIBIRD2020-CM program, S2018/BAA-4343—therapeutic precision nutrition formulas to fight cancer]. IMDEA Food RI on FS. 2019. URL: https://www.food.imdea.org/blog/imdea-alimentaci%C3%B3n-participa-en-el-proyecto-alibird-2020-que-busca-f%C3%B3rmulas-terap%C3%A9uticas-de [accessed 2019-07-29]
31. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17 Suppl 1(Suppl 1):11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
32. Pfaeffli L, Maddison R, Whittaker R, et al. A mHealth cardiac rehabilitation exercise intervention: findings from content development studies. BMC Cardiovasc Disord 2012 May 30;12(1):36. [doi: 10.1186/1471-2261-12-36] [Medline: 22646848]
33. Trettin B, Danbjørg DB, Andersen F, Feldman S, Agerskov H. Development of an mHealth app for patients with psoriasis undergoing biological treatment: participatory design study. JMIR Dermatol 2021 May 10;4(1):e26673. [doi: 10.2196/26673] [Medline: 37632817]
34. Marent B, Henwood F, Darking M, EmERGE Consortium. Development of an mHealth platform for HIV care: gathering user perspectives through co-design workshops and interviews. JMIR Mhealth Uhealth 2018 Oct 19;6(10):e184. [doi: 10.2196/mhealth.9856] [Medline: 30339132]
35. Kokol P. Agile software development in healthcare: a synthetic scoping review. Appl Sci (Basel) 2022;12(19):9462. [doi: 10.3390/app12199462]

XSL•FO
RenderX

36. Ramos Cardozzo D. Desarrollo de Software: Requisitos, Estimaciones y Análisis, 3rd edition: CreateSpace Independent Publishing Platform; 2018.
37. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
38. Muscaritoli M, Arends J, Bachmann P, et al. ESPEN practical guideline: Clinical Nutrition in cancer. Clin Nutr 2021 May;40(5):2898-2913. [doi: 10.1016/j.clnu.2021.02.005] [Medline: 33946039]
39. Bozzetti F, Mariani L, Lo Vullo S, et al. The nutritional risk in oncology: a study of 1,453 cancer outpatients. Support Care Cancer 2012 Aug;20(8):1919-1928. [doi: 10.1007/s00520-012-1387-x] [Medline: 22314972]
40. Baldwin C, Spiro A, Ahern R, Emery PW. Oral nutritional interventions in malnourished patients with cancer: a systematic review and meta-analysis. J Natl Cancer Inst 2012 Mar 7;104(5):371-385. [doi: 10.1093/jnci/djr556] [Medline: 22345712]
41. Rivadeneira DE, Evoy D, Fahey TJ, Lieberman MD, Daly JM. Nutritional support of the cancer patient. CA Cancer J Clin 1998;48(2):69-80. [doi: 10.3322/canjclin.48.2.69] [Medline: 9522822]
42. Martínez-González MA, García-Arellano A, Toledo E, et al. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects: the PREDIMED trial. PLoS One 2012;7(8):e43134. [doi: 10.1371/journal.pone.0043134] [Medline: 22905215]
43. Common terminology criteria for adverse events (CTCAE) version 50. National Cancer Institute, US Department of Health and Human Services. 2017. URL: https://dctd.cancer.gov/research/ctep-trials/trial-development [accessed 2025-07-24]
44. Koller M, Warncke S, Hjermstad MJ, et al. Use of the lung cancer-specific Quality of Life Questionnaire EORTC QLQ-LC13 in clinical trials: a systematic review of the literature 20 years after its development. Cancer 2015 Dec 15;121(24):4300-4323. [doi: 10.1002/cncr.29682] [Medline: 26451520]
45. Cederholm T, Jensen GL, Correia M, et al. GLIM criteria for the diagnosis of malnutrition—a consensus report from the global clinical nutrition community. J Cachexia Sarcopenia Muscle 2019 Feb;10(1):207-217. [doi: 10.1002/jcsm.12383] [Medline: 30920778]
46. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8(1):115. [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
47. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4]
48. Kirsch M, Mitchell SA, Dobbels F, et al. Linguistic and content validation of a German-language PRO-CTCAE-based patient-reported outcomes instrument to evaluate the late effect symptom experience after allogeneic hematopoietic stem cell transplantation. Eur J Oncol Nurs 2015 Feb;19(1):66-74. [doi: 10.1016/j.ejon.2014.07.007] [Medline: 25190633]
49. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
50. Annunziata MA, Muzzatti B, Bidoli E, et al. Hospital Anxiety and Depression Scale (HADS) accuracy in cancer patients. Support Care Cancer 2020 Aug;28(8):3921-3926. [doi: 10.1007/s00520-019-05244-8] [Medline: 31858249]
51. Brooke J. SUS: A quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, Eds Usability Evaluation in Industry: London, UK: CRC Press; 1996:189-194.
52. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2024-04-07]
53. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/BF03195514]
54. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? Hum Factors 1992 Aug;34(4):457-468. [doi: 10.1177/001872089203400407]
55. Maguire R, Ream E, Richardson A, et al. Development of a novel remote patient monitoring system: the advanced symptom management system for radiotherapy to improve the symptom experience of patients with lung cancer receiving radiotherapy. Cancer Nurs 2015;38(2):E37-E47. [doi: 10.1097/NCC.0000000000000150] [Medline: 24836956]
56. Tang FWK, Chan CWH, Choy YP, Loong HHF, Chow KM, So WKW. A feasibility study on using tablet personal computers for self-reported symptom assessment in newly diagnosed lung cancer patients. Int J Nurs Pract 2018 Aug;24(4):e12658. [doi: 10.1111/ijn.12658] [Medline: 29642280]
57. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
58. Molinero N, Taladrid D, Zorraquín-Peña I, et al. Ulcerative colitis seems to imply oral microbiome dysbiosis. Curr Issues Mol Biol 2022 Mar 30;44(4):1513-1527. [doi: 10.3390/cimb44040103] [Medline: 35723361]
59. Taladrid D, de Celis M, Belda I, Bartolomé B, Moreno-Arribas MV. Hypertension- and glycaemia-lowering effects of a grape-pomace-derived seasoning in high-cardiovascular risk and healthy subjects. Interplay with the gut microbiome. Food Funct 2022 Feb 21;13(4):2068-2082. [doi: 10.1039/d1fo03942c] [Medline: 35107113]
60. Almeida A, Mitchell AL, Boland M, et al. A new genomic blueprint of the human gut microbiota. Nature New Biol 2019 Apr;568(7753):499-504. [doi: 10.1038/s41586-019-0965-1] [Medline: 30745586]

XSL•FO
RenderX

61. Rinninella E, Raoul P, Cintoni M, et al. What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases. Microorganisms 2019 Jan 10;7(1):14. [doi: 10.3390/microorganisms7010014] [Medline: 30634578]
62. Latorre-Pérez A, Hernández M, Iglesias JR, et al. The Spanish gut microbiome reveals links between microorganisms and Mediterranean diet. Sci Rep 2021 Nov 10;11(1):21602. [doi: 10.1038/s41598-021-01002-1] [Medline: 34759297]
63. Casén C, Vebø HC, Sekelja M, et al. Deviations in human gut microbiota: a novel diagnostic test for determining dysbiosis in patients with IBS or IBD. Aliment Pharmacol Ther 2015 Jul;42(1):71-83. [doi: 10.1111/apt.13236] [Medline: 25973666]
64. Kirkpatrick S, Davey Z, Wright PR, Henshall C. Supportive eHealth technologies and their effects on physical functioning and quality of life for people with lung cancer: systematic review. J Med Internet Res 2024 Jul 26;26:e53015. [doi: 10.2196/53015] [Medline: 39059003]
65. Schreiweis B, Pobiruchin M, Strotbaum V, Suleder J, Wiesner M, Bergh B. Barriers and facilitators to the implementation of eHealth Services: systematic literature analysis. J Med Internet Res 2019 Nov 22;21(11):e14197. [doi: 10.2196/14197] [Medline: 31755869]
66. Fristedt S, Smith F, Grynne A, Browall M. Digi-Do: a digital information tool to support patients with breast cancer before, during, and after start of radiotherapy treatment: an RCT study protocol. BMC Med Inform Decis Mak 2021 Feb 25;21(1):76. [doi: 10.1186/s12911-021-01448-3] [Medline: 33632215]
67. Martin L, Birdsell L, Macdonald N, et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. J Clin Oncol 2013 Apr 20;31(12):1539-1547. [doi: 10.1200/JCO.2012.45.2722] [Medline: 23530101]
68. Takenaka Y, Oya R, Takemoto N, Inohara H. Predictive impact of sarcopenia in solid cancers treated with immune checkpoint inhibitors: a meta-analysis. J Cachexia Sarcopenia Muscle 2021 Oct;12(5):1122-1135. [doi: 10.1002/jcsm.12755] [Medline: 34337889]
69. Ticinesi A, Meschi T, Narici MV, Lauretani F, Maggio M. Muscle ultrasound and sarcopenia in older individuals: a clinical perspective. J Am Med Dir Assoc 2017 Apr 1;18(4):290-300. [doi: 10.1016/j.jamda.2016.11.013] [Medline: 28202349]
70. Micó V, Díez-Ricote L, Daimiel L. Nutrigenetics and nutrimiromics of the circadian system: the time for human health. Int J Mol Sci 2016 Feb 26;17(3):299. [doi: 10.3390/ijms17030299] [Medline: 26927084]
71. Porzi M, Burton-Pimentel KJ, Walther B, Vergères G. Development of personalized nutrition: applications in lactose intolerance diagnosis and management. Nutrients 2021 Apr 29;13(5):1503. [doi: 10.3390/nu13051503] [Medline: 33946892]
72. Liu F, Li J, Guan Y, et al. Dysbiosis of the gut microbiome is associated with tumor biomarkers in lung cancer. Int J Biol Sci 2019;15(11):2381-2392. [doi: 10.7150/ijbs.35980]
73. Zhuang H, Cheng L, Wang Y, et al. Dysbiosis of the gut microbiome in lung cancer. Front Cell Infect Microbiol 2019;9:112. [doi: 10.3389/fcimb.2019.00112] [Medline: 31065547]
74. Zheng Y, Fang Z, Xue Y, et al. Specific gut microbiome signature predicts the early-stage lung cancer. Gut Microbes 2020 Jul 3;11(4):1030-1042. [doi: 10.1080/19490976.2020.1737487] [Medline: 32240032]
75. Zhao F, An R, Wang L, Shan J, Wang X. Specific gut microbiome and serum metabolome changes in lung cancer patients. Front Cell Infect Microbiol 2021;11:725284. [doi: 10.3389/fcimb.2021.725284] [Medline: 34527604]
76. Liu NN, Ma Q, Ge Y, et al. Microbiome dysbiosis in lung cancer: from composition to therapy. NPJ Precis Oncol 2020 Dec 10;4(1):33. [doi: 10.1038/s41698-020-00138-z] [Medline: 33303906]
77. Zhang M, Zhou H, Xu S, et al. The gut microbiome can be used to predict the gastrointestinal response and efficacy of lung cancer patients undergoing chemotherapy. Ann Palliat Med 2020 Nov;9(6):4211-4227. [doi: 10.21037/apm-20-2183]

Abbreviations

aLC: advanced lung cancer
LC: lung cancer
MEDAS: Mediterranean Diet Adherence Screener
mHealth: mobile health
NSCLC: non–small cell lung cancer
PREM: patient-reported experience measure
PRO: patient-reported outcome
PROM: patient-reported outcome measure
QoL: quality of life
REDCap: Research Electronic Data Capture
SNP: single nucleotide polymorphism
SUS: System Usability Scale

XSL•FO
RenderX

Edited by N Cahill; submitted 10.Dec.2024; peer-reviewed by A Rodríguez, S Tungare, WKH Mok; revised version received 16.Nov.2025; accepted 17.Nov.2025; published 11.Feb.2026.

Please cite as:
Iniesta-Chamorro JM, Sereno M, Garrido-Rubiales B, Villarino Sanz M, Espinosa-Salinas I, Falagán Martínez S, Rubio Romero G, Moreno-Rubio J, Tabarés B, Martín Y, Hernández de Córdoba I, Morales Parga M, Molinero N, Moreno-Arribas MV, Reglero Rada G, Ramírez de Molina A, Gómez EJ, Casado E
Evaluation of the ALIBIRD mHealth Platform for Care of Patients With Lung Cancer: Prospective Pilot Study
JMIR Cancer 2026;12:e69525
URL: https://cancer.jmir.org/2026/1/e69525
doi:10.2196/69525

© Jose M Iniesta-Chamorro, Maria Sereno, Beatriz Garrido-Rubiales, Marta Villarino Sanz, Isabel Espinosa-Salinas, Sandra Falagán Martínez, Gustavo Rubio Romero, Juan Moreno-Rubio, Beatriz Tabarés, Yolanda Martín, Irene Hernández de Córdoba, Maria Morales Parga, Natalia Molinero, M Victoria Moreno-Arribas, Guillermo Reglero Rada, Ana Ramírez de Molina, Enrique J Gómez, Enrique Casado. Originally published in JMIR Cancer (https://cancer.jmir.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of a Mobile Nutrition App on Dietary Outcomes in Cancer Survivors: Pilot Feasibility Study

Sang Won Park[1], PhD; Woo Jin Kim[2*], MD, PhD; Oh Beom Kwon[2*], MD, PhD; Inhyeok Yim[3*], MD

[1]Institute of Clinical Research, Biolink Inc, Daegu, Republic of Korea

[2]Department of Pulmonary Medicine, Kangwon National University, Chuncheon, Gangwon-do, Republic of Korea

[3]Department of Family Medicine, Kangwon National University Hospital, 156, Baengnyeong-ro, Chuncheon, Gangwon-do, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Inhyeok Yim, MD
Department of Family Medicine, Kangwon National University Hospital, 156, Baengnyeong-ro, Chuncheon, Gangwon-do, Republic of Korea

Abstract

Background: Cancer survivors frequently face persistent nutrition-related challenges after treatment. Mobile health tools may extend access to dietary self-management support beyond clinic settings, but feasibility and preliminary effects remain insufficiently characterized in this population.

Objective: This study aims to evaluate the feasibility, user engagement, and preliminary effects of a 4-week mobile nutrition app on dietary behavior and quality of life (QoL) among cancer survivors, and to explore whether higher engagement is associated with greater improvements.

Methods: A single-arm, prospective pilot feasibility study was conducted at a tertiary cancer center in Korea. Participants used a mobile nutrition app that provided dietary feedback and self-monitoring features. In-app log data were analyzed to determine engagement metrics (session frequency, duration, and gap regularity) using an elbow-based 10-minute session threshold. Primary outcomes included the Nutrition Quotient for Adults (NQ-2021) and European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30) scores, measured before and after 4 weeks. Nonparametric paired analyses assessed changes, and exploratory correlations examined relationships between engagement and outcomes.

Results: Among 27 enrolled participants, 24 cancer survivors (88.9%) completed the intervention and postassessment; the majority were female (17/24, 70.8%), with a mean age of 58.5 (SD 8.7) years. Breast cancer was the most common diagnosis (11/24, 45.8%), and most participants reported no diet-related adverse effects (20/24, 83.3%) and stable body weight during the study period. Participants averaged 2.3 app sessions per day and a median cumulative use of 177.5 minutes. Retention was 88.9%, and median adherence to daily self-monitoring exceeded 85%. The NQ Moderation domain improved significantly (mean 76.6, SD 17.5 → mean 81.0, SD 13.7; *P*=.02), while Balance and Dietary Behavior showed positive trends (mean 63.7, SD 16.1 → mean 65.5, SD 13.3, *P*=.14; mean 64.9, SD 17.3 → mean 67.1, SD 15.9, *P*=.10). In QoL outcomes, appetite loss decreased (mean 17.9, SD 22.0 → mean 7.7, SD 14.6; *P*=.03) and global health status increased modestly (mean 68.5, SD 21.4 → mean 72.9, SD 20.1; *P*=.08). Higher engagement correlated with improved moderation (*r*=0.46; *P*=.02) and reduced appetite loss (*r*=–0.42; *P*=.04). Exploratory subgroup analyses suggested stronger effects among participants aged 60 years and older (ΔNQ Moderation +7.9; *P*=.04) and those with longer cancer survivorship (>3 years; *P*=.047). No adverse events were reported.

Conclusions: This pilot feasibility study demonstrates high user engagement, satisfactory retention, and preliminary improvements in nutritional behavior and QoL among cancer survivors using a mobile nutrition app. These findings indicate the feasibility of a larger controlled trial to confirm the app's effectiveness and explore long-term adherence strategies.

(JMIR Cancer 2026;12:e79215) doi:10.2196/79215

KEYWORDS

mobile health; cancer; nutrition; digital health; quality of life; mobile phone

Introduction

Poor nutritional status is a highly prevalent and serious problem in cancer care, affecting 40% - 80% of patients depending on cancer type and treatment phase [1,2]. This high prevalence reflects both tumor-related metabolic alterations and treatment-induced side effects, such as appetite loss or gastrointestinal dysfunction, which compromise dietary intake. Malnourished patients experience poorer treatment tolerance, higher complication rates, delayed recovery, and diminished

XSL•FO
RenderX

quality of life (QoL). Poor nutritional status has also been associated with reduced survival [3,4]. Nutritional support interventions—including early dietetic counseling and individualized nutrition plans—can mitigate these effects and improve both clinical and patient-reported outcomes [1,5]. Accordingly, current oncology guidelines recommend routine nutrition screening and intervention across all phases of cancer care [2,6,7].

Despite these recommendations, access to professional dietary support remains limited. Many cancer survivors report unmet nutritional needs due to lack of referrals, time constraints, and system-level barriers. In 1 study, only 23% of survivors had consulted a dietitian after treatment, most citing difficulty accessing expert guidance [8]. Consequently, many survivors turn to unverified online information, underscoring the urgent need for accessible, evidence-based digital tools to support nutritional self-management [9]. Importantly, nutritional challenges in cancer survivorship are heterogeneous and extend beyond inadequate intake alone. While many survivors experience appetite loss and reduced dietary intake during or after treatment, others—particularly survivors of breast, colorectal, and endometrial cancers—are prone to excess weight gain or unhealthy dietary patterns that are associated with increased risks of recurrence and metabolic comorbidities. Accordingly, contemporary survivorship guidelines, including those of the American Cancer Society (ACS), emphasize not only sufficient energy and protein intake but also the adoption of balanced, high-quality dietary patterns and moderation of excessive or nutrient-poor consumption. Nutritional support in survivorship, therefore, requires an approach that addresses both insufficient intake and maladaptive overconsumption, rather than focusing solely on weight loss or caloric increase.

Mobile health (mHealth) apps represent a promising and scalable approach to extend nutrition support beyond the clinic. These apps can provide personalized dietary education, facilitate self-monitoring of food intake, and enable real-time symptom tracking beyond traditional clinical settings [9,10]. Digital symptom monitoring during active cancer treatment has been associated with improved patient outcomes—including better QoL, fewer emergency visits, and even survival benefits—by enabling earlier supportive interventions [3,11,12]. Building on these advances, early studies of mHealth-based nutrition interventions in cancer care have reported encouraging outcomes. App-based programs targeting diet or weight management have demonstrated improvements in dietary quality, weight maintenance, and patient satisfaction in small pilot trials [12-14]. For instance, a mobile app for head and neck cancer survivors showed high engagement and acceptability for diet and symptom tracking. However, despite these encouraging findings, the overall evidence base remains limited. Most prior studies have been small or short-term, with few using rigorous designs or enrolling diverse cancer populations. Moreover, the relationship between app engagement and nutritional or QoL outcomes remains largely unexplored. These gaps underscore the need for a rigorously designed pilot study to establish feasibility parameters and explore engagement-outcome relationships before a future definitive trial.

Given these knowledge gaps and the widespread nutrition challenges among cancer survivors, we developed a new mHealth-based dietary intervention. This intervention was designed to support healthy dietary pattern formation across a broad spectrum of survivorship needs, including individuals experiencing appetite loss and insufficient intake as well as those requiring dietary balance and moderation to prevent excess weight gain. Before long-term efficacy can be meaningfully evaluated, it is therefore necessary to establish whether a candidate intervention is feasible in real-world survivorship settings, whether users engage with it consistently, and whether early behavioral and symptom changes can be detected. Accordingly, we conducted a structured pilot feasibility study to evaluate this intervention with a specific focus on recruitment, retention, engagement patterns, and short-term behavioral responsiveness. The objectives were not to establish long-term efficacy but to assess feasibility and preliminary effects and to generate critical parameters—such as adherence rates, engagement intensity, and engagement-outcome associations—required to inform the design of a future randomized, long-term trial. By integrating multidimensional dietary guidance with detailed, log-based engagement analytics and validated patient-reported outcomes, this study aims to address a key methodological gap in prior mHealth nutrition research and to provide foundational evidence for the development of rigorously designed, longer-term intervention studies in cancer survivorship.

Methods

Study Design and Sample

This pilot study was conducted as a single-arm, prospective feasibility trial in Korea. The study followed the CONSORT (Consolidated Standards of Reporting Trials) 2010 extension for pilot and feasibility trials, with no protocol modifications or interim analyses after trial initiation [15]. To enhance completeness of descriptive reporting, we also considered selected STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) items relevant to participant characteristics and outcome reporting; however, CONSORT served as the principal reporting framework for this feasibility intervention [16]. Participants were recruited from Family Medicine and Hematology-Oncology outpatient clinics of a tertiary cancer center between October and December 2024. Treating physicians identified potentially eligible cancer survivors during follow-up visits and referred them to a research nurse with more than 3 years of oncology care experience. The nurse conducted eligibility screening, obtained written informed consent, provided training on app installation and use, and monitored participants throughout the study period. Recruitment was conducted consecutively via physician referrals and informational leaflets displayed in the clinics. A single research nurse provided standardized app training and follow-up support for all participants to ensure procedural consistency. Eligible participants were adult (aged 20 - 69 years) cancer survivors who had completed curative-intent treatment—surgery, chemotherapy, and/or radiotherapy for any cancer type—and owned a smartphone capable of running the study app. All participants had undergone curative surgical resection as their

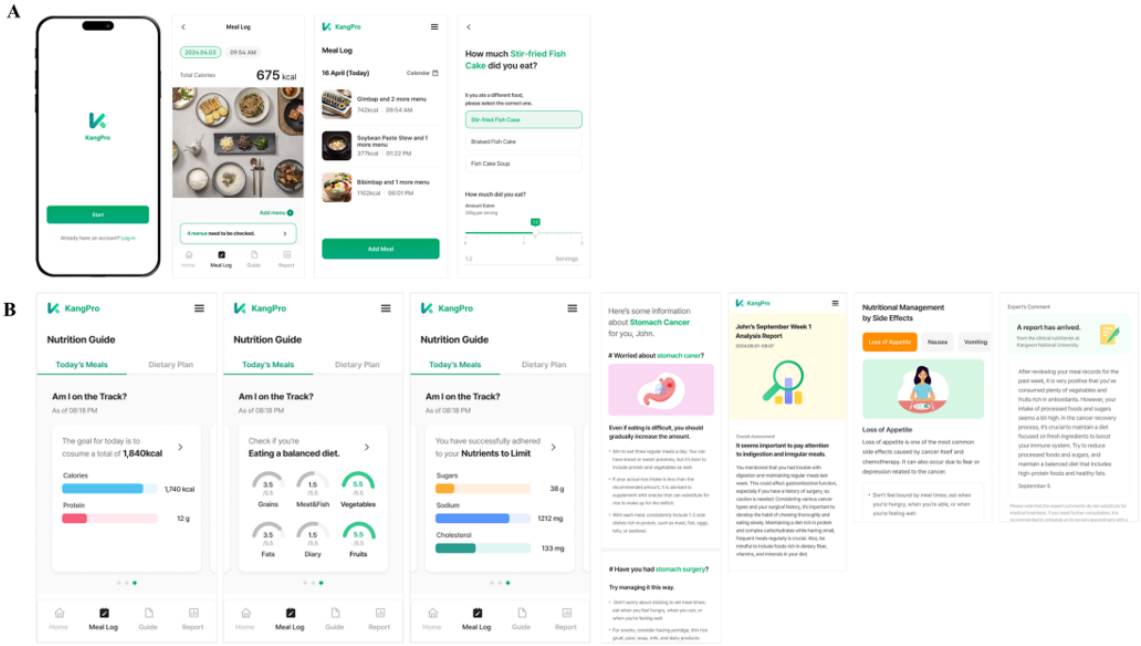
XSL•FO
RenderX

primary treatment and were within 5 years of completing curative therapy. Except for 4 individuals with mild physical limitations, most participants were able to perform light daily activities without restriction. Information on employment status and comorbidities was not collected in this pilot study. Key exclusion criteria were ongoing active cancer treatment or hospice or palliative care, refusal to participate, significant cognitive impairment precluding informed consent, and inability to use the smartphone app, even with caregiver assistance. Participants could withdraw at any time without penalty. Based on institutional registry data (average 2153 new cancer diagnoses annually, 2019 - 2023), a simple proportion estimate with a 0.10 margin of error and 90% CI suggested a required sample of approximately 66 participants for adequate precision. However, as this pilot aimed primarily to evaluate feasibility outcomes (recruitment, adherence, and usability) rather than statistical efficacy, a pragmatic target of approximately 50 participants was set, accounting for potential attrition. A total of 27 participants were enrolled, of whom 24 completed the 28-day intervention and provided both pre- and postintervention data for analysis. Recruitment and follow-up were planned over a 4-week intervention window per participant. Feasibility endpoints included recruitment rate, retention, and adherence; these parameters were predefined to assess study progression. No outcome measures were modified after trial initiation.

Mobile App Intervention and Data Acquisition

Participants received access to a custom mobile app designed to support dietary management in cancer survivorship. The app was intentionally designed to support balanced, sustainable eating behaviors across heterogeneous survivorship needs—including both insufficient intake due to appetite loss and excessive or unbalanced intake associated with weight gain—rather than focusing solely on weight loss (Figure 1A). Nutrient composition and feedback were based on the Korea Food and Drug Safety Food and Nutrient Database. The app generated artificial intelligence–based nutrition reports analyzing meal composition and nutrient balance using this database (Figure 1A and B). Participants could also report symptoms, which investigators monitored in real time. The app allowed users to record their meals and dietary intake through text search or photograph uploads and to view a personalized summary of their nutritional patterns. It also automatically generated artificial intelligence–based nutrition reports that analyzed meal composition and nutrient balance using a national food composition database. The app also enabled remote symptom monitoring, as participants could report symptoms that investigators could oversee in real time.

Figure 1. Nutrition reports by generative artificial intelligence in mobile apps for patients with cancer. (**A**) Recording meals process (**B**) Result nutrition analysis and reports by generative artificial intelligence.



Although the app generated quantitative dietary intake estimates for user feedback, these values were not prespecified as study end points and were therefore not analyzed for efficacy evaluation. The primary quantitative outcomes were validated questionnaire-based measures of dietary quality based on Nutrition Quotient for Adults (NQ-2021) and European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30), along with app engagement metrics. Participants were instructed to use the app as desired over the 28-day study period, and weekly phone or text message check-ins by study staff were used to encourage engagement. These check-ins were standardized brief contacts focused on technical support and general encouragement, without providing individualized dietary counseling or clinical advice, and were applied uniformly to all participants. Participants who had difficulty navigating the app could be assisted by a family member or caregiver.

All user interactions with the app were automatically recorded in time-stamped log files. For data analysis, raw usage logs were filtered to include records only from valid study

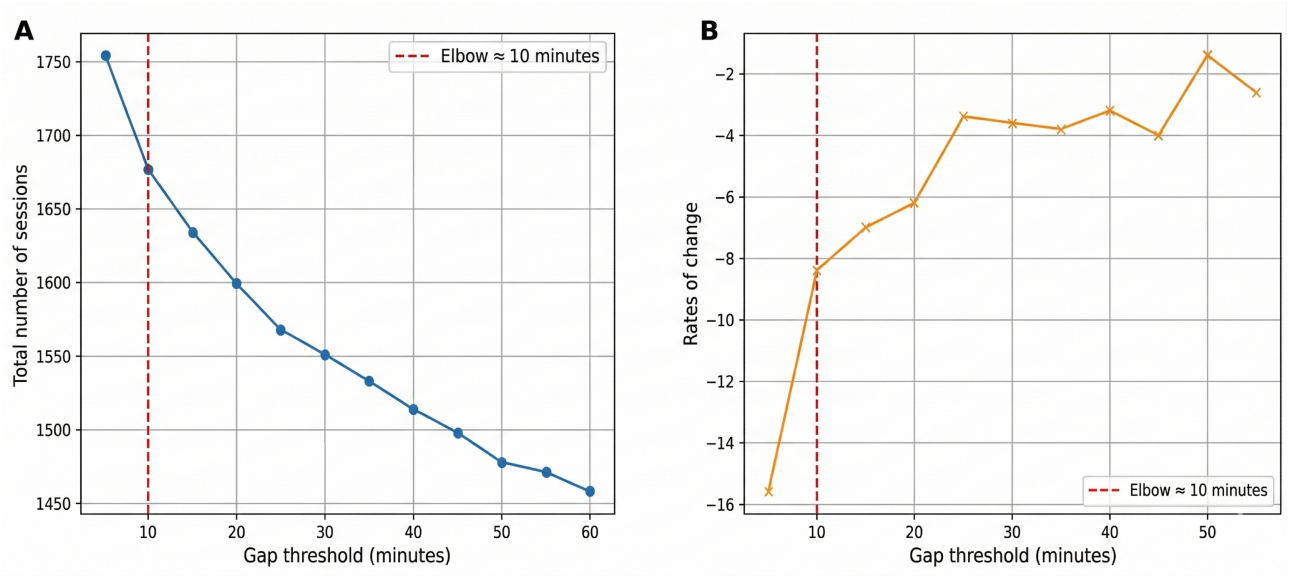
XSL•FO
RenderX

participants and restricted to each participant's 28-day intervention window. An app usage session was defined as a series of consecutive interactions separated by less than 30 minutes. In other words, any app events occurring within a 30-minute interval were grouped as 1 session, and a break of 30 minutes or more initiated a new session.

Operational Definition of App Sessions

To define user sessions from time-stamped login data, we applied a threshold-based approach that dynamically identifies session boundaries [17,18]. Specifically, we examined a range of time gaps (0 - 60 minutes in 10-minute increments) between consecutive login events to identify an appropriate threshold for segmenting discrete usage sessions (Figure 2A and B). For each gap threshold, user activity was aggregated such that any login occurring after a period longer than the defined threshold was considered the start of a new session. This procedure was repeated for all users, and the total number of sessions was calculated for each threshold value. To identify the optimal cutoff, we analyzed the rate of change in total session counts across thresholds by calculating the first derivative. This approach quantified how quickly session counts stabilized as the threshold increased. The elbow point—where increasing the threshold produced only minimal further reduction in the total session count—was interpreted as the balance between splitting continuous usage into too many fragments and merging clearly separate sessions into one. This inflection occurred at a 10-minute gap, which we therefore adopted as the operational session boundary. In simple terms, very small gaps overdivide normal continuous use, while very large gaps merge independent activities into a single session; the 10-minute "elbow" provides the most stable midpoint between these extremes. This data-driven procedure provides an objective and reproducible threshold, analogous to the elbow method commonly used in clustering analysis [19]. Our approach further aligns with prior studies on user sessionization heuristics, which typically use gap-based thresholds (around 25 - 30 minutes) to segment user activity into meaningful behavioral units [20].

Figure 2. Identification of optimal session gap threshold using the elbow method. (A) Total number of app usage sessions across gap thresholds from 0 to 60 minutes. As the allowed time gap between consecutive interactions increases, separate activities are more likely to be grouped into the same session, leading to a gradual decrease in the total number of sessions. (**B**) First derivative showing rate of change in session count. The elbow point at 10 minutes (red dashed line) was selected as the optimal threshold for defining discrete app usage sessions.



QoL Assessment

QoL was measured using the EORTC QLQ-C30 (version 3.0) [21]. The survey for EORTC QLQ-C30 was administered before and after app use following a detailed, face-to-face explanation of the survey procedures provided by a specialized research nurse (SJP), with additional consultation from clinical specialists (WJK, OBK, and IHY). This validated, cancer-specific instrument includes 30 items covering 5 functional domains, 3 multi-item symptom domains, 1 global health domain, and 6 single-item symptom scales. Scores were transformed to standardized 0 - 100 scales according to the EORTC scoring manual [22]; higher functional and global health scores indicated better QoL, whereas higher symptom scores reflected greater symptom burden.

Nutrition Assessment

Dietary intake and behavior were evaluated using the NQ-2021, a nutrition assessment tool developed and validated by the Korean Nutrition Society [23]. The survey for NQ-2021 was administered before and after app use following a detailed, face-to-face explanation of the survey procedures provided by a specialized research nurse, with additional consultation from clinical specialists (WJK, OBK, and IHY). The NQ-2021 consists of 18 items in 3 domains: Balance (consumption frequency of recommended food groups), moderation (frequency of unhealthy foods or eating behaviors), and practice (healthy eating habits). Each item score is weighted according to an established algorithm, and domain scores are summed and standardized to a 0 - 100 scale. Higher NQ-2021 scores reflect better overall diet quality and healthier eating behaviors.

XSL•FO
RenderX

Statistical Analysis

Because of the behavioral nature of the intervention, neither participants nor investigators were blinded. All statistical analyses were descriptive and exploratory. All statistical analyses were descriptive and exploratory. Within-subject pre-post comparisons were conducted using the Wilcoxon signed rank test, appropriate for small samples with nonnormal distributions. Higher postintervention functional or nutritional scores were interpreted as improvement, whereas lower symptom scores indicated reduced symptom burden. Exploratory subgroup analyses examined differences in outcomes by engagement level and participant demographics.

Engagement metrics included total usage time, maximum consecutive days of app use (≥14,≥21, and ≥28 days), average sessions per day, mean intersession gap time, and the standard deviation of intersession gaps as a measure of variability. Analyses were performed on complete cases; missing data were described but not imputed. Statistical analyses were conducted using Python (version 3.9; Python Software Foundation) and R (version 4.1; R Core Team).

Ethical Considerations

The study was approved by the Kangwon National University Hospital Institutional Review Board (KNUH-2024-07-010) and conducted in accordance with the Declaration of Helsinki and relevant Korean privacy regulations. Written informed consent was obtained from all participants. The original consent permitted secondary analyses of deidentified data without additional approval. All data were anonymized using coded identifiers and stored on encrypted, access-restricted servers. Participants received a modest gift card upon study completion, as approved by the institutional review board. No protocol modifications or interim analyses were performed after trial initiation.

Results

Study Population and Baseline Characteristics

A total of 24 participants completed both pre- and postintervention assessments, including 17 (70.8%) females (Table 1). Except for 4 participants with minor physical limitations, all were able to perform light daily activities without restriction. The mean age was 61.57 (SD 8.22) years for males and 55.94 (SD 8.21) years for females. Primary tumors included breast cancer (11/24, 45.83%), colorectal cancer (3/24, 12.50%), gastric cancer (3/24, 12.50%), lung cancer (3/24, 12.50%), thyroid cancer (2/24, 8.33%), lymphoma (1/24, 4.17%), and renal cancer (1/24, 4.17%). Most participants used the app for weight maintenance (n=20, 83.3%), while 4 (16.7%) aimed for weight loss. In terms of physical activity, 12 (50.0%) reported unrestricted activity, 8 (33.3%) light-intensity activity, and 4 (16.7%) a predominantly sedentary lifestyle. Meal frequency was reported as 3 times daily by 18 (75.0%) participants, twice daily by 5 (20.8%), and 4 times by 1 (4.2%). Most participants (20/24, 83.3%) were classified as not applicable with respect to diet-related adverse effects. Body weight remained stable, with a mean of 60.8 (SD 11.5) kg before and 61.0 (SD 11.4) kg after the intervention.

XSL•FO
RenderX

Table . Patient characteristics.

Variables	Values (N=24)
Sex (female), n (%)	17 (70.8)
Age (years), mean (SD)	
Male	61.57 (8.22)
Female	55.94 (8.21)
Weight (kg), mean (SD)	
Before (mobile app usage)	60.8 (11.5)
After (mobile app usage)	61.0 (11.4)
After operation, mean (SD)	3.2 (1.9)
App usage purpose for weight, n (%)	
Maintenance	20 (83.3)
Loss	4 (16.7)
Activity, n (%)	
Light-intensity physical activity	8 (33.3)
Unrestricted physical activity	12 (50.0)
Predominantly sedentary lifestyle	4 (16.7)
Meals (in a day), n (%)	
Twice	5 (20.8)
Three times	18 (75.0)
Four times	1 (4.2)
Dietary-related adverse effects, n (%)	
Not applicable	20 (83.33)
Immunosuppression and throat pain	1 (16.7)
Constipation	1 (16.7)
Multisymptomatic (constipation, xerostomia, throat pain, etc)	1 (16.7)
Xerostomia	1 (16.7)
Primary tumor, n (%)	
Thyroid	2 (8.33)
Colorectal	3 (12.50)
Lymphoma	1 (4.17)
Renal	1 (4.17)
Gastric	3 (12.50)
Breast	11 (45.83)
Lung	3 (12.50)

Feasibility Outcomes—Engagement and App Usage Characteristics

A total of 27 participants completed baseline assessments and initiated the intervention. Of these, 24 (88.9%) participants completed the full 4-week intervention and postintervention assessments. Three participants discontinued app use within the first 2 weeks of the study and did not complete the postintervention questionnaires. All 3 noncompleters completed baseline nutrition and QoL assessments, but postintervention data were missing due to early discontinuation of app use. No partial postintervention questionnaires were recorded. Participants who did not complete the study (n=3) did not differ markedly from completers with respect to age, sex, cancer type, or baseline nutrition and QoL scores. However, all noncompleters exhibited early disengagement from app use.

Among participants who completed the intervention, using the predefined sessionization criterion, a total of 82,385 raw app interaction log entries were consolidated into 1551 distinct usage sessions across the 24 participants during the 4-week intervention period. Overall engagement with the app was high. Participants used the app a mean of 2.3 sessions per day, with

XSL•FO
RenderX

a median cumulative usage time of 177.5 minutes over the intervention period. Engagement was sustained throughout the study, with no marked decline in usage over time. Adherence to daily self-monitoring was high, with a median adherence rate exceeding 85%, and 88.9% (24/27) of enrolled participants completed the full 4-week intervention, indicating satisfactory retention. Patterns of use varied between individuals, but most participants demonstrated regular engagement characterized by repeated short sessions distributed across the day, consistent with meal logging and dietary review behaviors. Analysis of engagement intensity showed that participants differed in session frequency, cumulative usage time, and regularity of use. These engagement metrics were subsequently used to stratify participants into subgroups (eg, by consecutive days of use, total usage time, daily usage frequency, and intersession gap regularity) for exploratory analyses examining associations between engagement patterns and nutritional and QoL outcomes.

Changes in Nutrition and QoL After App Use

After the 4-week app intervention, participants showed significant improvements in specific nutritional and symptom outcomes. In particular, the nutrition quotient moderation domain score increased from 76.61 (SD 15.36) at baseline to 81.04 (SD 12.69) postintervention, a statistically significant improvement (P=.02) (Table 2). This indicates a meaningful reduction in the frequency of unhealthy food choices and overeating behaviors after using the app. In contrast, other nutrition domain scores—including balance (fruits or vegetables intake) and implementation (healthy eating practices)—and the overall nutrition score did not change significantly. Among QoL measures, the most notable change was a decrease in appetite loss symptoms. The mean appetite loss score improved from 17.94 (SD 22.0) (higher scores indicate worse symptoms) to 7.69 (SD 14.62) after the intervention (P=.03), suggesting that participants experienced significantly less loss of appetite following app-guided dietary management. Other EORTC QLQ-C30 domains, including global health status and functional scales (physical, role, emotional, cognitive, and social), showed no statistically significant pre-post differences. Symptom scales such as fatigue, pain, nausea or vomiting, insomnia, constipation, diarrhea, and financial difficulties likewise exhibited no significant changes.

Table . Comparison of nutrition and quality of life before and after app use[a].

Characteristics	Before (n=24), mean (SD)	After (n=24), mean (SD)	*P* value
Nutrition			
Moderation	76.61 (15.36)	81.04 (12.69)	.02[a]
Balance	59.25 (16.95)	60.29 (16.04)	.20
Implementation	78.5 (13.22)	75.99 (11.74)	.88
Nutritional score	72.16 (10.63)	72.8(10.9)	.28
Quality of life			
Global health	53.2 (12.53)	47.43 (16.1)	.84
Appetite loss	17.94 (22.0)	7.69 (14.62)	.03[a]
Cognitive	82.03 (12.65)	82.05 (17.3)	.08
Emotional	78.2 (18.17)	69.87 (23.21)	.72
Social	94.87 (10.51)	76.92 (17.4)	>.99
Physical	86.16 (12.3)	80.51 (17.52)	.80
Role	83.34 (24.52)	79.49 (15.45)	.77
Nausea and vomiting	2.57 (6.27)	3.85 (9.99)	.20
Constipation	20.51 (25.6)	25.64 (33.76)	.84
Diarrhea	17.93 (17.28)	12.82 (16.88)	.67
Pain	15.38 (17.29)	12.82 (15.45)	.20
Dyspnea	23.07 (31.58)	25.64 (27.74)	.94
Insomnia	28.18 (18.49)	30.77 (21.35)	.96
Fatigue	32.45 (23.33)	42.73 (23.06)	.95
Financial difficulties	10.25 (16.0)	20.51 (21.68)	.92

[a]Statistically significant results (P<.05).

XSL•FO
RenderX

Changes in Nutrition and QoL Based on App Usage Intensity

Greater usage intensity was generally linked to greater improvements in nutrition and QOL measures. Consecutive days of app use (Table 3 and Table S1 in Multimedia Appendix 1) showed a clear dose-response relationship with outcomes. Participants who used the app for at least 14 consecutive days showed a significant increase in their Moderation score compared with baseline (*P*=.02). Those with ≥21 consecutive days of usage had a significant improvement in their moderation domain score (*P*=.04), reflecting better intake of recommended food groups. Furthermore, participants who engaged with the app for the full 28 days experienced significant gains in the balance domain (*P*=.02) and also in physical functioning (*P*=.047)—a QoL improvement not seen with shorter usage durations.

Table . Characteristics according to consecutive days of app use[a].

Days of app use	Before use	After use	*P* value
14 days (nutrition)			
Moderation	75.11 (17.51)	80.58 (13.74)	.02[b]
21 days (nutrition)			
Moderation	72.71 (20.14)	77.99 (14.25)	.04[b]
28 days (nutrition)			
Balance	55.8 (15.07)	65.29 (15.75)	.02[b]
28 days (quality of life)			
Physical	84.74 (10.69)	87.62 (7.13)	.047

[a]Results are expressed as mean (SD).

[b]Statistically significant results (*P*<.05).

Total accumulated usage time was another important factor (Table 4). Participants in the high total use time group (≥177.5 minutes of cumulative app use over 4 weeks) achieved significant improvements in the moderation domain and cognitive functioning scores (both *P*=.02). In contrast, those with lower total usage time (<177.5 minutes) did not show any significant changes in these outcomes, implying that a minimum threshold of engagement time may be necessary to produce measurable benefits.

XSL•FO
RenderX

Table . Comparison based on total app usage time[a].

Characteristics	Total app use time <177.5 minutes (n=12)			Total app use time ≥177.5 minutes (n=12)		
	Before	After	*P* value	Before	After	*P* value
Nutrition						
Moderation	76.45 (16.06)	82.33 (11.48)	.08	76.77 (15.33)	79.75 (14.18)	.02[b]
Balance	56.84 (19.08)	56.66 (16.62)	.28	61.66 (14.97)	63.92 (15.26)	.28
Implementation	73.35 (10.90)	72.51 (13.05)	.60	83.65 (13.74)	79.47 (9.55)	.92
Nutritional score	69.33 (11.97)	70.70 (11.92)	.26	74.99 (8.68)	74.89 (9.84)	.48
Quality of life						
Global health	49.30 (12.02)	47.92 (13.82)	.64	56.94 (15.43)	51.39 (16.22)	.86
Appetite loss	16.66 (22.47)	8.33 (15.07)	.11	5.55 (12.96)	0.00 (0.00)	.08
Cognitive	73.59 (22.96)	72.22 (28.72)	.39	81.93 (11.13)	86.11 (9.62)	.03[b]
Emotional	68.05 (26.79)	68.06 (31.35)	.25	81.24 (15.12)	75.00 (14.65)	.91
Social	86.12 (27.36)	79.17 (25.74)	.94	91.66 (15.08)	77.78 (19.24)	.99
Physical	85.01 (10.29)	85.56 (11.31)	.53	87.77 (12.66)	83.33 (17.17)	.83
Role	88.89 (14.78)	80.56 (21.12)	.93	81.95 (25.07)	86.11 (15.62)	.34
Nausea and vomiting	5.57 (8.22)	4.17 (10.36)	.25	1.39 (4.82)	1.39 (4.81)	.33
Constipation	24.99 (32.18)	27.78 (37.15)	.98	27.77 (31.25)	27.78 (27.83)	.57
Diarrhea	16.65 (17.39)	11.11 (16.41)	.46	5.55 (12.96)	5.55 (12.97)	.92
Pain	12.50 (14.42)	9.72 (11.14)	.29	22.22 (23.93)	20.83 (20.26)	.25
Dyspnea	13.88 (17.15)	16.67 (22.47)	.98	19.44 (33.21)	19.44 (26.43)	.61
Insomnia	36.09 (26.44)	38.89 (23.93)	.90	24.98 (20.72)	27.78 (19.24)	.91
Fatigue	38.87 (22.98)	43.52 (28.21)	.70	25.91 (24.31)	31.48 (18.25)	.95
Financial difficulties	22.22 (32.83)	16.67 (22.47)	.50	11.10 (16.40)	19.44 (22.28)	.99

[a]Values are expressed as mean (SD).

[b]Statistically significant results (*P*<.05).

Changes in Nutrition and QoL Based on App Usage Frequency

Daily usage frequency showed similar trends in Table 5. High-frequency users who used the app on average at least 2.5 times per day (eg, multiple log-ins or meal recordings per day) demonstrated significant improvements in various areas: their moderation and balance scores both increased significantly (*P*=.002 and .003, respectively), and their overall nutrition score (NQ total) also rose markedly (*P*=.008). Additionally, this high-frequency group showed a significant enhancement in cognitive QoL scores (*P*=.04) and pain (*P*=.047). In contrast, participants with low daily usage frequency (<2.5 uses per day) did not exhibit any statistically significant changes in nutrition scores or QoL measures. These results indicate that more frequent interaction with the app was associated with greater dietary improvements and perceived cognitive benefits.

XSL•FO
RenderX

Table . Comparison according to app usage counts in a day[a].

Characteristics	Low frequency (<2.5 per day, n=12)			High frequency (≥2.5 per day, n=12)		
	Before	After	*P* value	Before	After	*P* value
Nutrition						
Moderation	75.79 (16.13)	76.21 (14.80)	.44	77.42 (15.21)	85.88 (8.16)	.002[b]
Balance	56.22 (19.83)	51.01 (14.48)	.90	62.28 (13.69)	69.58 (11.84)	.003[b]
Implementation	77.14 (12.06)	73.58 (12.21)	.84	79.87 (14.69)	78.40 (11.24)	.75
Nutritional score	70.46 (11.61)	67.60 (11.18)	.96	73.86 (9.75)	77.99 (8.02)	.008[b]
Quality of life						
Global health	50.00 (13.76)	48.61 (15.82)	.66	56.24 (14.28)	50.69 (14.42)	.80
Appetite loss	19.43 (22.28)	8.33 (15.07)	.06	2.77 (9.61)	0.00 (0.00)	.16
Cognitive	72.21 (22.83)	72.22 (27.83)	.37	83.32 (10.04)	86.11 (11.96)	.04[b]
Emotional	68.74 (26.86)	64.58 (28.67)	.76	80.55 (15.61)	78.47 (17.21)	.57
Social	86.11 (26.42)	72.22 (23.92)	.98	91.67 (16.66)	84.72 (19.41)	.97
Physical	83.34 (11.89)	78.89 (17.49)	.95	89.43 (10.43)	90.00 (7.25)	.17
Role	81.95 (25.07)	75.00 (19.46)	.95	88.89 (14.78)	91.67 (13.29)	.23
Nausea and vomiting	2.78 (6.50)	5.56 (10.86)	.64	4.17 (7.55)	N/A[c]	.04[b]
Constipation	27.77 (31.25)	33.33 (37.61)	.93	24.99 (32.18)	22.22 (25.95)	.50
Diarrhea	13.88 (17.15)	11.11 (16.41)	.76	8.32 (15.06)	5.55 (12.97)	.50
Pain	16.66 (15.88)	16.67 (15.89)	.52	18.07 (24.06)	13.89 (18.58)	.047[b]
Dyspnea	24.99 (32.18)	25.00 (28.87)	.74	8.32 (15.06)	11.11 (16.41)	.97
Insomnia	33.32 (28.43)	36.11 (22.29)	.89	27.76 (19.24)	30.55 (22.29)	.91
Fatigue	40.73 (28.96)	46.30 (28.76)	.83	24.05 (14.84)	28.70 (14.57)	.91
Financial difficulties	13.88 (22.28)	22.22 (25.95)	.90	19.43 (30.01)	13.89 (17.16)	.83

[a]Results are presented as mean (SD).

[b]Statistically significant results (*P*<.05).

[c]N/A: not applicable.

Changes in Nutrition and QoL Based on Regular App Usage

Participants who used the app consistently throughout the day—defined as daily use with intervals of 7 hours or less (eg, morning, afternoon, and evening)—demonstrated significantly better dietary outcomes (Table 6). As shown in Table 6, this regular-use group exhibited improvements in moderation (*P*=.002), balance (*P*=.046), and overall nutrition score (*P*=.01). By contrast, those with irregular usage patterns (gaps >7 hours and variable timing) did not show improvements in these domains. These results indicate that consistent, frequent app engagement is associated with improved dietary behavior and cognitive outcomes. Notably, however, the irregular-use group reported a significant reduction in appetite loss symptoms (*P*=.04), while the regular-use group had minimal baseline appetite issues, leaving less room for improvement. No other QoL domains differed significantly by usage pattern.

XSL•FO
RenderX

Table . Comparison by app usage routine pattern.

Characteristics	Irregular usage pattern (>7hours, n=12)			Regular usage pattern (≤7 hours, n=12)		
	Before	After	*P* value	Before	After	*P* value
Nutrition						
Moderation	80.75 (8.86)[a]	80.00 (10.35)	.64	72.46 (19.42)	82.08 (15.07)	.002[b]
Balance	55.90 (19.24)	53.67 (13.16)	.69	62.60 (14.35)	66.92 (16.40)	.046[b]
Implementation	77.57 (13.49)	74.86 (11.68)	.70	79.43 (13.47)	77.12 (12.20)	.83
Nutritional score	72.02 (11.50)	70.04 (10.13)	.88	72.29 (10.19)	75.55 (11.36)	.01[b]
Quality of life						
Global health	53.48 (16.84)	51.39 (15.82)	.71	52.77 (11.43)	47.92 (14.27)	.76
Appetite loss	16.66 (22.47)	5.55 (12.97)	.04[b]	5.55 (12.96)	2.78 (9.62)	.33
Cognitive	74.98 (21.89)	76.39 (26.07)	.21	80.54 (13.90)	81.94 (18.06)	.12
Emotional	68.74 (23.87)	68.75 (22.51)	.36	80.55 (19.89)	74.31 (26.46)	.88
Social	84.72 (26.06)	70.83 (23.70)	.98	93.06 (16.60)	86.11 (8.58)	.97
Physical	84.45 (2.81)	83.89 (14.90)	.74	88.32 (9.90)	85.00 (14.25)	.73
Role	83.34 (25.61)	79.17 (20.26)	.85	87.50 (14.42)	87.50 (16.09)	.47
Nausea and vomiting	2.78 (6.50)	4.17 (10.36)	.50	4.17 (7.55)	1.39 (4.81)	.06
Constipation	24.98 (28.86)	25.00 (28.87)	.92	27.78 (34.33)	30.56 (36.12)	.60
Diarrhea	13.88 (17.15)	11.11 (16.41)	.76	8.32 (15.06)	5.55 (12.97)	.50
Pain	20.82 (17.58)	16.67 (15.89)	.289	13.90 (22.29)	13.89 (18.58)	.25
Dyspnea	19.43 (30.01)	19.44 (26.43)	.83	13.88 (22.28)	16.67 (22.47)	.92
Insomnia	33.32 (28.43)	36.11 (22.29)	.89	27.76 (19.24)	30.55 (22.29)	.90
Fatigue	40.73 (28.96)	39.81 (26.99)	.48	24.05 (14.84)	35.18 (21.63)	.98
Financial difficulties	13.88 (22.28)	22.22 (25.95)	.90	19.43 (30.01)	13.89 (17.16)	.83

[a]These values are expressed as mean (SD).

[b]Statistically significant results (*P*<.05).

Exploratory Subgroup Analyses

Exploratory subgroup analyses revealed variations in intervention effects by age, sex, and cancer type (Tables S1-S3 in Multimedia Appendix 2). Among older participants (aged 60 years and older), significant improvements were observed in overall diet quality—reflected by higher Moderation subscores (*P*=.02)—and in cognitive functioning (*P*=.02), whereas younger participants showed no comparable changes. Female survivors also demonstrated a significant postintervention increase in Moderation (*P*=.04) that was not evident in males. In addition, survivors with non–breast cancers (eg, colorectal, gastric, or lung) exhibited a marked improvement in diet quality (*P*=.008) and greater reductions in appetite loss symptoms, while breast cancer survivors—who had higher baseline diet quality and minimal appetite complaints—showed little change. These subgroup findings suggest that age, sex, and cancer type may moderate the magnitude of benefit derived from the app-based dietary intervention.

Discussion

Principal Results

Although this study involved a short 4-week intervention, its primary purpose was to generate feasibility evidence and behavioral insights to inform the design of subsequent long-term trials, rather than to establish long-term efficacy. While some participants expressed an interest in weight control, the primary focus of this study was to improve dietary quality, appetite regulation, and overall well-being rather than to promote weight loss per se. In this 4-week pilot study, a mobile dietary self-management app for cancer survivors was associated with improvements in key nutrition-related behaviors and symptoms. Participants' overall diet quality improved significantly, as shown by an increase in the Nutrition Quotient Moderation subscore (reflecting reduced consumption of unhealthy foods), and they reported a marked reduction in appetite loss symptoms. These positive changes are notable, given that many cancer survivors face challenges with inadequate dietary intake and appetite suppression during and after treatment.

XSL•FO
RenderX

The pronounced reduction in appetite loss may reflect not only improved nutritional intake but also partial alleviation of psychosocial factors (such as treatment-related stress or social isolation). The app provided personalized meal recommendations, diverse healthy recipes, and motivational messages, offering a multidimensional, individualized approach to stimulate appetite and encourage regular eating. This outcome aligns with survivorship nutrition guidelines, as the ACS encourages survivors to consume a variety of vegetables, fruits, and high-quality protein and to optimize the mealtime environment to help maintain or regain appetite—practices shown to support appetite recovery [24]. These findings illustrate that the guideline—implementation—effectiveness pathway proposed by the ACS can be operationalized through a digital platform, demonstrating that evidence-based nutrition guidance can translate into measurable behavioral and clinical improvements. By delivering tailored support consistent with these recommendations, the app may help survivors overcome appetite loss, a complex and common issue in cancer survivorship. Notably, maintaining a healthy weight is often recommended as part of recurrence prevention in breast, prostate, and endometrial cancer survivors. In this context, weight awareness or moderate control behaviors complement—rather than contradict—the goal of restoring appetite and balanced eating. A modest but significant improvement in pain symptoms was also noted among high-frequency app users, suggesting that consistent dietary self-monitoring may provide ancillary benefits for perceived physical well-being. Another key finding was a significant improvement in dietary moderation, indicating a shift away from unhealthy eating patterns (eg, frequent overeating, binge eating, or strong food aversions) toward more balanced dietary habits. This finding supports the emphasis of established dietary guidelines on moderation and limiting nutrient-poor, energy-dense foods. For example, ACS guidelines warn against Western-style diets high in fat and sugar and instead highlight the importance of high-fiber, low-fat diets with increased intake of plant-based foods while restricting processed, calorie-dense items [24]. Our results suggest that a digital intervention can facilitate these healthy dietary shifts in a real-world context, helping survivors adopt and sustain recommended eating behaviors. In addition, we observed an improvement in cognitive functioning scores among participants aged 60 years and older. Older cancer survivors are particularly susceptible to cancer-related cognitive impairment due to factors such as chemotherapy, chronic inflammation, sleep disturbances, and depression. The app included features that might support cognitive health, including suggestions for “brain-friendly” foods rich in omega-3 fatty acids and antioxidants, encouragement of regular meal routines, and tools for self-monitoring dietary intake. These components may have helped impose daily structure and improve mental clarity in older users. While this subgroup finding is exploratory, it raises the possibility that digital dietary interventions could confer ancillary cognitive benefits in older patients with cancer, warranting further investigation into the underlying mechanisms and durability of this effect. Importantly, our analysis found that greater engagement with the app was associated with larger improvements in outcomes. Participants who used the app more frequently, for longer total durations, and with more consistency (ie, fewer and shorter gaps between usage sessions) tended to achieve greater gains in healthy eating behaviors and QoL scores than those with lower usage. To our knowledge, this study is among the first in the cancer survivorship setting to capture such detailed app usage metrics (using time-stamped log data to define user sessions) and to link these engagement indicators to standardized nutrition and QoL outcomes. This analytic approach provides preliminary evidence that engagement may mediate the effectiveness of digital interventions. However, because this was an uncontrolled, short-term pilot, the observed relationships between engagement and outcomes should be interpreted as associative rather than causal, serving primarily as hypothesis-generating evidence for subsequent trials. This pilot was designed to assess short-term feasibility, adherence, and early symptom response rather than to verify long-term efficacy; the findings thus provide foundational data to inform the design and power estimation of future randomized controlled trials (RCTs).

Overall, these findings suggest that an app-based dietary program can serve as a comprehensive supportive tool that addresses not only survivors’ nutritional needs but also aspects of their psychological, cognitive, and behavioral health. The app’s capacity to deliver personalized nutrition guidance, promote beneficial behavior changes, and provide accessible support indicates that such digital interventions could be further refined and tailored to specific cancer types, age groups, and phases of survivorship for maximum benefit. Although the intervention period was limited to 4 weeks, this study was intentionally designed to align with the identified gap in the existing literature. Prior mHealth nutrition studies in cancer survivorship have been constrained not only by short follow-up durations but also by interventions narrowly focused on single components—such as calorie tracking or symptom logging—without examining how users engage with more comprehensive, behavior-oriented tools. In contrast, this study evaluated a multidimensional dietary intervention that integrates personalized nutrition guidance, lifestyle-related behavioral feedback, and symptom awareness, while explicitly quantifying real-world engagement patterns using objective log data. Within this framework, the 4-week time frame served as a necessary and methodologically appropriate first step to determine intervention feasibility, user adherence, and early behavioral signal detection—prerequisites for designing and justifying longer-term efficacy trials. Thus, rather than attempting to address long-term outcomes directly, this pilot study contributes to the literature by establishing whether the intervention demonstrates sufficient feasibility and engagement-driven responsiveness to warrant subsequent large-scale, long-term evaluation. This stepwise approach is consistent with recommended frameworks for the development and evaluation of complex digital health interventions.

Comparison With Prior Work

Our findings are consistent with, and extend, prior research on mHealth nutrition interventions for cancer survivors. Previous pilot studies of app-based diet or weight management programs have reported improvements in dietary quality, weight maintenance, and patient-reported outcomes, and a recent

XSL•FO
RenderX

systematic review of randomized trials concluded that mHealth interventions show promise for improving diet and physical activity in cancer survivor populations [12]. Similarly, the improvements observed in our study—such as enhanced dietary moderation and reduced symptom burden—mirror these earlier reports, reinforcing that app-based guidance can help meet survivors' nutritional needs. Distinct from most earlier studies, this work quantitatively linked real-world app engagement metrics (frequency, duration, and continuity of use) to measurable changes in dietary and QoL outcomes. This analytic approach provides empirical groundwork for defining optimal adherence thresholds and intervention intensity in future trials. We add to the literature by demonstrating a direct link between participants' engagement with the app and the magnitude of benefits achieved, a relationship that has seldom been examined in prior studies. Emerging evidence from other digital health studies also highlights the importance of sustained app engagement for successful behavior change and symptom management. For example, in a recent mobile weight loss program for breast cancer survivors, those who engaged more frequently with app features (such as logging meals or reading educational content) were significantly more likely to achieve clinically meaningful weight loss [25]. Likewise, Yunis et al [8] found that patients with cancer and caregivers had high adherence to a symptom-tracking app over 1 month, and their study demonstrated that detailed app use metadata (eg, time-stamped login information) can inform remote monitoring of patient engagement and well-being between clinic visits. Furthermore, adding interactive components to boost engagement can yield additional benefits: in an RCT, breast cancer survivors who participated in an app-based peer support community (vs using a tracking app alone) experienced significantly lower distress levels and increased physical activity [26].

Beyond these trials, recent systematic and observational evidence clarifies where this study advances the field (Table S1 in Multimedia Appendix 3). Shen et al [25] demonstrated that higher real-world engagement with a behavioral app predicted meaningful weight loss, supporting engagement as a key behavioral driver, although cognitive outcomes were not assessed. Wang et al [12] summarized 23 RCTs of mHealth diet and activity programs, confirming short-term benefits but highlighting limited mechanistic data and underrepresentation of older adults. Lu et al [27] found that nutritional interventions—including anti-inflammatory and fasting-mimicking diets and polyunsaturated fatty acids supplementation—can improve cognitive outcomes, yet these trials were investigator-driven rather than self-managed. Coro et al [28] qualitatively showed that survivors view diet as linked to mental clarity but struggle to maintain structured routines due to cognitive fog. Crowder et al [29] observed that higher monounsaturated fatty acids or saturated fatty acids ratios were associated with better short-term cognitive performance during chemotherapy, while McLeod et al [30] reported complex, context-dependent associations between Mediterranean diet adherence and cognitive outcomes in older adults. In contrast, this study uniquely integrates detailed engagement analytics with dietary, symptom, and QoL outcomes. By linking frequency, duration, and continuity of app use to improvements in diet quality and appetite loss, it provides concrete evidence supporting the engagement-outcome mechanism. Moreover, the inclusion of older survivors (aged 60 years and older) extends digital survivorship research into an age group seldom represented in prior studies and suggests potential cognitive and structural benefits associated with sustained app use. Taken together, these reports and our results underscore that maximizing patient engagement is vital for mHealth interventions to achieve their full potential in cancer care. Ensuring that survivors not only adopt an app but continue to use it actively and regularly appears to be a prerequisite for attaining meaningful improvements in health behaviors and QoL.

Limitations

Several limitations of this pilot study must be acknowledged. First, this study had a modest sample size (n=24 completers), an uncontrolled single-arm design, and a short 4-week intervention, which limit generalizability and preclude causal inference. The short follow-up also meant that we could not determine the longevity of the observed behavior changes. Second, all outcomes were measured via self-reported questionnaires, introducing potential recall bias and social desirability bias. Third, participants received weekly check-in calls or text messages from study staff to encourage app use; this human support likely improved adherence and also confounds attribution of the outcomes to the app alone (ie, higher engagement may have been partly driven by these external prompts). Indeed, blended approaches that combine digital tools with human coaching are known to boost user engagement and can potentially amplify outcomes [31], so our study cannot disentangle the app's independent effects from the influence of this added support. In addition, app-generated quantitative dietary intake data were not analyzed, as they were not prespecified efficacy end points and were variably incomplete due to early discontinuation of app use among noncompleters. Finally, while our sample included a range of ages and cancer types, the study was not powered to examine differences across subgroups; any exploratory observations of trends by age or diagnosis should therefore be interpreted with caution, as they were not the focus of this analysis. Another limitation is that detailed clinical and sociodemographic data—including employment status, comorbidities, and the exact interval since surgery—were not comprehensively collected. Although all participants were disease-free and within 5 years after curative surgery, the lack of these variables may limit contextual interpretation of outcomes. Despite these limitations, the completion rate (88.9%) and adherence rate (≥70%) exceeded predefined feasibility benchmarks, confirming the practicality and acceptability of the intervention. Notably, the absence of a control group and the short duration of follow-up mean that the durability of the behavior change is uncertain. This is a common issue in early mHealth studies, and a recent review highlighted that evidence for the long-term efficacy of nutrition apps in cancer care remains inconclusive [32]. Accordingly, our positive findings should be viewed as preliminary and hypothesis-generating. Although quantitative food record data were partially collected, response completeness was insufficient for reliable analysis. Therefore, this pilot

XSL•FO
RenderX

prioritized a qualitative dietary assessment tool that captures overall meal quality and behavioral changes, which better reflect the app's behavioral focus and feasibility objectives. Future large-scale studies will integrate quantitative nutrient analyses for comprehensive validation.

Future Directions

Building on this pilot, future research should validate and expand these findings using more robust study designs. In particular, larger RCTs are needed to confirm the relationships observed between app engagement and outcomes and to determine whether actively boosting user engagement can causally improve nutritional status or clinical results. Such trials should also include longer follow-up periods to assess whether improvements in diet and QoL are sustained over time, and they should incorporate objective outcome measures (eg, weight change or nutritional biomarkers) alongside self-reports for a more comprehensive evaluation of the intervention's impact. Another priority for future interventions is to leverage engagement-tailored personalization. Our data suggest that users vary in how actively and consistently they use the app, and these usage patterns likely influence the benefits gained. Tailoring the program to an individual's engagement profile could amplify its effectiveness. For example, the app could use just-in-time adaptive features such as real-time alerts or motivational feedback when a lapse in use is detected, or dynamically adjust the frequency of prompts and content based on each user's behavior. By using analytics to identify when a participant is at risk of disengaging, the intervention could deliver timely, personalized support (such as encouragement from a coach, simplified tasks, or new content to recapture interest). Such adaptive personalization may help prevent drop-off and maximize the "dose" of the intervention each participant receives, ultimately leading to better outcomes. Additionally, future efforts should integrate mHealth dietary apps into standard survivorship care. Rather than functioning in isolation, apps like ours can be embedded within oncology nutrition services and survivorship care plans to extend support beyond the hospital setting. For instance, integrating the app with routine follow-up visits or with electronic health records could enable nutritionists and clinicians to remotely monitor patients' dietary logs and symptoms and intervene earlier when issues arise [8]. Care teams might receive automated summaries of a survivor's app activity or alerts if a patient's engagement or nutritional status falls below a certain threshold. This type of health system integration would facilitate a more continuous model of care. By providing an accessible, scalable platform for dietary self-management that remains connected to professional support, a validated nutrition app could reach survivors who otherwise might not receive ongoing nutritional guidance.

Conclusions

This pilot study demonstrates the feasibility of a mobile app–based dietary intervention for cancer survivors and provides preliminary evidence of its benefits. Participants achieved improvements in diet-related measures and QoL, particularly in those domains (such as healthy eating habits and symptom relief) where the app's guidance was most directly applied. Notably, greater user engagement was associated with greater improvements, highlighting the critical role of sustained engagement in the success of digital health interventions. While these results require confirmation in controlled trials, they underscore the potential of personalized, engagement-driven mHealth tools to support the long-term health and well-being of cancer survivors.

Acknowledgments

Generative artificial intelligence (AI) was used to assist in improving spelling and grammar in the initial draft of this paper. The authors have made some revisions using generative AI (GPT Version 5) to produce a coherent and readable text of their research findings. Accountability of those using generative AI, transparency regarding full disclosure, and confidentiality are essential to maintaining the integrity of the scientific record and are core elements of JMIR Publications' editorial policy. With rapid advancements in AI technology, related policies, regulations [33], research, and best practices [34], the authors anticipate that *JMIR Cancer* will continue to lead and advance as an innovator in scientific publishing. The final version was subsequently proofread by a professional editor.

Funding

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant HI23C0733). This work was supported by the Gangwon Information & Culture Industry Promotion Agency (GICA) and National IT Industry Promotion Agency (NIPA), grant funded by the Korea government (MSIT).

Data Availability

Data are available on reasonable request to the corresponding authors and with appropriate ethical approvals.

Authors' Contributions

IHY, SWP, and WJK had full access to all of the data in this study and took responsibility for the integrity of the data and the accuracy of the data analysis. IHY and WJK reviewed this paper, supervised this study, and are the guarantors for this study. IHY and WJK contributed equally to this study as corresponding authors. IHY, SWP, and WJK did the study concept and design;

XSL•FO
RenderX

worked on the acquisition, analysis, or interpretation of data; drafted this paper; and handled the statistical analysis. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. All authors approved the final version before submission and critically revised this paper for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparison of nutrition and quality of life before and after app use by sex.
[DOCX File, 26 KB - cancer_v12i1e79215_app1.docx]

Multimedia Appendix 2
Comparison of nutrition and quality of life before and after app use by age group.
[DOCX File, 38 KB - cancer_v12i1e79215_app2.docx]

Multimedia Appendix 3
Systematic review.
[DOCX File, 18 KB - cancer_v12i1e79215_app3.docx]

Checklist 1
CONSORT (Consolidated Standards of Reporting Trials) checklist.
[PDF File, 335 KB - cancer_v12i1e79215_app4.pdf]

References

1. Ravasco P. Nutrition in cancer patients. J Clin Med 2019 Aug 14;8(8):1211. [doi: 10.3390/jcm8081211] [Medline: 31416154]
2. Arends J, Bachmann P, Baracos V, et al. ESPEN guidelines on nutrition in cancer patients. Clin Nutr 2017 Feb;36(1):11-48. [doi: 10.1016/j.clnu.2016.07.015] [Medline: 27637832]
3. Zhang X, Tang T, Pang L, et al. Malnutrition and overall survival in older adults with cancer: a systematic review and meta-analysis. J Geriatr Oncol 2019 Nov;10(6):874-883. [doi: 10.1016/j.jgo.2019.03.002] [Medline: 30917937]
4. Lis CG, Gupta D, Lammersfeld CA, Markman M, Vashi PG. Role of nutritional status in predicting quality of life outcomes in cancer--a systematic review of the epidemiological literature. Nutr J 2012 Apr 24;11(1):27. [doi: 10.1186/1475-2891-11-27] [Medline: 22531478]
5. Odelli C, Burgess D, Bateman L, et al. Nutrition support improves patient outcomes, treatment tolerance and admission characteristics in oesophageal cancer. Clin Oncol 2005 Dec;17(8):639-645. [doi: 10.1016/j.clon.2005.03.015]
6. Baguley BJ, Benna-Doyle S, Drake S, Curtis A, Stewart J, Loeliger J. Access to nutrition services and information after active cancer treatment: a mixed methods study. J Cancer Surviv 2024 Feb;18(1):176-185. [doi: 10.1007/s11764-023-01352-x] [Medline: 36823493]
7. Aapro M, Bossi P, Dasari A, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct;28(10):4589-4612. [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
8. Yunis R, Fonda SJ, Aghaee S, et al. Mobile app activity engagement by cancer patients and their caregivers informs remote monitoring. Sci Rep 2024;14(1). [doi: 10.1038/s41598-024-53373-w]
9. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565. [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
10. Ma TM, Yang T, Philipson R, Kishan AU, Lee P, Raldow AC. Web-based symptom monitoring with patient-reported outcomes during definitive radiation therapy with chemotherapy (SYMPATHY): a prospective single-center phase 1 study. Adv Radiat Oncol 2023;8(3):101073. [doi: 10.1016/j.adro.2022.101073] [Medline: 37273926]
11. Davis SW, Oakley-Girvan I. Achieving value in mobile health applications for cancer survivors. J Cancer Surviv 2017 Aug;11(4):498-504. [doi: 10.1007/s11764-017-0608-1]
12. Wang L, Langlais CS, Kenfield SA, et al. mHealth interventions to promote a healthy diet and physical activity among cancer survivors: a systematic review of randomized controlled trials. Cancers (Basel) 2022 Aug 6;14(15):3816. [doi: 10.3390/cancers14153816]
13. Ng KLS, Munisamy M, Lim JBY, Alshagga M. The effect of nutritional mobile apps on populations with cancer: systematic review. JMIR Cancer 2025 Feb 5;11:e50662 [FREE Full text] [doi: 10.2196/50662] [Medline: 39908548]
14. Zini EM, Lanzola G, Quaglini S, Bossi P, Licitra L, Resteghini C. A pilot study of a smartphone-based monitoring intervention on head and neck cancer patients undergoing concurrent chemo-radiotherapy. Int J Med Inform 2019 Sep;129:404-412. [doi: 10.1016/j.ijmedinf.2019.06.004]

XSL•FO
RenderX

15. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24;355:i5239. [doi: 10.1136/bmj.i5239] [Medline: 27777223]
16. Elm EV, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD]
17. Beard E, Marsden J, Brown J, et al. Understanding and using time series analyses in addiction research. Addiction 2019 Oct;114(10):1866-1884. [doi: 10.1111/add.14643] [Medline: 31058392]
18. Halfaker A, Keyes O, Kluver D, et al. User session identification based on strong regularities in inter-activity time. arXiv. Preprint posted online on Aug 4, 2019. [doi: 10.48550/arXiv.1411.2878]
19. Thorndike RL. Who belongs in the family? Psychometrika 1953 Dec;18(4):267-276. [doi: 10.1007/BF02289263]
20. Jones R, Rey B, Madani O, Greiner W. Generating query substitutions. 2006 Presented at: Proceedings of the 15th International Conference on World Wide Web; May 23-26, 2006; Edinburgh, Scotland p. 387-396. [doi: 10.1145/1135777.1135835]
21. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 3;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
22. Fayers PM, Aaronson N, Bjordal K, Groenvold M, Curran D, Bottomley A. EORTC QLQ-C30 Scoring Manual: This Manual Is Intended to Assist Users With Scoring Procedures for the QLQ-C30 Version 3 and Earlier, and the QLQ Supplementary Modules, 3rd edition: EORTC; 2001.
23. Yook SM, Lim YS, Lee JS, et al. Revision of Nutrition Quotient for Korean adults: NQ-2021. J Nutr Health 2022;55(2):278. [doi: 10.4163/jnh.2022.55.2.278]
24. Rock CL, Thomson CA, Sullivan KR, et al. American Cancer Society nutrition and physical activity guideline for cancer survivors. CA Cancer J Clin 2022 May;72(3):230-262. [doi: 10.3322/caac.21719] [Medline: 35294043]
25. Shen S, Salehi E, White C, Chen Y, Iyengar NM. Evaluation of a mobile behavior change program for weight loss in breast cancer survivors. NPJ Breast Cancer 2024 Jun 29;10(1):53. [doi: 10.1038/s41523-024-00659-x] [Medline: 38951532]
26. Jung M, Lee SB, Lee JW, et al. The impact of a mobile support group on distress and physical activity in breast cancer survivors: randomized, parallel-group, open-label, controlled trial. J Med Internet Res 2023 Aug 7;25:e47158. [doi: 10.2196/47158] [Medline: 37549004]
27. Lu Y, Yuan H, Li Y, et al. Effects of nutritional interventions on cognitive function in adult cancer survivors: a systematic review. J Clin Nurs 2024 Nov;33(11):4227-4253. [doi: 10.1111/jocn.17371] [Medline: 39021041]
28. Coro DG, Hutchinson AD, Banks S, Coates AM. Diet and cognitive function in cancer survivors with cancer‐related cognitive impairment: a qualitative study. Eur J Cancer Care 2020 Nov;29(6). [doi: 10.1111/ecc.13303]
29. Crowder SL, Welniak TL, Hoogland AI, et al. Diet quality indices and changes in cognition during chemotherapy. Support Care Cancer 2023 Jan;31(1):75. [doi: 10.1007/s00520-022-07513-5]
30. McLeod AP, Gabel K, Chakos K, Wefel J, Tussing-Humphreys LM, Oddo V. The effect of Mediterranean diet adherence on the relationship between cognitive performance and cancer survivorship. Public Health Healthcare 2025 Feb 11. [doi: 10.20944/preprints202502.0840.v1]
31. Dhar E, Bah AN, Chicchi Giglioli IA, et al. A scoping review and a taxonomy to assess the impact of mobile apps on cancer care management. Cancers (Basel) 2023 Mar 15;15(6):1775. [doi: 10.3390/cancers15061775] [Medline: 36980661]
32. Ng KLS, Munisamy M, Lim JBY, Alshagga M. The effect of nutritional mobile apps on populations with cancer: systematic review. JMIR Cancer 2025 Feb 5;11:e50662. [doi: 10.2196/50662] [Medline: 39908548]
33. EU AI Act: first regulation on artificial intelligence. European Parliament. 2023. URL: https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence [accessed 2026-03-01]
34. Coiera EW, Verspoor K, Hansen DP. We need to chat about artificial intelligence. Med J Aust 2023 Aug 7;219(3):98-100. [doi: 10.5694/mja2.51992] [Medline: 37302124]

Abbreviations

ACS: American Cancer Society
CONSORT: Consolidated Standards of Reporting Trials
QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
mHealth: mobile health
NQ-2021: Nutrition Quotient for Adults
QoL: quality of life
RCT: randomized controlled trial
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by M Balcarras, N Cahill; submitted 17.Jun.2025; peer-reviewed by C Lozano, K Fitzner; revised version received 06.Jan.2026; accepted 26.Jan.2026; published 31.Mar.2026.

Please cite as:
Park SW, Kim WJ, Kwon OB, Yim I
Impact of a Mobile Nutrition App on Dietary Outcomes in Cancer Survivors: Pilot Feasibility Study
JMIR Cancer 2026;12:e79215
URL: https://cancer.jmir.org/2026/1/e79215
doi:10.2196/79215

© Sang Won Park, Woo Jin Kim, Oh Beom Kwon, Inhyeok Yim. Originally published in JMIR Cancer (https://cancer.jmir.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating a Telehealth Coaching and Mobile-Based Digital Engagement Intervention for People With Cancer Using the Patient-Reported Outcomes Measurement Information System Global Health: Pilot Questionnaire Study

Joanne Lewis[1,2], PhD; Shen Wang[1]; Toni Rose Jue[1], PhD; Tim Atkins[1]; Scott A Irwin[3], MD, PhD; Puneet Nanda[1], MD, MPH; Davinder Sangar[1,4], BSc, MbCHb, MRCGP; Lynette Mackenzie[5], PhD; Raghav Murali-Ganesh[1], MBBS, FRANZCR

[1]Osara Health, Eveleigh, NSW, Australia

[2]Faculty of Medicine, Nursing, Midwifery and Health Sciences, The University of Notre Dame, Sydney, Australia

[3]Cedars-Sinai Cancer, and Department of Psychiatry and Behavioral Neurosciences, Cedars-Sinai Health System, Los Angeles, CA, United States

[4]University of Bristol, Bristol, United Kingdom

[5]Faculty of Medicine and Health, University of Sydney, Sydney, Australia

Corresponding Author:
Raghav Murali-Ganesh, MBBS, FRANZCR
Osara Health
4 Cornwallis St
Eveleigh, NSW, 2015
Australia
Phone: 61 0429724428
Email: raghav@osarahealth.com

Abstract

Background: People with cancer often experience unmet needs during treatment and survivorship, which can impact their ability to carry out daily tasks, reduce their quality of life, and limit their participation in work and social activities. Cancer Coach by CancerAid (now known as Osara Health) is a digital health intervention designed to address these needs through a combination of synchronous telehealth coaching and an asynchronous mobile app that supports behavior change and emotional well-being.

Objective: This study aimed to evaluate the impact of Cancer Coach by CancerAid on the mental and physical health of people with cancer using patient-reported outcomes.

Methods: Participants were referred to the program via insurers and hospital clinics. Health coaches administered the Patient-Reported Outcomes Measurement Information System 10-item Global Health Short Form via telephone at both the beginning and end of the intervention. This tool measures global physical health (GPH) and global mental health (GMH). Pre- and postintervention scores were analyzed using Wilcoxon signed rank tests. Independent 2-tailed *t* tests assessed whether changes in GPH and GMH scores were associated with the use of health coaching alone or in combination with the mobile app.

Results: Statistically significant improvements were observed in both GPH ($z=-4.97$; $P<.001$; $r=0.37$) and GMH ($z=-4.53$; $P<.001$; $r=0.34$), indicating moderate effect sizes in the 89 participants. The average T score point changes of 4.43 for GPH and 4.58 for GMH represented a minimal important change for participants. The improvement in the group GMH T score was reflected in the move from “good” to “very good” mental health status. Participants who engaged with both health coaching and the mobile app showed greater improvements in physical health, whereas those who received health coaching alone exhibited higher gains in mental health. This suggests that the mode of support may influence specific health outcomes.

Conclusions: The use of the Patient-Reported Outcomes Measurement Information System 10-item Global Health Short Form showed that participants had significant improvements in physical and mental health after participating in the Cancer Coach by CancerAid intervention. The integration of telehealth coaching with app-based support may enhance overall well-being and address holistic needs during cancer treatment and survivorship.

(JMIR Cancer 2026;12:e72647) doi:10.2196/72647

XSL•FO
RenderX

KEYWORDS

cancer; digital; telehealth; coaching; mental health; physical health

Introduction

Countries with strong health care systems have seen increased cancer survival rates due to improvements in early detection and treatment of cancer [1]. However, cancer treatment side effects can impact physical, cognitive, psychological, and emotional well-being, not only in the acute phase of treatment but also for months or years afterward [2]. Managing the immediate and long-term side effects can be further complicated by health care funding shortages and a lack of services [3]. Worldwide, rehabilitation options for addressing the long-term holistic survivorship needs of people with cancer, including resuming everyday roles and occupations and improving quality of life, are limited [4]. Health coaching and various digital health options have emerged as a way of addressing some of the unmet needs of people with cancer [5].

Health coaching is a collaborative and client-centered intervention that aims to improve the physiological, behavioral, psychological, and social well-being of people with chronic illnesses [6]. For people with cancer, it offers a holistic approach to managing overall health and well-being beyond that of the primary oncology care team [7], assisting with informational, behavioral, and social concerns, particularly after treatment. A longitudinal study of 3-month telephone health coaching with survivors of cancer showed significant improvements in patient-reported depression, anxiety, exercise, and quality of life. The improvements continued for 6 months after coaching ceased, with slight declines at 12 months [8]. Participants indicated that health coaching assisted with improving exercise and healthier eating, with the most helpful aspect being the motivation and feedback provided by a personal health coach [8]. Health coaching for survivors of cancer was also supported by a systematic review [9], which found improvements in quality of life, mood, and physical activity.

Health coaches use various techniques, including motivational interviewing, goal setting, and behavior change strategies, to support and empower people to improve their overall health [10]. Technological advances have increased digital and telehealth options and reduced the need for traditional in-person health coaching, particularly since the global COVID-19 pandemic as a way of staying connected with people during periods of isolation [11]. Digital and telehealth options are also beneficial to those in rural areas with community access difficulties or caring or work responsibilities [12]. Purcell et al [13] used the Patient-Reported Outcomes Measurement Information System 10-item Global Health Short Form (PROMIS-10) to assess changes in veterans' health status following participation in a telehealth coaching program. The study found significant improvements in PROMIS-10 physical and mental health scores after coaching, indicating that the telehealth coaching program positively influenced veterans' overall health and well-being [13].

For people with cancer, digital and telehealth support options can help alleviate the financial burdens associated with attending in-person appointments [14] and provide valuable assistance during treatment periods when they may be immunocompromised. Mobile health (mHealth) options in cancer care can encompass activation and support of self-management, exercise, and enablement of survivorship care delivery [15]. Specifically, mHealth applications vary in use from treatment side effect and symptom management, virtual medical consultations, exercise prescription, support groups, and education and care navigation [16]. The method of delivery can be synchronous (eg, telephone calls or virtual care) or asynchronous (eg, wearable devices, web-based platforms, mobile-based apps, mobile-based digital games, and SMS text messages) [16,17]. Digital health programs can improve global, physical, and mental health while providing support in an ambulatory setting for people with cancer in managing their disease symptoms and side effects [16] by enhancing patients' self-efficacy, as well as helping health care professionals track and respond to patient-reported outcomes [7]. Furthermore, digital cancer health coaching programs have also been found to improve physical and mental health scores [5,18,19]. However, most studies focus on a single mHealth intervention, and there is a paucity of research evaluating interventions that combine synchronous and asynchronous mHealth interventions.

Cancer Coach by CancerAid integrates synchronous coaching with asynchronous mobile support guided by the transtheoretical model [20]. The transtheoretical model conceptualizes behavior change as a progression through stages, from precontemplation to maintenance. This framework informed the structure of the intervention by aligning synchronous coaching with the earlier stages of change, incorporating motivational interviewing and goal setting to enhance readiness and intention. The asynchronous mobile app was designed to support participants across all stages but, in particular, in the action and maintenance stages through reminders, self-monitoring tools, and educational content that reinforces ongoing behavior change [20].

Previously, Cancer Coach by CancerAid has been found to improve people's ability to return to work compared with age- and disease-matched individuals who did not participate in the program [21]. The primary aim of this pilot study was to determine whether Cancer Coach by CancerAid (now known as Osara Health) improved participants' self-reported physical and mental health using PROMIS-10 scores [22-27], and a secondary aim was to explore which programmatic aspect contributed to these changes. This was to inform program improvements and modifications for an international cohort.

Methods

Study Recruitment

People with cancer were informed about the Cancer Coach by CancerAid program by their insurer, employer, or health service and self-enrolled into the program. CancerAid supports people at any time during their cancer experience. CancerAid coaches contacted people who self-enrolled and advised them about the program and the study.

XSL•FO
RenderX

Participants were matched using email addresses and mobile phone numbers to prevent multiple identities for users. Data were collected from August 2020 to November 2022 using secure online forms and databases. This period was chosen due to the consistency of having the same 3 health coaches. The end point was set to analyze data on broader app changes to enhance app integration into synchronous health coaching and ensure its relevance and applicability across an international cohort.

Eligibility criteria included a cancer diagnosis, completion of the PROMIS-10 survey before and after the program, and participation in at least 2 health coaching calls. Health coaches administered the PROMIS-10 questionnaire via phone before the first and after the final coaching call using a standard script for consistency (Multimedia Appendix 1). Use of the CancerAid app (Multimedia Appendix 2) was optional.

Study Design

This study used a nonexperimental evaluative research design to assess the effectiveness of the overall intervention and the programmatic elements that contributed to its effectiveness [28]. All participants in the study received health coaching and access to the CancerAid app. During the study period, there were minor app bug fixes, content changes, and downtimes to enhance performance and security. Details about this can be found in Multimedia Appendix 3. These interruptions did not significantly affect the participants during the study period. As no new features were introduced and app maintenance changes were back end in nature, these were not expected to materially affect user experience or outcomes. Therefore, they were not controlled for in the statistical analysis.

PROMIS-10 Measure

The PROMIS-10 is a standardized and validated tool [29] that established the methodological rigor for this study. The PROMIS-10 contains 10 questions related to the patient's physical, mental, and social health and well-being. The complete PROMIS-10 questionnaire can be found in Multimedia Appendix 1. Patients rate their well-being on a 5-point Likert scale. One question related to pain is rated on a scale from 1 to 10. This score for pain is recomputed to a score out of 5. All 10 question scores are then recalculated to present 2 domain scores: global physical health (GPH) and global mental health (GMH). Both GPH and GMH have a maximum possible score of 20 points. For all items, a higher score equals more of the concept being measured, and a lower score indicates less of the concept being measured. Multimedia Appendix 1 also provides inclusions for calculating GPH and GMH scores.

Although the PROMIS-10 is primarily used to generate composite scores for GPH and GMH, this study also analyzed individual item responses to explore specific symptom changes. While not standard practice, item-level analysis can yield clinically relevant insights [30], particularly in cancer populations where symptoms such as pain, fatigue, and emotional distress are common. These insights may help refine health coaching strategies and enhance understanding of how coaching influences distinct aspects of functioning.

The GPH and GMH scores were recalibrated using the same underlying metric, which corresponds to a normalized T score. The T scores are standardized to the general US population. The average T score for the US population is 50 points, with an SD of 10 points [26,29,31]. PROMIS-10 domains use different cutoff score ranges. The PROMIS-10 GPH score ranges are as follows: excellent (T score>58), very good (T score=50-57), good (T score=42-59), fair (T score=41-35), and poor (T score<35). The PROMIS-10 GMH score ranges are as follows: excellent (T score>57), very good (T score=48-56), good (T score=40-47), fair (T score=39-29), and poor (T score<29) [26].

To identify minimal important change [32] or "meaningful change" [26] in PROMIS-10 scores, a review of the literature found that a 2- to 6-point change in T scores for PROMIS-10 measures could be reasonably considered as an indicator [32]. Minimal important change has been defined as a threshold for a minimal within-person change over a period during which patients perceive themselves to have experienced an important change [32].

Intervention

Cancer Coach by CancerAid is a flexible, evidence-based program designed to support people with cancer over approximately 6 to 12 weeks. It combines telehealth coaching with a mobile app and resources to address key areas such as symptom tracking, nutrition, sleep, exercise, psychosocial well-being, and return to work. The educational content is based on American Society of Clinical Oncology guidelines and findings from large randomized trials, ensuring that topics such as lifestyle interventions and digital symptom tracking are backed by the latest evidence. Support is delivered via a combination of asynchronous digital resources (app, SMS text messages, and emails) and synchronous coaching calls, of which there are approximately 3 throughout the program.

Each participant is paired with a qualified health coach (with backgrounds in allied health, nursing, or medicine) trained in standardized methods that ensure the scope of practice while allowing for a client-centered approach. Coaches provide tailored interventions based on the individual's diagnosis, stage of cancer, and specific needs. They use evidence-based behavior change techniques, including goal setting and review, symptom monitoring, motivational interviewing, feedback and accountability, positive reinforcement, lifestyle education, and social and emotional support. The approach is based on the transtheoretical model of change, with an emphasis on empowering participants to take proactive steps toward improving their health and well-being. Regular communication and digital support help overcome barriers to face-to-face care, creating a supportive framework for sustained self-management and behavior change.

Further information about the Cancer Coach by CancerAid program can be found in the study by Lo et al [21].

Health coaching is complemented by the CancerAid app (now known as the Osara Health app), an mHealth app providing symptom tracking, health information articles, and supportive care resources. Key features include symptom logs, evidence-based articles, and a note-taking feature (Multimedia Appendix 2). The app was designed using principles of

XSL•FO
RenderX

behavioral science and patient-centered care, incorporating elements such as self-monitoring and educational empowerment. The development process of the CancerAid iOS and Android app followed the typical iterative agile development process of planning, design, development, testing, deployment, and review. The app was tested through an automated test suite in addition to manual testing by developers and stakeholders. Throughout the study period, no new features were introduced, and the new releases were limited to maintenance updates. The app was relatively stable during the study period and is archived in a private code repository owned by CancerAid. The app is free and publicly accessible via the Apple App Store and Google Play Store. Participants were encouraged by health coaches to use the app to supplement their health coaching calls.

Statistical Analysis

Physical and mental health were evaluated using the PROMIS-10 measure to determine the impact of Cancer Coach by CancerAid. An examination of the data distribution revealed no evidence of a ceiling effect; scores were well distributed across the full range, indicating adequate variability for analysis. Descriptive statistics were used to establish ranges and means of variables. Tests for normality showed that the data did not fit normal distributions, and therefore, nonparametric testing was used. Raw PROMIS-10 GPH and GMH scores were converted to T scores to compare with US population norms [26,29,31]. A T score of 50 (–10 to +10) is considered to be a healthy score, and well-being descriptors are ascribed based on scores and associated categories (see PROMIS-10 Measure section) [26,33]. Two-tailed *t* tests were used to determine the correlation between the number of calls and app use and changes in GPH and GMH scores. For an exploratory study on individual item score changes, the Wilcoxon signed rank test was used to compare the paired groups to determine whether there was a statistically significant difference ($P \leq .05$). The effect size of the Wilcoxon signed rank test was calculated and considered in terms of the classification of effect sizes by Cohen [34] (0.1=small effect; 0.3=moderate effect; ≥0.5=large effect) [29,34]. Data were analyzed using SPSS (version 29; IBM Corp).

Ethical Considerations

Ethics approval was obtained from the University of Sydney Human Research Ethics Committee (approval 2024/HE000297). Informed consent was obtained from all individual participants to take part in the research. Participation in the study was voluntary. Participant privacy and confidentiality were protected throughout the study in accordance with institutional, national, and international guidelines for human subjects research. No compensation was provided. All data were deidentified prior to analysis, and no identifying information is included in the published results.

Results

Participants

Between August 2020 and November 2022, a total of 200 participants completed a PROMIS-10 survey at either the start or end of the intervention, 93 (46.5%) of whom had both pre- and posttest data. A total of 2% (4/200) of the participants were excluded due to missing data. The 55% (107/200) attrition rate was due to participants not completing the postprogram PROMIS-10 survey for health, family, or work reasons, resulting in a final dataset of 89 participants. Table 1 provides demographic data.

Analysis between the group included in the study (completed pre- and posttest PROMIS-10 questionnaires; n=89) and those not included (did not complete the posttest PROMIS-10 questionnaire; n=111) was conducted. An independent-sample 2-tailed *t* test compared the mean ages of the 2 groups, revealing no significant difference (mean age 51.44 years; $P=.35$). Cross-tabulation analyses of gender showed no significant association (33/111, 30% male; $P=.73$). However, the chi-square test indicated a significant difference between the groups on diagnosis ($P<.001$). Notably, the nonstudy group had 4 times as many cases of an unknown cancer diagnosis compared to the study group. The top 5 cancer types in the group not included in the study were breast, unknown, bowel, blood, and lung. Employers comprised a larger proportion of the referral sources in the nonstudy group (employer: 11/111, 10%; insurer: 92/111, 83%; hospital: 8/111, 7%).

XSL•FO
RenderX

Table 1. Participant demographics (N=89).

Characteristic	Values
Sex, n (%)	
Male	32 (35.9)
Female	57 (64)
Intersex	0 (0)
Age (years), mean (SD; range)	52.86 (9.94; 23-82)
Cancer type, n (%)	
Breast	34 (38.2)
Blood	14 (15.7)
Bowel	8 (9)
Head and neck	8 (9)
Prostate	5 (5.6)
Brain	4 (4.5)
Gynecological	4 (4.5)
Melanoma	4 (4.5)
Unknown	4 (4.5)
Kidney	2 (2.2)
Other (single cases: pancreas, liver, lung, spinal cord, and neuroendocrine)	5 (5.6)
Referral source, n (%)	
Insurer	78 (87.6)
Hospital	10 (11.2)
Employer	1 (1.1)
Duration of program (days), mean (SD)	52 (32.5)

PROMIS-10 Global Health Scores and T Scores

A statistically significant mean change of 4.43 points was observed in the GPH T score (P<.001); however, the score remained within the "Good" category of well-being. Similarly, the mean GMH T score increased by 4.48 points following participation in the Cancer Coach by CancerAid intervention, a change that was also statistically significant (P<.001). The mean postintervention GMH T score was 49.052 (SD 7.4992), indicating an improvement in overall mental health status from "Good" to "Very good." Both the 4.43-point change in the GPH T score and the 4.48-point change in the GMH T score are considered clinically significant. Table 2 outlines changes in both GPH and GMH T scores.

Table 2. Global physical health and global mental health T scores.

	Before the intervention	After the intervention	Mean change (points)	*P* value
Global physical health			4.43	.001
T score, mean (SD)	42 (6.745)	47.09 (6.655)		
Category	Good	Good		
Global mental health			4.58	.001
T score, mean (SD)	44.47 (8.233)	49.052 (7.4992)		
Category	Good	Very good		

Pearson correlations revealed no significant relationship between age and changes in GPH score (r=0.003; P=.98) or GMH score (r=−0.178; P=.10). Independent-sample 2-tailed t tests showed no significant gender differences in GPH score change for male (mean 2.16, SD 3.08) and female (mean 1.60, SD 2.93) individuals (P=.40). For GMH score change, means were 4.27 (SD 9.32) for male individuals and 4.76 (SD 8.10) for female individuals (P=.80). One-way ANOVA showed significant differences in both GPH score change (P=.009) and GMH score change (P=.001) across diagnostic groups. Post hoc analyses (Tukey honestly significant difference) indicated that participants with head and neck cancers had substantially higher GPH and GMH score changes compared with those in several other groups. The combined model including all 3 predictors

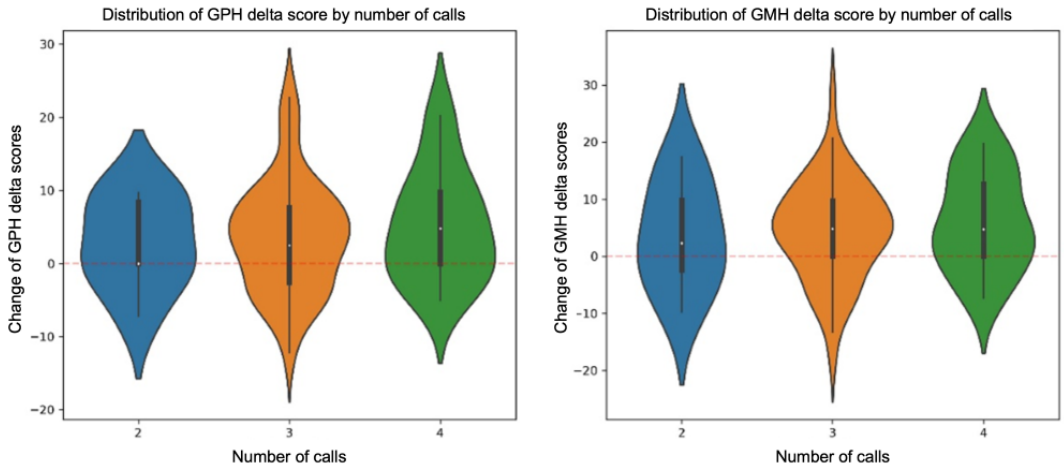
XSL•FO
RenderX

(age, sex, and diagnosis) was assessed using multiple linear regression analyses. Neither the model for GPH score change ($F_{3, 85}$=0.30; P=.83; R^2=0.01) or that for GMH score change ($F_{3, 85}$=1.16; P=.33; R^2=0.04) were significant. None of the individual predictors contributed significantly to either model.

Number of Health Coaching Calls and Changes in GPH and GMH Scores

For the call frequency and app usage analysis, 6 of the 89 participants were excluded because they completed more than 4 calls. These values fell outside the expected intervention range and were considered outliers in program engagement. The final analytic sample for this section, therefore, included 83 participants. Figure 1 illustrates the distribution of changes in physical health (GPH) and mental health (GMH) scores across different numbers of health coaching calls. While statistical analysis using ANOVA (Tukey honestly significant difference) showed no significant differences between call frequencies (GPH: P<.69; GMH: P<.74), the consistently wide range of positive score changes for 3, 4, and 5 calls suggests that patients can experience meaningful improvements regardless of call frequency. The presence of positive median values (shown by the white dots in Figure 1) for most call groups indicates an overall trend toward improvement in both physical and mental health scores, with slightly more improvement in physical and mental health with more calls.

Figure 1. Distribution of global physical health (GPH) and global mental health (GMH) scores according to number of health coaching calls.



Use of the App and Changes in GPH and GMH Scores

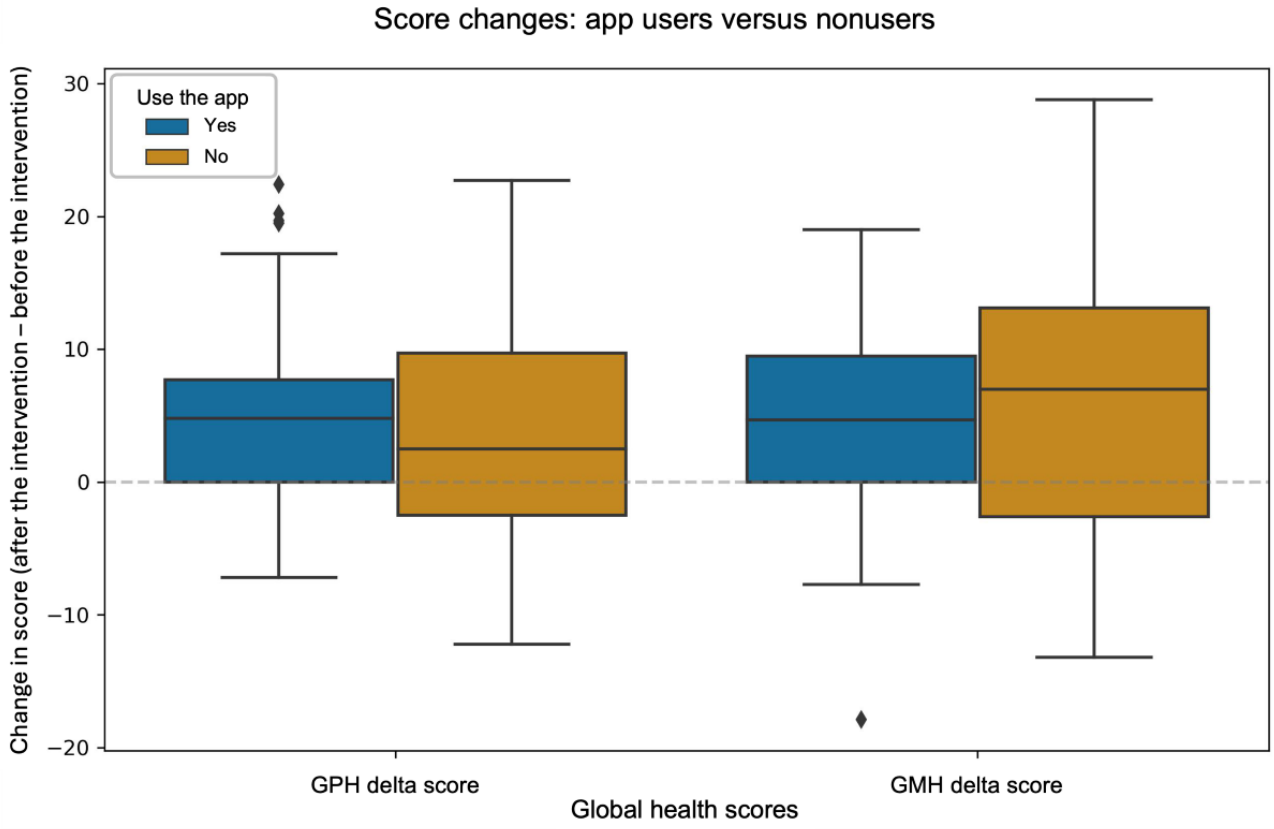
Table 3 outlines the change in GPH and GMH scores between app users (53/83, 64%) and nonusers (30/83, 36%) in the Cancer Coach by CancerAid intervention. Figure 2 shows that both groups exhibited overall positive improvements in their health scores, with median values notably above 0, indicating general health benefits regardless of app use. Non–app users (health coaching calls only) showed slightly higher median improvements in GMH scores 4.74 (SD 10.61) compared to app users (with health coaching calls) 4.63 (SD 7.29); whereas app users (with health coaching calls) had a higher median GPH score change 4.29 (SD 7.15) than participants who only had health coaching calls 3.22 (SD 8.07).

Table 3. Global physical health (GPH) and global mental health (GMH) scores for app users and non–app users.

	GPH point score change	GMH point score change
App users (median)	4.29 (SD 7.15)	4.63 (SD 7.29)
Non–app users (median)	3.22 (SD 8.07)	4.74 (SD 10.61)
t test (*df*)	0.62 (81)	–0.05 (81)
P value	.53	.96

XSL•FO
RenderX

Figure 2. Median and distribution of global physical health (GPH) and global mental health (GMH) scores for health coaching with app use and only use and health coaching calls with no app use.



Discussion

Principal Findings

This study aimed to evaluate the effectiveness of the Cancer Coach by CancerAid program using patient-reported outcomes and how these outcomes were related to the quantity of health coaching calls and mobile app use.

A significant number of participants had a higher GPH score upon completion of the Cancer Coach by CancerAid intervention, and the improvement was deemed clinically significant. GPH includes items of general physical health, ability to carry out daily activities, pain, and fatigue. Many mHealth programs aim to improve physical health with a focus on exercise oncology [35]. However, participants in the Cancer Coach by CancerAid program improved their physical well-being through a more holistic approach to mHealth via support for management of pain and fatigue levels and addressing a range of barriers to participation in everyday activities, including work [21,36]. Studies have shown mixed results in terms of mHealth programs reducing fatigue and pain levels [17,37,38]. The significant change in GPH score in this study could be related to improvements in fatigue and pain levels (see Multimedia Appendix 4 for individual score results in this study). Participating in the Cancer Coach by CancerAid program appears to be related to improved physical health through a personalized goal setting and psychoeducation approach rather than an exercise oncology focus. These findings contribute to the growing body of research supporting the benefits of nonpharmacological interventions to improve physical well-being and fatigue experienced by people with cancer. The magnitude of improvement observed in PROMIS-10 scores in this study is consistent with findings of other digital health coaching interventions. For example, there have been similar gains in PROMIS-10 GPH and GMH scores reported among veterans receiving telephone-based coaching [13]. Chow et al [5] also found improved mental health using a mobile coaching intervention among survivors of cancer. These improvements in PROMIS-10 scores, particularly in domains such as fatigue, emotional well-being, and social roles, have important implications for participants' broader cancer experiences [5]. They may reflect enhanced self-efficacy, functional recovery, and psychosocial adjustment across the different stages of diagnosis, treatment, and survivorship.

A significant number of participants had a higher GMH score upon completion of the Cancer Coach by CancerAid intervention, and the improvement was considered clinically significant. GMH includes items of quality of life, mental health (mood and cognition), social activities and relationships, and emotional well-being (anxiety and depression). Importantly, the average change in GMH T score moved the overall score of the group from a rating of "good" to a rating of "very good" mental health at the end of the program. Although not a counseling service, there are psychosocial aspects of the Cancer Coach by CancerAid intervention that support emotional well-being and improve participants' sense of agency, which was possibly related to the improved mental health scores. These results are similar to those of recent systematic reviews that have shown that mHealth interventions effectively reduce

XSL•FO
RenderX

anxiety and depression in patients with cancer [39] and improve quality of life [16].

Cancer Coach by CancerAid provides holistic care, addressing the physical and emotional support needs specific to the participant regardless of whether they are early or later in their cancer experience. The improved GPH and GMH scores indicate that the program could be providing this. Similarly, a holistic approach, such as health coaching, can also be beneficial in addressing aspects of health and well-being in a collective way [9,40]. A previous study of Cancer Coach by CancerAid also indicated that participation assisted people with cancer in resuming work activities [21]. This holistic approach could also provide a more effective form of health service delivery as it targets a wider group of people with cancer. If a digital and telehealth coaching program can assist with improving pain, fatigue, and mental and physical health, this may also have a cost benefit by reducing health care use, such as general practitioner consultations [41] and hospital visits [42,43].

There was inconclusive evidence as to which mHealth programmatic aspect of the Cancer Coach by CancerAid program contributed to the positive GPH and GMH scores. Participants who only had health coaching calls and no app use had slightly higher GMH scores than those who used both the app and the calls. However, participants who used the app and health coaching calls had a slightly higher GPH score than participants who only had health coaching calls. It is possible that the synchronous telephone health coaching aspect of Cancer Coach by CancerAid had more influence on participants' mental health and well-being than the asynchronous app, but this requires further investigation. Overall, the comparable outcomes between app users and nonusers suggest that patients can benefit from the Cancer Coach by CancerAid program through multiple engagement pathways, allowing for flexibility in how they choose to participate in their care journey. Additionally, the flexibility in call frequency may be beneficial for personalizing support schedules to individual patient needs and circumstances rather than requiring a fixed number of calls for all patients, acknowledging that higher GPH and GMH score changes were experienced by those with more calls, although this was not significant. These findings are similar to those of a systematic review that examined synchronous, asynchronous, and mixed approaches to delivering mHealth interventions to patients with cancer and found that the mode of delivery did not influence outcomes [17]. Understanding the most effective form of delivery is critical for developing evidence-based practice and cost-effective service delivery, so 3-arm randomized controlled trials are recommended [17] to better understand the effectiveness of mHealth programmatic aspects and health outcomes.

Limitations

Pilot studies are often conducted to obtain information in preparation for a larger trial. Therefore, the findings of this study should be interpreted with an understanding of the relatively small sample size and the absence of a control group. While all participants during this study completed the Cancer Coach by CancerAid program, the 55% (107/200) attrition rate refers specifically to noncompletion of the postprogram PROMIS-10 survey. This lack of follow-up data may have biased the study results toward those who experienced greater improvement as those who did not complete the survey may have had less positive outcomes. Therefore, it is possible that a higher survey completion rate could have revealed weaker overall effects.

Because Cancer Coach by CancerAid supports participants at any stage of the cancer journey, certain stages may result in better or worse scores. Although analysis showed no significant differences in age or gender between participants included in the study and those excluded due to missing postintervention PROMIS-10 data, there were notable differences in cancer diagnosis and referral source. A significantly higher proportion of excluded participants had an unknown cancer diagnosis, and the distribution of cancer types varied, potentially impacting the representativeness of the study sample. Additionally, a greater proportion of excluded participants were referred by employers rather than insurers or hospitals, suggesting possible differences in case characteristics or motivations for program enrollment. These factors may limit the generalizability of the findings to all program participants.

While there may be a lack of information about the stage of cancer and participants' details, this is seen as a benefit for our participants, who can be cautious about oversharing their personal and health information. This approach may also increase the number of people willing to enroll in the Cancer Coach by CancerAid program. The participants were 64% (57/89) female, with breast cancer being the dominant diagnosis, which may contribute to gender bias but also reflects men's reluctance to seek psychological or emotional help when dealing with cancer [44]. Additionally, race and ethnicity data were not collected as part of this study. This aligns with CancerAid's model of minimizing data collection to reduce barriers to participation but limits the ability to evaluate cultural or demographic variability in outcomes.

Health coaching is provided to meet individual needs and goals and uses techniques of goal setting and motivational interviewing, among others. All coaches were trained using the same models and methodologies to ensure consistency across techniques used and program reliability. Although the coaches followed consistent methodologies, differences in communication style and participant responsiveness could lead to variability in the results. To reduce potential bias with health coaches administering the PROMIS-10 survey, alternative approaches should be considered. A self-administered digital option within the app could offer greater privacy and autonomy, although it may risk lower completion rates. Another option is to have the survey administered by a neutral third party, which could help preserve objectivity while maintaining engagement.

Further studies on this program will need to focus on more robust data collection on app engagement, comparison with a control group, and extended use of the PROMIS-10 to evaluate sustained changes in functioning. Qualitative studies should consider content analysis of coaching calls to ascertain techniques affecting change and focus groups to gain participants' personal views and experiences of the Cancer Coach by CancerAid program. Given the multiple options in the delivery of mHealth solutions and the variance in reporting

XSL•FO
RenderX

interventions in research, more evidence is required to establish the effectiveness of specific modes of delivery as well.

Conclusions

Participants with various cancer diagnoses showed a significant and positive change in physical and mental health after taking part in the Cancer Coach by CancerAid program as indicated through the PROMIS-10 survey. These preliminary results suggest that clinical practice, including a mixed delivery of synchronous (telehealth coaching) and asynchronous (mobile app) interventions, provides holistic care and may improve physical and mental health and well-being for those with cancer. This type of intervention appears to constitute a potential option for patients with cancer who have unmet needs within traditional health care systems.

Acknowledgments

DS was affiliated with Osara Health at the time of the study and is now affiliated with the University of Bristol.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Ten-item Patient-Reported Outcomes Measurement Information System questionnaire.
[DOCX File , 28 KB - cancer_v12i1e72647_app1.docx]

Multimedia Appendix 2
Screenshots of the CancerAid (now known as Osara Health) app.
[DOCX File , 275 KB - cancer_v12i1e72647_app2.docx]

Multimedia Appendix 3
CancerAid iOS release notes.
[DOCX File , 346 KB - cancer_v12i1e72647_app3.docx]

Multimedia Appendix 4
Individual PROMIS-10 question score results.
[DOCX File , 30 KB - cancer_v12i1e72647_app4.docx]

References

1. Cancer. World Health Organization. URL: https://www.who.int/health-topics/cancer#tab=tab_1 [accessed 2025-05-29]
2. Markovic C, Mackenzie L, Lewis J, Singh M. Working with cancer: a pilot study of work participation among cancer survivors in Western Sydney. Aust Occup Ther J 2020 Dec 13;67(6):592-604. [doi: 10.1111/1440-1630.12685] [Medline: 32657435]
3. Paul K, Buschbacher R. Cancer rehabilitation. Am J Phys Med Rehabil 2011;90(5):S1-S4. [doi: 10.1097/PHM.0b013e31820be4f4]
4. Package of interventions for rehabilitation: module 7: malignant neoplasm. World Health Organization. URL: https://www.who.int/publications/i/item/9789240071254 [accessed 2025-05-29]
5. Chow PI, Drago F, Kennedy EM, Cohn WF. A novel mobile phone app intervention with phone coaching to reduce symptoms of depression in survivors of women's cancer: pre-post pilot study. JMIR Cancer 2020 Feb 06;6(1):e15750 [FREE Full text] [doi: 10.2196/15750] [Medline: 32027314]
6. Kivelä K, Elo S, Kyngäs H, Kääriäinen M. The effects of health coaching on adult patients with chronic diseases: a systematic review. Patient Educ Couns 2014 Nov;97(2):147-157. [doi: 10.1016/j.pec.2014.07.026] [Medline: 25127667]
7. Nguyen J, Maio V, Handley N. Improving cancer care through digital health coaching. Am J Med Qual 2022;37(4):369-370. [doi: 10.1097/JMQ.0000000000000059] [Medline: 35404305]
8. Galantino ML, Schmid P, Milos A, Leonard S, Botis S, Dagan C, et al. Longitudinal benefits of wellness coaching interventions for cancer survivors. Int J Interdiscip Soc Sci 2009;4(10):41-58. [doi: 10.18848/1833-1882/CGP/v04i10/53020]

XSL•FO
RenderX

9. Barakat S, Boehmer K, Abdelrahim M, Ahn S, Al-Khateeb AA, Villalobos N, et al. Does health coaching grow capacity in cancer survivors? A systematic review. Popul Health Manag 2018 Feb;21(1):63-81. [doi: 10.1089/pop.2017.0040] [Medline: 28636526]
10. Conn S, Curtain S. Health coaching as a lifestyle medicine process in primary care. Aust J Gen Pract 2019 Oct 01;48(10):677-680 [FREE Full text] [doi: 10.31128/AJGP-07-19-4984] [Medline: 31569315]
11. Markert C, Sasangohar F, Mortazavi BJ, Fields S. The use of telehealth technology to support health coaching for older adults: literature review. JMIR Hum Factors 2021 Jan 29;8(1):e23796 [FREE Full text] [doi: 10.2196/23796] [Medline: 33512322]
12. Ristevsk E, Trinh T, Vo N, Byrne A, Jamieson P, Greenall A, et al. I.CAN: health coaching provides tailored nutrition and physical activity guidance to people diagnosed with cancer in a rural region in West Gippsland, Australia. J Cancer Surviv 2020 Feb 08;14(1):48-52. [doi: 10.1007/s11764-019-00818-1] [Medline: 31705369]
13. Purcell N, Zamora K, Bertenthal D, Abadjian L, Tighe J, Seal KH. How VA whole health coaching can impact veterans' health and quality of life: a mixed-methods pilot program evaluation. Glob Adv Health Med 2021 Mar 05;10:2164956121998283 [FREE Full text] [doi: 10.1177/2164956121998283] [Medline: 33747639]
14. Financial toxicity (financial distress) and cancer treatment (PDQ®)–patient version. National Cancer Institute. URL: https://www.cancer.gov/about-cancer/managing-care/track-care-costs/financial-toxicity-pdq [accessed 2025-05-29]
15. Chan RJ, Howell D, Lustberg MB, Mustian K, Koczwara B, Ng CC, et al. Advances and future directions in the use of mobile health in supportive cancer care: proceedings of the 2019 MASCC Annual Meeting symposium. Support Care Cancer 2020 Sep 14;28(9):4059-4067. [doi: 10.1007/s00520-020-05513-x] [Medline: 32405966]
16. Sarbaz M, Manouchehri Monazah F, Eslami S, Kimiafar K, Mousavi Baigi SF. Effect of mobile health interventions for side effects management in patients undergoing chemotherapy: a systematic review. Health Policy Technol 2022 Dec;11(4):100680. [doi: 10.1016/j.hlpt.2022.100680]
17. Furness K, Sarkies MN, Huggins CE, Croagh D, Haines TP. Impact of the method of delivering electronic health behavior change interventions in survivors of cancer on engagement, health behaviors, and health outcomes: systematic review and meta-analysis. J Med Internet Res 2020 Jun 23;22(6):e16112 [FREE Full text] [doi: 10.2196/16112] [Medline: 32574147]
18. Wood WA, Bailey C, Castrogivanni B, Mehedint D, Bryant AL, Lavin K, et al. Piloting HealthScore: feasibility and acceptability of a clinically integrated health coaching program for people living with cancer. Cancer Med 2023 Apr 16;12(7):8804-8814 [FREE Full text] [doi: 10.1002/cam4.5625] [Medline: 36647557]
19. Handley NR, Wen K, Gomaa S, Brassil K, Shimada A, Leiby B, et al. A pilot feasibility study of digital health coaching for men with prostate cancer. JCO Oncol Pract 2022 Jul;18(7):e1132-e1140. [doi: 10.1200/OP.21.00712] [Medline: 35394806]
20. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997 Sep;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
21. Lo J, Ballurkar K, Fox S, Tynan K, Luu N, Boyer M, et al. A digital coaching intervention for cancer survivors with job loss: retrospective study. JMIR Cancer 2021 Nov 23;7(4):e31966 [FREE Full text] [doi: 10.2196/31966] [Medline: 34710853]
22. Ahmed O, Lee A, Vachharajani N, Chang SH, Park Y, Khan AS, et al. Reviewing patient-reported outcomes 1 year after orthotopic liver transplantation. J Am Coll Surg 2022 Jul 01;235(1):69-77. [doi: 10.1097/XCS.0000000000000244] [Medline: 35703964]
23. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194 [FREE Full text] [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]
24. Gutierrez C. PROMIS10 global health scale as part of a standard cancer rehabilitation consultation (CRC). Arch Phys Med Rehabil 2015 Oct;96(10):e84. [doi: 10.1016/j.apmr.2015.08.283]
25. Husain M, Phillips A, Tinoco G, Chen JL, Liebner DA. Patient-reported global health predicts adverse health outcomes in a cohort of patients with advanced sarcoma. J Clin Oncol 2020 May 20;38(15_suppl):e23547. [doi: 10.1200/JCO.2020.38.15_suppl.e23547]
26. PROMIS. Northwestern University Health Measures. URL: https://www.healthmeasures.net/ [accessed 2025-05-29]
27. Naughton MJ, Salani R, Lustberg MB, Peng J, Moon J, Loyan H, et al. PROMIS-10 scores at six months post-baseline among breast and gynecologic oncology patients participating in a text-based symptom monitoring program with patient navigation. J Clin Oncol 2020 May 20;38(15_suppl):e19173. [doi: 10.1200/jco.2020.38.15_suppl.e19173]
28. Patton M. Qualitative Research & Evaluation Methods. 4th edition. Thousand Oaks, CA: Sage Publications; 2015.
29. PROMIS - global health 10. Heart Beat Medical. URL: https://heartbeat-med.com/resources/promis-global-health-10/ [accessed 2025-05-29]
30. Lapin B, Katzan IL. PROMIS global health: potential utility as a screener to trigger construct-specific patient-reported outcome measures in clinical care. Qual Life Res 2023 Jan 10;32(1):105-113. [doi: 10.1007/s11136-022-03206-y] [Medline: 35947238]
31. Free online patient report outcome measurement information system (PROMIS 10) score calculator. Ortho ToolKit. URL: https://orthotoolkit.com/promis-10/ [accessed 2025-05-29]

XSL•FO
RenderX

32. Terwee CB, Peipert JD, Chapman R, Lai J, Terluin B, Cella D, et al. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. Qual Life Res 2021 Oct;30(10):2729-2754 [FREE Full text] [doi: 10.1007/s11136-021-02925-y] [Medline: 34247326]
33. Hays RD, Spritzer KL, Thompson WW, Cella D. U.S. general population estimate for "excellent" to "poor" self-rated health item. J Gen Intern Med 2015 Oct 2;30(10):1511-1516 [FREE Full text] [doi: 10.1007/s11606-015-3290-x] [Medline: 25832617]
34. Cohen J. A power primer. Psychol Bull 1992;112(1):155-159. [doi: 10.1037/0033-2909.112.1.155]
35. Stout NL, Baima J, Swisher AK, Winters-Stone KM, Welsh J. A systematic review of exercise systematic reviews in the cancer literature (2005-2017). PM R 2017 Sep 21;9(9S2):S347-S384 [FREE Full text] [doi: 10.1016/j.pmrj.2017.07.074] [Medline: 28942909]
36. Lewis J, Mackenzie L. Cognitive changes after breast cancer: a scoping review to identify problems encountered by women when returning to work. Disabil Rehabil 2022 Sep 11;44(18):5310-5328. [doi: 10.1080/09638288.2021.1919216] [Medline: 33974469]
37. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017 Dec;11(6):704-719 [FREE Full text] [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]
38. Hernandez Silva E, Lawler S, Langbecker D. The effectiveness of mHealth for self-management in improving pain, psychological distress, fatigue, and sleep in cancer survivors: a systematic review. J Cancer Surviv 2019 Feb;13(1):97-107. [doi: 10.1007/s11764-018-0730-8] [Medline: 30635865]
39. Zhang X, Sun S, Jiangenuer L, Zhao P, Lei H, Xu Z, et al. Effect of mobile health (mHealth) on improving anxiety, depression and quality of life in cancer patients: a systematic review and meta -analysis. J Affect Disord 2025 Apr 01;374:11-25. [doi: 10.1016/j.jad.2025.01.016] [Medline: 39793615]
40. Yun YH, Lim CI, Lee ES, Kim YT, Shin KH, Kim Y, et al. Efficacy of health coaching and a web-based program on physical activity, weight, and distress management among cancer survivors: a multi-centered randomised controlled trial. Psychooncology 2020 Jul 23;29(7):1105-1114. [doi: 10.1002/pon.5394] [Medline: 32307828]
41. Heins M, Korevaar J, Rijken P, Schellevis F. For which health problems do cancer survivors visit their General Practitioner? Eur J Cancer 2013 Jan;49(1):211-218. [doi: 10.1016/j.ejca.2012.07.011] [Medline: 22897842]
42. Mausbach BT, Yeung P, Bos T, Irwin SA. Health care costs of depression in patients diagnosed with cancer. Psychooncology 2018 Jul 25;27(7):1735-1741. [doi: 10.1002/pon.4716] [Medline: 29601657]
43. Gray H. Recommendations for an economic evaluation of health coaching. institute for employment studies. 2019. URL: https://www.employment-studies.co.uk/resource/recommendations-economic-evaluation-health-coaching [accessed 2025-05-29]
44. Aranda S. Why men with cancer need improved emotional support throughout their cancer experience. Cancer Council. 2019 Jul 5. URL: https://tinyurl.com/ytfhrwzw [accessed 2026-01-21]

Abbreviations

GMH: global mental health
GPH: global physical health
mHealth: mobile health
PROMIS-10: Patient-Reported Outcomes Measurement Information System 10-item Global Health Short Form

Edited by J Lau; submitted 14.Feb.2025; peer-reviewed by Y Chu, H Derry-Vick; comments to author 02.May.2025; revised version received 16.Dec.2025; accepted 18.Dec.2025; published 01.Apr.2026.

Please cite as:
Lewis J, Wang S, Jue TR, Atkins T, Irwin SA, Nanda P, Sangar D, Mackenzie L, Murali-Ganesh R
Evaluating a Telehealth Coaching and Mobile-Based Digital Engagement Intervention for People With Cancer Using the Patient-Reported Outcomes Measurement Information System Global Health: Pilot Questionnaire Study
JMIR Cancer 2026;12:e72647
URL: https://cancer.jmir.org/2026/1/e72647
doi:10.2196/72647
PMID:41920589

©Joanne Lewis, Shen Wang, Toni Rose Jue, Tim Atkins, Scott A Irwin, Puneet Nanda, Davinder Sangar, Lynette Mackenzie, Raghav Murali-Ganesh. Originally published in JMIR Cancer (https://cancer.jmir.org), 01.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/),

XSL•FO
RenderX

which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Gait Changes After a Mobile Health Exercise Intervention in Older Adults With Myeloid Neoplasms: Single-Arm Pilot Trial

Ying Wang[1], MS, PhD; Marielle Jensen-Battaglia[1], PT, DPT, PhD; Po-Ju Lin[2], RD, MPH, PhD; Ian Kleckner[3], MPH, PhD; Elisabeth Hayward[4], BS; Jason Mendler[5], MD, PhD; Jane Liesveld[5], MD; Marissa LoCastro[6], BS, MD; Cait Fallone Sharma[7], MA; Soroush Mortaz Hedjri[7], MS, MD; Richard Dunne[5], MS, MD; Karen Mustian[2], MA, MPH, PhD; Kah Poh Loh[5], BAO, MBBCh, MS

[1]Department of Public Health Sciences, School of Medicine and Dentistry, University of Rochester, Rochester, NY, United States

[2]Department of Surgery, University of Rochester Medical Center, Rochester, NY, United States

[3]Department of Pain & Translational Symptom Science, School of Nursing, University of Maryland, Baltimore, Baltimore, MD, United States

[4]School of Medicine and Dentistry, University of Rochester, Rochester, NY, United States

[5]Division of Hematology/Oncology, Department of Medicine, James P. Wilmot Cancer Institute, University of Rochester, Rochester, NY, United States

[6]Department of Internal Medicine, University of Wisconsin, Madison, WI, United States

[7]James P. Wilmot Cancer Institute, Rochester, NY, United States

Corresponding Author:
Kah Poh Loh, BAO, MBBCh, MS
Division of Hematology/Oncology, Department of Medicine
James P. Wilmot Cancer Institute
University of Rochester
601 Elmwood Avenue, Box 704
Rochester, NY, 14642
United States
Phone: 1 585 276 4353
Fax: 1 585 273 1042
Email: kahpoh_loh@urmc.rochester.edu

Abstract

Background: Myeloid neoplasms (MNs) are most frequently diagnosed among adults aged 60 years and older. Cancer and chemotherapy can cause gait disturbances and increase fall risk in older adults with MNs. Exercise may improve gait, but there is a lack of research among older adults with MNs undergoing active chemotherapy.

Objective: We explored gait changes following a home-based mobile health exercise intervention during 2 cycles of outpatient chemotherapy (8-12 weeks).

Methods: In a single-arm pilot study, we included adults aged 60 years and older with MNs undergoing chemotherapy. Geriatric Oncology-Exercise for Cancer Patients intervention integrates a progressive aerobic walking and resistance exercise program with a mobile app. We assessed gait by using a waist-worn G-Walk motion sensor during a 6-minute walk at the preintervention and postintervention time points. Spatiotemporal outcomes included cadence (steps per minute), velocity (meters per minute), normalized stride (stride length normalized over height), and swing duration (percentage of the gait cycle during which a foot is in the air when walking). Regularity outcomes that measure gait rhythm included variability of normalized stride and variability of swing duration. Variability for both outcomes was quantified as the SD across all gait cycles. We calculated Cohen *d* effect sizes (ESs) for change in gait outcomes and used the Spearman rank correlation to correlate changes in daily steps and resistance exercise duration with gait outcomes.

Results: We included 13 patients (mean age 71, SD 4.8 years); most were male (n=8, 61.5%), White individuals (n=12, 92.3%), and non-Hispanic individuals (n=13, 100%). Average daily steps were 3084 (SD 1765.5) at the preintervention time point and 3757 (SD 2623.6) at the postintervention time point. Patients performed resistance exercises for 25 minutes per day, 4 days per week at low intensity (mean rating of perceived exertion 3/10, SD 1.3). At the postintervention time point, we observed numerical changes in gait outcomes, including increased cadence (mean +4.6, SD 14.6 steps per minute; *P*=.24; ES=0.38) and decreased variability in normalized stride (mean −1.4%, SD 8.5%; *P*=.34; ES=−0.18) and swing duration (mean −0.1%, SD 1.1%; *P*=.54; ES=−0.15), although these improvements were not statistically significant. Increased daily steps significantly correlated with

XSL•FO
RenderX

decreased swing duration variability (r=−0.72; P=.01). Resistance exercise duration significantly correlated with increased cadence (r=0.54; P=.06) and velocity (r=0.56; P=.05).

Conclusions: In our exploratory analyses, better adherence to exercise correlated with improved gait outcomes. Our ongoing pilot randomized controlled trial (ClinicalTrials.gov identifier: NCT04981821) will further examine the effects of the Geriatric Oncology-Exercise for Cancer Patients intervention on gait outcomes in this population.

(JMIR Cancer 2026;12:e80909) doi:10.2196/80909

KEYWORDS

exercise intervention; gait; mobile health; older adults; myeloid neoplasms

Introduction

Gait is an important indicator of overall health and functional status [1-3]. Although walking speed is a well-recognized marker of many health outcomes, it arises from more granular aspects of gait, including both spatiotemporal measures (eg, gait speed, stride length, cadence, and symmetry) and regularity measures (eg, variability of stride length) [1,4,5]. Cancer and its treatments can contribute to disturbances in these more detailed aspects of gait long before measurable loss in gait speed or function [6]. Possible pathways include musculoskeletal effects, including fatigue and sarcopenia, which influence body movement, and neurological effects, including direct tumor infiltration or compression and chemotherapy-induced neurotoxicity [6]. For example, chemotherapy-induced peripheral neuropathy in older adults with cancer is commonly associated with shorter step length, loss of symmetry or synchrony, and greater step width variability in addition to slower gait speed [3,7]. Gait disturbances have been linked to poor clinical outcomes, including falls, unplanned hospitalizations, emergency department visits, and mortality [2,3]. While habitual physical activity may mitigate some adverse effects of cancer and its treatments, gait disturbances still occur in older adults with cancer [8]. Therefore, there is a need for interventions that can help maintain normal gait in this population.

Myeloid neoplasms (MNs; eg, acute myeloid leukemia [AML] and myelodysplastic syndrome [MDS]) are most frequently diagnosed among adults aged 60 years and older [9]. Older adults with MNs are vulnerable to gait disturbances due to the combined effects of aging, cancer, and treatment-induced toxicities [3,7,10]. Despite the growing number of studies on exercise interventions in older adults with cancer, the impact of specific exercise programs on gait outcomes in older adults with MNs undergoing active chemotherapy is understudied. We previously established the feasibility and usability of a single-arm mobile health exercise program for older adults with MNs [11]. Using data from this pilot study, we aimed to explore (1) changes in gait outcomes following exercise and (2) correlations of exercise adherence with changes in gait outcomes.

Methods

Study Design, Setting, and Participants

We conducted a single-arm pilot study at the University of Rochester Medical Center (ClinicalTrials.gov identifier: NCT04035499) to evaluate the feasibility of a mobile health exercise intervention, Geriatric Oncology-Exercise for Cancer Patients (GO-EXCAP), for older adults with MNs. Patients were screened and recruited between February 2020 and July 2021. Eligible patients were approached after confirming with their treating oncology team. Eligibility criteria were (1) age of 60 years and older; (2) being English speakers; (3) diagnosis of an MN, such as AML, MDS, or MDS and myeloproliferative neoplasm overlap; (4) undergoing outpatient chemotherapy; (5) a physician-verified Eastern Cooperative Oncology Group (ECOG) performance status of 0 to 2; (6) ability to walk 4 m; and (7) capacity to provide informed consent. Individuals were excluded from the study if they (1) had medical contraindications to exercise as determined by their treating oncologists, (2) had a low platelet count on the most recent complete blood count (ie, ≤10,000 platelets per microliter), and (3) did not receive platelet transfusion (eg, due to platelet refractoriness). Findings from the primary study have been previously reported [11].

This study is a secondary analysis of data from the parent study [11]. We report this analysis following the TREND (Transparent Reporting of Evaluations With Nonrandomized Designs) guidelines for nonrandomized evaluations of behavioral and public health interventions (Table S1 in Multimedia Appendix 1).

Intervention and Study Procedures

GO-EXCAP is a mobile health exercise intervention developed to promote successful behavior change by using evidence-based approaches to improve exercise self-efficacy (ie, verbal persuasion, social modeling, imaginal experiences, and provision of regular monitoring and feedback on mastery as well as emotional and physiological states).

GO-EXCAP integrates the registered and copyrighted home-based exercise intervention Exercise for Cancer Patients (EXCAP) with the GO-EXCAP mobile app. The EXCAP intervention was designed and delivered by American College of Sports Medicine–certified clinical exercise physiologists and has been tested in multiple studies [12-14]. It consists of progressive aerobic walking and resistance band exercise prescribed by a clinical exercise physiologist during a one-on-one in-person teaching session. The exercise physiologist generated individually tailored exercise prescriptions for each participant, increasing their baseline daily steps weekly by 5% to 20%. Between 5% and 20% weekly increase in steps would allow individuals classified as sedentary (ie, <5000 steps) to progress toward or achieve an active classification (ie, ≥10,000

XSL•FO
RenderX

steps) over an 8- to 12-week intervention period. The 5% to 20% step progression rate was individualized based on three key criteria: (1) average daily step counts at baseline, (2) participants' physical capability and perceived readiness, and (3) weekly performance during the intervention. The strategy used to determine the weekly progression rate was described in Multimedia Appendix 2. For resistance band exercise, participants also received individualized prescriptions based on their baseline fitness level. The protocol began with 1 set of 8 to 15 repetitions for 16 exercises, using color-coded resistance bands appropriate for individual strength. Under the instruction of exercise physiologists, participants progressed weekly by increasing sets and repetitions until reaching 4 sets of 15 repetitions. Upon reaching this milestone, participants transitioned to a heavier resistance band and repeated the progression. Throughout the intervention, participants were instructed to perform both aerobic walking and resistance band exercises at a low to moderate intensity, corresponding to a rating of perceived exertion (RPE) of 3 to 5 on a scale from 1 to 10 (1="not tired at all"; 10="extremely tired") [15]. We provided participants with a GO-EXCAP kit containing a printed manual, an activity tracker to record daily steps (Garmin Forerunner 35), 3 therapeutic bands with resistance levels corresponding to 2 to 14 pounds, and a tablet preloaded with the GO-EXCAP app. The exercise physiologist monitored participant-entered exercise data on the app and adjusted prescriptions as needed.

The EXCAP intervention lasted 6 weeks in prior trials [12-14]. Because chemotherapy for MNs is typically 4 weeks per cycle, the intervention duration was extended to align with 2 chemotherapy cycles. Given that chemotherapy delays are common in clinical practice, the GO-EXCAP intervention lasted 8 to 12 weeks, or 2 chemotherapy cycles.

Measures

Demographic and Clinical Characteristics

Demographic and clinical characteristics collected at baseline included age, sex, race, ethnicity, marital status, educational attainment, ECOG performance status, cancer type, and chemotherapy cycle at the initiation of the intervention. The ECOG performance status is a validated and widely used scale to assess a patient's ability to perform daily activities, with scores ranging from 0 (fully active) to 5 (deceased). ECOG scores are associated with various health outcomes in patients with cancer, including emergency department visits, hospitalizations, and mortality [16].

Exercise Adherence

During the intervention, daily exercise data collected through the GO-EXCAP app included the number of steps from the activity tracker and self-reported minutes of resistance band exercise and RPE on a scale from 1 to 10 for EXCAP exercises [17]. These metrics provide a comprehensive representation of exercise adherence across exercise type, frequency, duration, and intensity.

Gait Outcomes

Participants wore a G-Walk motion sensor (BTS Bioengineering) around the waist while performing the 6-minute walk at baseline and the postintervention time point. All assessments were administered and supervised by clinical exercise physiologists. The exercise physiologists instructed participants to walk as fast as they could on level ground for up to 6 minutes. The G-Walk motion sensor is a noninvasive, lightweight, waist-worn device equipped with a triaxial accelerometer, gyroscope, magnetometer, and sensor fusion. It collects gait data and transmits them to the G-Studio software (BTS Bioengineering) through a Bluetooth connection for functional evaluations of walking. Prior to performing any assessments, all clinical exercise physiologists completed study-specific training and followed standardized operating procedures to ensure consistency and adherence to the protocol. All assessments were periodically audited by a second trained clinical exercise physiologist to ensure procedural fidelity and interrater reliability across exercise physiology personnel.

Figure 1 describes a gait cycle consisting of stance and swing phases. We measured both the spatiotemporality and regularity of gait (Table 1) [18]. The spatiotemporal outcomes included walking distance covered in 6 minutes (meters), cadence (steps per minute), velocity (meters per minute), stride length (meters), normalized stride (percentage of height), general symmetry index (GSI), stance duration (percentage of gait cycle), and swing duration (percentage of gait cycle). Regularity outcomes that measured gait rhythm included variability of normalized stride (percentage of height) and variability of swing duration (percentage of gait cycle).

XSL•FO
RenderX

Figure 1. An overview of the walking gait cycle (with the right leg as the reference leg).

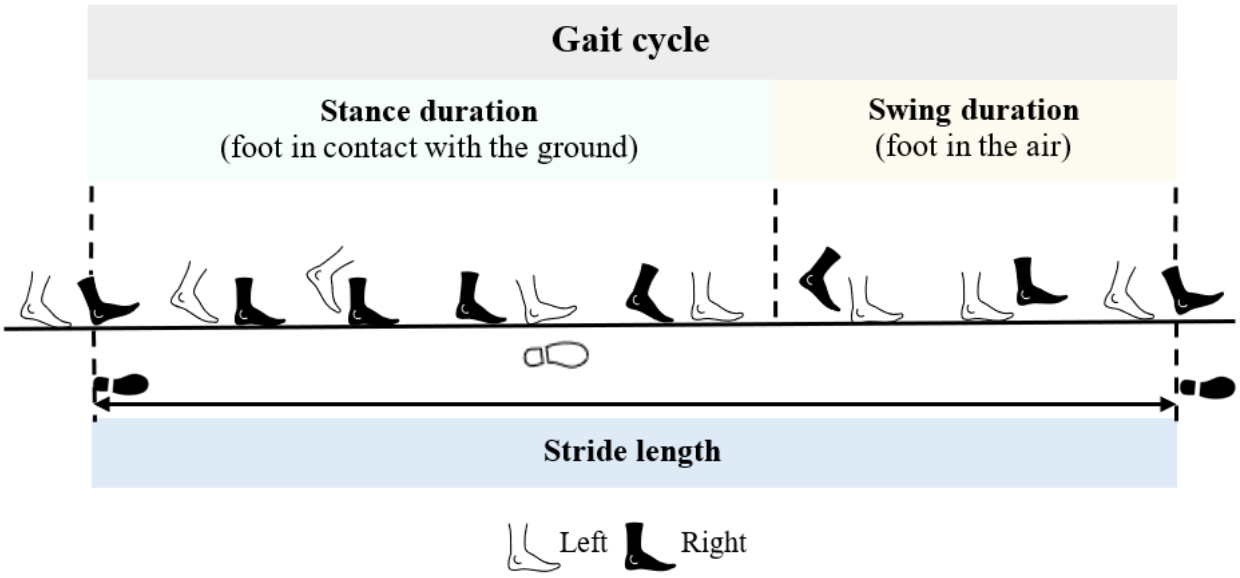


Table 1. Gait outcomes measured during the 6-minute walk test[a].

Dimension and outcome	Definition
Spatiotemporality	
Walking distance	The distance covered in 6 min (meters); higher is better
Cadence	The number of steps per minute; higher is better
Velocity	The distance covered per minute (meters per min); higher is better
Stride length	The distance covered from the moment one foot hits the ground to the next time that same foot touches the ground again (meters), equivalent to the length of a full gait cycle; higher is better
Normalized stride	Normalized over height: calculated as normalized stride length = stride length (meters)/height (meters); higher is better
General symmetry index	An index ranging from 0% to 100% indicating symmetry between the right and left gait cycles based on the comparison of the antero-posterior acceleration component; higher is better
Stance duration	The percentage of time for which a foot remains in contact with the ground (optimal range: 60%-62%); bilateral stance duration is calculated using the average of the left and right stance durations
Swing duration	The percentage of time for which a foot is in the air during walking (optimal range: 38%-40%); bilateral swing duration is calculated using the average of the left and right swing durations
Regularity	
Variability of normalized stride	The SD of stride length across all gait cycles; lower is better
Variability of swing duration	The SD of swing duration across all gait cycles; lower is better

[a]The variability of stance and swing duration is identical due to the statistical relationship between them (stance + swing duration = 100%). Therefore, we reported only the variability of swing duration in this study.

Statistical Analysis

Participants with gait outcomes available at both baseline and the postintervention time point were included in the analyses. To evaluate potential selection bias, we compared baseline characteristics between participants included in this analysis and those who were not. For participants included in this analysis, we described their characteristics, exercise adherence, and gait outcomes. Depending on the data distribution, we used paired 2-tailed *t* tests or Wilcoxon signed-rank tests to evaluate within-subject changes in each gait outcome from baseline to the postintervention time point. We also calculated Cohen *d* effect sizes (ESs) to quantify changes from baseline to the postintervention time point (small effect: 0.20 to <0.50; moderate effect: ≥0.50 to <0.80; large effect: ≥0.80) [19]. Spearman correlation analyses were conducted to assess the correlations of changes in daily steps and the duration of resistance band exercise with changes in gait outcomes. Individual data were presented in scatterplots to detect potential influential observations. In sensitivity analysis, we repeated the Spearman correlation analyses after excluding one observation

XSL•FO
RenderX

that was more than 1.5 IQRs below the first quartile and did not follow the general trend of the rest of the data.

Given the exploratory nature of this analysis and the small sample size, no imputation methods were applied. A complete-case approach was used for all analyses. We performed hypothesis testing at an α value of .10 (2-tailed). All analyses were conducted using SAS (version 9.4; SAS Institute).

Ethical Considerations

The parent study was approved by the Research Subjects Review Board of the University of Rochester (study ID: STUDY00003945), and all participants provided informed consent. To ensure participant privacy and confidentiality, all data were deidentified and managed via REDCap (Research Electronic Data Capture; Vanderbilt University), a secure web application accessible only to authorized research personnel. Participants received US $50 in compensation after completing all baseline and postintervention activities. The original ethics approval covers secondary analysis, and thus, no additional consent was required for this study.

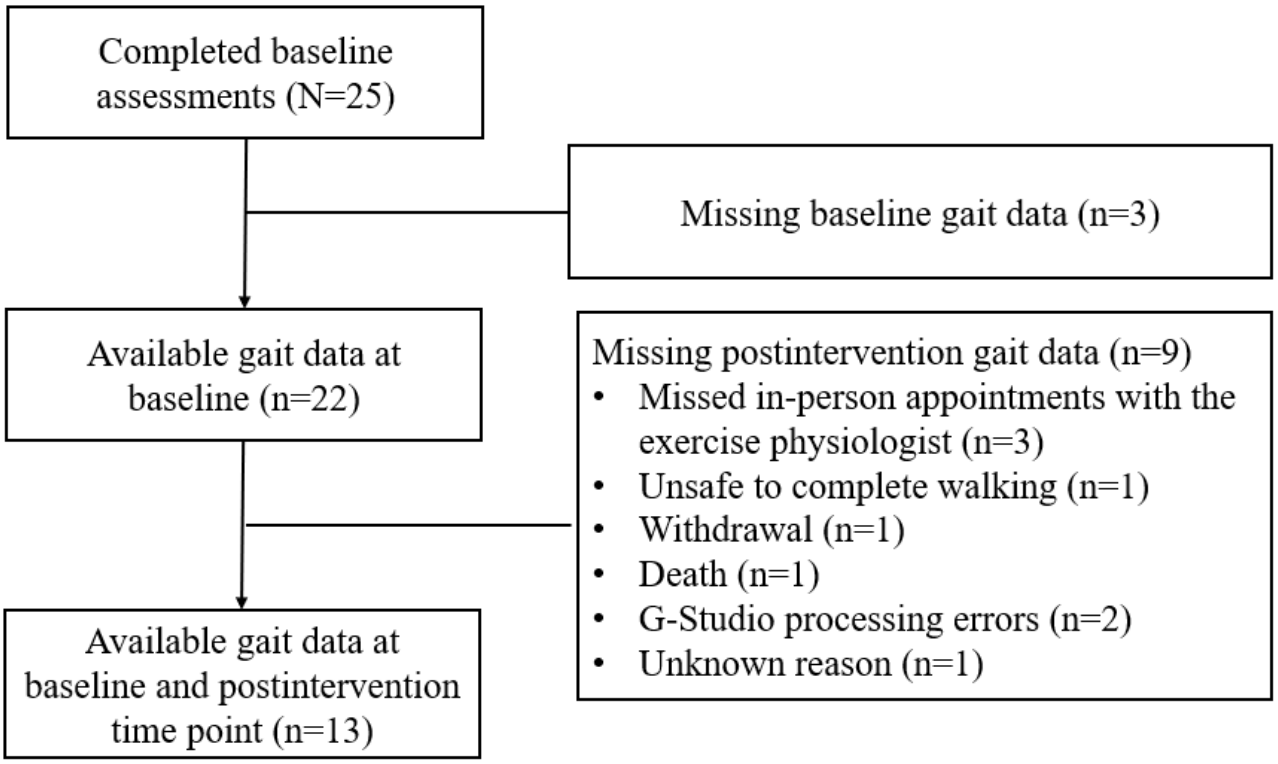
Results

Participant Characteristics and Exercise Adherence

We included 13 participants in this analysis (Figure 2). Of the 25 participants enrolled, 3 (12%) had missing baseline gait data due to logistical issues (eg, Bluetooth connection error). Of the 22 participants with available baseline gait data, 13 (59.1%) completed postintervention gait measures. Reasons for missing postintervention gait data included missed in-person appointments with the exercise physiologist due to health decline (3/22, 13.6%), being unsafe to complete walking (1/22, 4.5%), withdrawal (1/22, 4.5%), death (1/22, 4.5%), G-Studio processing errors (2/22, 9.1%), and unknown reasons (1/22, 4.5%). Comparison of baseline characteristics between participants included in the analysis and those who were not did not reveal substantial differences (Table S1 in Multimedia Appendix 3).

Table 2 describes the characteristics and exercise adherence of the 13 participants with both baseline and postintervention gait data. These participants had a mean age of 71 (SD 4.8) years, and most were White (n=12, 92.3%) and non-Hispanic (n=13, 100%) individuals. More than half were male (n=8, 61.5%), married (n=9, 69.2%), and college graduates or of higher educational levels (n=7, 53.8%). Most participants were diagnosed with AML (n=7, 53.8%) or MDS (n=5, 38.5%). These 13 participants had a mean of 3084 (SD 1765.5) daily steps at baseline and 3757 (SD 2623.6) daily steps at the postintervention time point. On average, they performed resistance band exercises for 4 days per week, 25 minutes per day, with a mean RPE of 3/10 (SD 1.3). No adverse events were reported related to the intervention [11]. Overall, this intervention demonstrated excellent tolerability among older adults with MNs.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.



XSL•FO
RenderX

Table 2. Demographic and clinical characteristics (n=13).

Variable	Values
Age (y), mean (SD; range)	71 (4.8; 65 to 80)
Sex, n (%)	
Male	8 (61.5)
Female	5 (38.5)
Race, n (%)	
Black or African American	1 (7.7)
White	12 (92.3)
Ethnicity (not Hispanic or Latino), n (%)	13 (100)
Marital status, n (%)	
Married	9 (69.2)
Divorced or widowed	1 (7.7)
Single	3 (23.1)
Educational level, n (%)	
High school or lower	1 (7.7)
Some college or post–high school training	4 (30.8)
College graduate or higher	7 (53.8)
Other or preferred not to say	1 (7.7)
ECOG[a] performance status, n (%)	
0	1 (7.7)
1	8 (61.5)
2	4 (30.8)
Diagnosis, n (%)	
Acute myeloid leukemia	7 (53.8)
Chronic myelomonocytic leukemia	1 (7.7)
Myelodysplastic syndrome	5 (38.5)
Chemotherapy cycle at initiation of the intervention, n (%)	
1	3 (23.1)
2	4 (30.8)
3	3 (23.1)
4	1 (7.7)
>5	2 (15.4)
Daily steps at baseline, mean (SD; range)	3084 (1765.5; 649.8 to 6546.4)
Daily steps at the postintervention time point, mean (SD; range)	3757 (2623.6; 307.3 to 10,554.9)
Change in daily steps, mean (SD; range)	673 (1270.1; −1099.8 to 4008.4)
Frequency of resistance exercise (d per wk), mean (SD; range)	4 (2.2; 0.1 to 7.0)
Bout of resistance exercise (min), mean (SD; range)	25 (10.8; 8.3 to 46.1)
RPE[b] during resistance exercise, mean (SD; range)	3 (1.3; 0.8 to 5.6)

[a]ECOG: Eastern Cooperative Oncology Group.

[b]RPE: rating of perceived exertion (1=“not tired at all”; 10=“extremely tired”).

XSL•FO
RenderX

Change in Gait Outcomes From Baseline to the Postintervention Time Point

Table 3 describes gait outcomes at baseline and the postintervention time point and changes in these measures. At baseline, participants walked an average of 408.0 (SD 128.8) m in the 6-minute walk test, with a cadence of 110.9 (SD 12.4) steps per minute and a velocity of 82.0 (SD 21.8) meters per minute. The mean GSI was 95.1% (SD 3%). Stance duration accounted for an average of 59.9% (SD 1.9%) of the gait cycle, with a mean stance duration variability of 0.8% (SD 0.9%).

From baseline to the postintervention time point, we observed a numerical increase in average cadence (mean +4.6, SD 14.6 steps per minute; *P*=.24; ES=0.38) and velocity (mean +1.6, SD 8.2 meters per minute; *P*=.24; ES=0.07). We also observed a numerical increase in stance duration (mean +0.4%, SD 2.3%; *P*=.55; ES=0.20) and a numerical decrease in swing duration (mean −0.4%, SD 2.3%; *P*=.55; ES=0.20), which together brought stance and swing durations closer to their optimal ranges (Table 3). In terms of regularity outcomes, we observed a numerical decrease in variability for both normalized stride (mean −1.4%, SD 8.5%; *P*=.34; ES=−0.18) and swing duration (mean −0.1%, SD 1.1%; *P*=.54; ES=−0.15). None of these outcomes were statistically significant.

Table 3. Change in gait outcomes from baseline to the postintervention time point (n=13).

	Baseline		Postintervention time point		Change from baseline to postintervention time point			
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	*P* value	Effect size
Spatiotemporal outcomes								
6-min walking distance (meters)	408.0 (128.8)	434.1 (362.7 to 485.3)	411.9 (121.7)	449.6 (287.7 to 485.5)	+3.9 (74.5)	+15.5 (−14.7-28.2)	.85	0.03
Cadence (steps per min)	110.9 (12.4)	113.0 (110.2 to 115.3)	115.5 (14.1)	115.5 (110.4 to 125.9)	+4.6 (14.6)	+2.0 (−0.8 to 5.0)	.24	0.38
Velocity (meters per min)	82.0 (21.8)	84.6 (68.8 to 94.6)	83.5 (24.7)	86.4 (76.3 to 97.0)	+1.6 (8.2)	+3.7 (−1.6 to 6.1)	.24	0.07
Stride length (meters)	1.5 (0.3)	1.6 (1.2 to 1.7)	1.4 (0.3)	1.5 (1.3 to 1.7)	−0.1 (0.2)	−0.0 (−0.1 to 0.0)	.54	−0.16
Stride length (% of height)	89.1 (19.2)	93.4 (73.6 to 102.0)	86.0 (19.2)	90.4 (73.6 to 98.2)	−3.1 (10.8)	−0.1 (−5.1 to 1.2)	.54	−0.16
General symmetry index (%)	95.1 (3)	95.7 (92.9 to 98.0)	95.1 (2.4)	95.8 (94.2 to 96.4)	−0.0 (2.1)	−0.8 (−1.2 to 1.8)	.99	−0.00
Swing duration (% of cycle)	40.1 (1.9)	40.2 (39.4 to 41.5)	39.7 (2.5)	40.2 (37.1 to 41.6)	−0.4 (2.3)	−0.2 (−1.2 to 0.9)	.55	−0.20
Stance duration (% of cycle)	59.9 (1.9)	59.8 (58.5 to 60.6)	60.3 (2.5)	59.8 (58.4 to 62.9)	+0.4 (2.3)	+0.2 (−0.9 to 1.2)	.55	0.20
Regularity outcomes								
Stride length variability (% of height)	5.9 (7.6)	4.1 (2.6 to 5.5)	4.5 (2.8)	3.3 (3.0 to 4.5)	−1.4 (8.5)	+0.8 (0.2 to 1.1)	.34	−0.18
Swing duration variability (% of cycle)	0.8 (0.9)	0.5 (0.4 to 0.8)	0.7 (0.6)	0.5 (0.4 to 0.6)	−0.1 (1.1)	−0.1 (−0.3 to 0.2)	.54	−0.15

Correlations Between Both Change in Daily Steps and Resistance Exercise Duration and Changes in Gait Outcomes

Table 4 presents the correlations between changes in daily steps and duration of resistance band exercise with changes in gait outcomes. Increased daily steps significantly correlated with decreased swing duration variability (*r*=−0.72; *P*=.01). Duration of resistance band exercise significantly correlated with increased distance covered in the 6-minute walk test (*r*=0.62; *P*=.03), cadence (*r*=0.54; *P*=.06), velocity (*r*=0.56; *P*=.05), and stance duration (*r*=0.58; *P*=.04). We did not observe significant correlations for other gait outcomes.

Scatterplots showing the relationship between changes in gait outcomes and changes in exercise are shown in Figures S1 and S2 in Multimedia Appendix 3. In sensitivity analysis excluding one potential influential observation (Table S2 in Multimedia Appendix 3), the correlations remained largely consistent. Increased daily steps continued to be correlated with reduced swing duration variability (*r*=−0.64; *P*=.02). Similarly, longer resistance band exercise duration was correlated with improved walking distance (*r*=0.59; *P*=.04) and velocity (*r*=0.54; *P*=.07). Correlations of resistance band exercise duration with cadence (*r*=0.45; *P*=.14) and stance duration (*r*=0.50; *P*=.10) remained in the same direction.

XSL•FO
RenderX

Table 4. Correlations between change in gait outcomes and change in exercise (n=13).

Gait parameter	Change in daily steps		Duration of resistance training (min)	
	r	*P* value	*r*	*P* value
Change in spatiotemporal outcomes				
6-min walking distance (meters)	0.16	.59	0.62	.03
Cadence (steps per min)	0.20	.51	0.54	.06
Velocity (meters per min)	0.24	.43	0.56	.05
Stride length (meters)	0.05	.87	0.03	.93
Stride length (% of height)	0.05	.87	0.03	.93
General symmetry index (%)	0.10	.75	0.19	.54
Swing duration (% of cycle)	0.26	.38	−0.58	.04
Stance duration (% of cycle)	−0.26	.38	0.58	.04
Change in regularity outcomes				
Stride length variability (% of height)	−0.24	.43	−0.15	.62
Swing duration variability (% of cycle)	−0.72	.01	0.10	.75

Discussion

In this pilot mobile health exercise study for older adults with MNs, we observed numerical changes in several gait outcomes, including cadence (mean +4.6, SD 14.6 steps per minute; *P*=.24; ES=0.38), stance duration (mean +0.4%, SD 2.3%; *P*=.55; ES=0.20), swing duration (mean −0.4%, SD 2.3%; *P*=.55; ES=0.20), and variability in normalized stride (mean −1.4%, SD 8.5%; *P*=.34; ES=−0.18) and swing duration (mean −0.1%, SD 1.1%; *P*=.54; ES=−0.15). Better intervention adherence significantly correlated with improved gait outcomes. These findings suggest that adherence to the GO-EXCAP mobile health exercise intervention shows promise in improving gait in this population.

To our knowledge, no studies have evaluated gait outcomes following exercise interventions in older adults with MNs. Therefore, we compare our results to those of previous research on older adults with other types of cancer. In a 12-week program incorporating electromyostimulation, light-intensity exercise, and dietary support for patients with advanced-stage solid cancers (mean age 62.4 years), participants in the intervention arm (n=26) showed improvements in mean cadence (+4.2 steps per minute) and stance duration (0.5% closer to the optimal range) [20]. In another 8-week home-based resistance training for older patients with advanced cancer (n=30; median age 75 years), mean gait velocity increased by 2.4 meters per minute among those with ≥80% intervention compliance [21]. While these interventions differ from ours in exercise type and duration, the favorable changes in gait outcomes observed are comparable. Taken together, these studies suggest that exercise interventions, including GO-EXCAP, have the potential to address spatiotemporal outcomes in older adults with cancer and need to be studied further.

In addition to commonly analyzed spatiotemporal outcomes, our study expands the literature by evaluating gait symmetry in older adults with MNs. Gait asymmetries are clinically relevant as they have been linked to a high risk of fall [22,23]. There are several likely explanations for the lack of significant change in symmetry in our sample. First, the GSI reported in our study (95.1%) is slightly higher than that reported for healthy adults aged 61 years and older (92%-94%) [24]. Although there is a lack of normative data for this measure, this value suggests that our study sample may have preserved physical function, reflected in a relatively good gait symmetry at baseline, leaving limited room for improvement. In addition, symmetry can be assessed in different movement directions, such as anterior or posterior (evaluated in our study) and medial or lateral. Several studies have linked pathological gait patterns to a higher risk of falls in the medial or lateral direction, suggesting that non–anterior or posterior assessments may provide more valuable insights into gait symmetry [4]. Future studies are warranted to determine which gait symmetry measures are most sensitive in detecting clinically relevant pathological gait outcomes.

Adherence to GO-EXCAP shows promise in improving the variability of stride and swing duration, which quantifies the automaticity and regularity of gait, among older patients with MNs. Increased gait variability is associated with a higher risk of falls, frailty, and incident mobility disability [25-27]. Notably, gait variability may be a more sensitive marker than spatiotemporal outcomes (eg, gait velocity) for detecting pathological gait in older adults with cancer who report peripheral neuropathy [7]. Despite its predictive values, we found only one previous study that evaluated gait variability following an exercise intervention, and the authors reported no significant changes [20]. Thus, our study expands the literature by demonstrating a statistically significant correlation between adherence to the GO-EXCAP intervention and improved gait regularity in older adults with MNs. Increased exercise, including progressive aerobic walking, may improve gait regularity through enhancing motor control and gait automaticity [28]. More studies are needed to better understand these underlying mechanisms.

XSL•FO
RenderX

The strengths of this study include the use of a user-friendly and noninvasive motion sensor to measure gait outcomes, which allowed for efficient gait dataset management, and the comprehensive measurement of gait by including both spatiotemporal and regularity outcomes. Several limitations should also be noted. First, given the pilot nature of this study, the correlation analyses were exploratory and hypothesis generating. Second, the single-arm design without a control group prevents us from determining whether observed gait improvements are attributed directly to the intervention. Third, the small sample size limited our ability to adjust for potential confounders or provide CI estimates for correlation analysis. Additionally, to our knowledge, established minimal clinically important difference thresholds for gait outcomes are not available for general older adults or for older adults with cancer. Therefore, it remains unclear whether the observed improvements in gait outcomes reached a clinically meaningful threshold. Finally, there is no consensus on clinically meaningful differences in gait spatiotemporality and regularity outcomes. Therefore, our correlation analysis between change in gait outcomes and change in exercise should be interpreted as exploratory and is primarily useful to guide future larger analyses of similar outcomes in this population.

In conclusion, we found that greater exercise adherence to GO-EXCAP significantly correlated with improved gait outcomes, suggesting that the combination of aerobic and resistance training has the potential to improve gait quality in inactive older adults with cancer. Our ongoing pilot randomized controlled trial (ClinicalTrials.gov identifier: NCT04981821) will further examine the effects of the GO-EXCAP intervention on gait outcomes in this population and assess whether improvements in gait are associated with better functional outcomes, including reduced falls.

Acknowledgments

The authors would like to thank Susan Rosenthal, MD, for her editorial assistance.

Funding

This work was supported by National Institutes of Health National Cancer Institute grants K99CA237744 and R00CA237744 to KPL, UG1CA189961 to GRM and KMM, and K07CA22193 to IK.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
TREND checklist.
[PDF File (Adobe PDF File), 686 KB - cancer_v12i1e80909_app1.pdf]

Multimedia Appendix 2
Strategy used to determine the weekly progression rate.
[DOCX File , 17 KB - cancer_v12i1e80909_app2.docx]

Multimedia Appendix 3
Baseline characteristic comparisons between participants, scatterplots showing the relationship between changes in gait outcomes and changes in exercise, and sensitivity analysis.
[DOCX File , 1851 KB - cancer_v12i1e80909_app3.docx]

References

1. Hollman JH, McDade EM, Petersen RC. Normative spatiotemporal gait parameters in older adults. Gait Posture 2011 May;34(1):111-118 [FREE Full text] [doi: 10.1016/j.gaitpost.2011.03.024] [Medline: 21531139]
2. Liu MA, DuMontier C, Murillo A, Hshieh TT, Bean JF, Soiffer RJ, et al. Gait speed, grip strength, and clinical outcomes in older patients with hematologic malignancies. Blood 2019 Jul 25;134(4):374-382 [FREE Full text] [doi: 10.1182/blood.2019000758] [Medline: 31167800]
3. Marshall TF, Zipp GP, Battaglia F, Moss R, Bryan S. Chemotherapy-induced-peripheral neuropathy, gait and fall risk in older adults following cancer treatment. J Cancer Res Pract 2017 Dec;4(4):134-138. [doi: 10.1016/j.jcrpr.2017.03.005]
4. Viteckova S, Kutilek P, Svoboda Z, Krupicka R, Kauler J, Szabo Z. Gait symmetry measures: a review of current and prospective methods. Biomed Signal Process Control 2018 Apr;42:89-100. [doi: 10.1016/j.bspc.2018.01.013]

XSL•FO
RenderX

5. Hulleck AA, Menoth Mohan D, Abdallah N, El Rich M, Khalaf K. Present and future of gait assessment in clinical practice: towards the application of novel trends and technologies. Front Med Technol 2022 Dec 16;4:901331 [FREE Full text] [doi: 10.3389/fmedt.2022.901331] [Medline: 36590154]
6. Hojan K. Gait and balance disorders in cancer patients. Pol Orthop Traumatol 2012 Sep 21;77:73-76. [Medline: 23306290]
7. Hsieh KL, Trinh L, Sosnoff JJ. Gait variability is altered in cancer survivors with self-reported neuropathy. Gait Posture 2019 Jul;72:206-210. [doi: 10.1016/j.gaitpost.2019.06.014] [Medline: 31254773]
8. Winters-Stone KM, Horak F, Jacobs PG, Trubowitz P, Dieckmann NF, Stoyles S, et al. Falls, functioning, and disability among women with persistent symptoms of chemotherapy-induced peripheral neuropathy. J Clin Oncol 2017 Aug 10;35(23):2604-2612 [FREE Full text] [doi: 10.1200/JCO.2016.71.3552] [Medline: 28586243]
9. Cancer stat facts: leukemia — acute myeloid leukemia (AML). National Cancer Institute. URL: https://seer.cancer.gov/statfacts/html/amyl.html [accessed 2026-04-06]
10. Nonnekes J, Post E, Imbalzano G, Bloem BR. Gait changes with aging: an early warning sign for underlying pathology. J Neurol 2025 Mar 07;272(4):257. [doi: 10.1007/s00415-025-12995-4] [Medline: 40053183]
11. Loh KP, Sanapala C, Watson EE, Jensen-Battaglia M, Janelsins MC, Klepin HD, et al. A single-arm pilot study of a mobile health exercise intervention (GO-EXCAP) in older patients with myeloid neoplasms. Blood Adv 2022 Jul 12;6(13):3850-3860 [FREE Full text] [doi: 10.1182/bloodadvances.2022007056] [Medline: 35320340]
12. Kamen C, Heckler C, Janelsins MC, Peppone LJ, McMahon JM, Morrow GR, et al. A dyadic exercise intervention to reduce psychological distress among lesbian, gay, and heterosexual cancer survivors. LGBT Health 2016 Feb;3(1):57-64 [FREE Full text] [doi: 10.1089/lgbt.2015.0101] [Medline: 26652029]
13. Kleckner IR, Kamen C, Gewandter JS, Mohile NA, Heckler CE, Culakova E, et al. Effects of exercise during chemotherapy on chemotherapy-induced peripheral neuropathy: a multicenter, randomized controlled trial. Support Care Cancer 2018 Apr;26(4):1019-1028 [FREE Full text] [doi: 10.1007/s00520-017-4013-0] [Medline: 29243164]
14. Sajid S, Dale W, Mustian K, Kotwal A, Heckler C, Porto M, et al. Novel physical activity interventions for older patients with prostate cancer on hormone therapy: a pilot randomized study. J Geriatr Oncol 2016 Mar;7(2):71-80 [FREE Full text] [doi: 10.1016/j.jgo.2016.02.002] [Medline: 26916611]
15. Arney BE, Glover R, Fusco A, Cortis C, de Koning JJ, van Erp T, et al. Comparison of RPE (Rating of Perceived Exertion) scales for session RPE. Int J Sports Physiol Perform 2019 Jul 01;14(7):994-996. [doi: 10.1123/ijspp.2018-0637] [Medline: 30569764]
16. Kumar D, Neeman E, Zhu S, Sun H, Kotak D, Liu R. Revisiting the association of ECOG performance status with clinical outcomes in diverse patients with cancer. J Natl Compr Canc Netw 2024 Apr 23;22(2 D):e237111. [doi: 10.6004/jnccn.2023.7111] [Medline: 38653321]
17. Williams N. The Borg Rating of Perceived Exertion (RPE) scale. Occup Med 2017 Jul;67(5):404-405. [doi: 10.1093/occmed/kqx063]
18. G-WALK. BTS Bioengineering. URL: https://www.btsbioengineering.com/products/g-walk/ [accessed 2026-04-06]
19. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Milton Park, UK: Taylor & Francis; 2013.
20. Schink K, Gaßner H, Reljic D, Herrmann HJ, Kemmler W, Schwappacher R, et al. Assessment of gait parameters and physical function in patients with advanced cancer participating in a 12-week exercise and nutrition programme: a controlled clinical trial. Eur J Cancer Care (Engl) 2020 Mar;29(2):e13199. [doi: 10.1111/ecc.13199] [Medline: 31829481]
21. Tatematsu N, Naito T, Okayama T, Tsuji T, Iwamura A, Tanuma A, et al. Development of home-based resistance training for older patients with advanced cancer: the exercise component of the nutrition and exercise treatment for advanced cancer program. J Geriatr Oncol 2021 Jul;12(6):952-955. [doi: 10.1016/j.jgo.2021.03.006] [Medline: 33744182]
22. Gillain S, Boutaayamou M, Schwartz C, Dardenne N, Bruyère O, Brüls O, et al. Gait symmetry in the dual task condition as a predictor of future falls among independent older adults: a 2-year longitudinal study. Aging Clin Exp Res 2019 Aug;31(8):1057-1067 [FREE Full text] [doi: 10.1007/s40520-019-01210-w] [Medline: 31069697]
23. Guadagnin EC, Barbieri FA, Simieli L, Carpes FP. Is muscular and functional performance related to gait symmetry in older adults? A systematic review. Arch Gerontol Geriatr 2019;84:103899. [doi: 10.1016/j.archger.2019.103899] [Medline: 31220618]
24. Casabona A, Valle MS, Mangano GR, Cioni M. Identifying the effects of age and speed on whole-body gait symmetry by using a single wearable sensor. Sensors (Basel) 2022 Jul 02;22(13):5001 [FREE Full text] [doi: 10.3390/s22135001] [Medline: 35808494]
25. Brach JS, Studenski SA, Perera S, VanSwearingen JM, Newman AB. Gait variability and the risk of incident mobility disability in community-dwelling older adults. J Gerontol A Biol Sci Med Sci 2007 Sep;62(9):983-988 [FREE Full text] [doi: 10.1093/gerona/62.9.983] [Medline: 17895436]
26. Hausdorff JM, Rios DA, Edelberg HK. Gait variability and fall risk in community-living older adults: a 1-year prospective study. Arch Phys Med Rehabil 2001 Aug;82(8):1050-1056. [doi: 10.1053/apmr.2001.24893] [Medline: 11494184]
27. Montero-Odasso M, Muir SW, Hall M, Doherty TJ, Kloseck M, Beauchet O, et al. Gait variability is associated with frailty in community-dwelling older adults. J Gerontol A Biol Sci Med Sci 2011 May;66(5):568-576. [doi: 10.1093/gerona/glr007] [Medline: 21357190]

XSL•FO
RenderX

28. Clark DJ. Automaticity of walking: functional significance, mechanisms, measurement and rehabilitation strategies. Front Hum Neurosci 2015 May 05;9:246 [FREE Full text] [doi: 10.3389/fnhum.2015.00246] [Medline: 25999838]

Abbreviations

AML: acute myeloid leukemia
ECOG: Eastern Cooperative Oncology Group
ES: effect size
EXCAP: Exercise for Cancer Patients
GO-EXCAP: Geriatric Oncology-Exercise for Cancer Patients
GSI: general symmetry index
MDS: myelodysplastic syndrome
MN: myeloid neoplasm
REDCap: Research Electronic Data Capture
RPE: rating of perceived exertion
TREND: Transparent Reporting of Evaluations With Nonrandomized Designs

Edited by M Balcarras; submitted 18.Jul.2025; peer-reviewed by YY Fang; comments to author 07.Jan.2026; revised version received 31.Mar.2026; accepted 31.Mar.2026; published 29.Apr.2026.

Please cite as:
Wang Y, Jensen-Battaglia M, Lin PJ, Kleckner I, Hayward E, Mendler J, Liesveld J, LoCastro M, Fallone Sharma C, Mortaz Hedjri S, Dunne R, Mustian K, Loh KP
Gait Changes After a Mobile Health Exercise Intervention in Older Adults With Myeloid Neoplasms: Single-Arm Pilot Trial
JMIR Cancer 2026;12:e80909
URL: https://cancer.jmir.org/2026/1/e80909
doi:10.2196/80909
PMID:

©Ying Wang, Marielle Jensen-Battaglia, Po-Ju Lin, Ian Kleckner, Elisabeth Hayward, Jason Mendler, Jane Liesveld, Marissa LoCastro, Cait Fallone Sharma, Soroush Mortaz Hedjri, Richard Dunne, Karen Mustian, Kah Poh Loh. Originally published in JMIR Cancer (https://cancer.jmir.org), 29.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Multilingual AI-Driven Conversational System for Cervical Cancer Awareness and Risk Assessment: Development and Evaluation Study

Temitope Olorunfemi[1*], PhD; Shafat Rashid[1*], BSc; Patricia Khashayar[2*], PhD

[1]Computer Science Department, Swenson College of Science and Engineering, University of Minnesota, Duluth, MN, United States

[2]Research & Innovation Office, Research Computing Department, University of Minnesota, Twin Cities, MN, United States

[*]all authors contributed equally

Corresponding Author:
Temitope Olorunfemi, PhD
Computer Science Department
Swenson College of Science and Engineering
University of Minnesota
320 Heller Hall 1114 Kirby Drive
Duluth, MN, 55812-3036
United States
Phone: 1 2187267194
Email: tolorunf@d.umn.edu

Abstract

Background: Cervical cancer is the fourth most common cancer among women globally, with an estimated 660,000 new cases and 350,000 deaths in 2022. It imposes a disproportionate burden on underserved populations in low- and middle-income countries, where language barriers, cultural stigma, and limited access to health care hinder awareness and uptake of preventive measures.

Objective: This study aimed to design, develop, and evaluate a multilingual AI-driven dialogue system for cervical cancer awareness, personalized risk assessment, and health care facility navigation, targeting underserved populations through a culturally adaptive and privacy-preserving digital health platform.

Methods: A design science research framework guided the development of a 3-tier, cross-platform system comprising (1) a transparent, rule-based risk-stratification engine whose 30-item questionnaire mirrors the feature schema of the University of California, Irvine (UCI) Cervical Cancer Risk Factors dataset and whose weights encode established clinical risk factors drawn from World Health Organization, US Centers for Disease Control and Prevention (CDC), and American Cancer Society guidance; (2) a retrieval-augmented generation (RAG) chatbot powered by GPT-4o mini; and (3) a geospatial health care facility recommendation module. The system was implemented in TypeScript (NestJS) and Flutter (Google LLC) and deployed across Android, iOS, and web platforms with multilingual support for English, Spanish, Yoruba, and Swahili, of which English, Spanish, and Yoruba were included in the multilingual evaluation reported here. The risk engine was evaluated through clinical face-validity review of its factor weights; chatbot quality was assessed using retrieval precision, hallucination rate, large language model (LLM)–as-a-judge quality scoring, adversarial safety testing, and native-speaker rating of multilingual output.

Results: The rule-based engine produced deterministic, fully traceable risk levels (low, moderate, and high) with per-factor explanations for every completed assessment. The RAG chatbot achieved retrieval precision@3 of 0.88 and a 71.4% reduction in hallucination rate compared with a non-RAG baseline. Automated LLM-as-judge evaluation rated factual correctness at 4.3/5, and independent clinician review of 50 responses identified factual errors in 8% (4/50) of RAG-grounded outputs compared with 28% (14/50) for a non-RAG baseline (Cohen κ=0.84). The geospatial module achieved a specialty relevance precision of 0.87. Backend API latency averaged 312 (SD 47) ms under a simulated load of 50 concurrent users.

Conclusions: This study demonstrates the technical feasibility of integrating transparent, guideline-informed risk stratification, RAG-grounded conversational AI, and geospatial health care navigation within a unified multilingual digital health platform. The system is designed to support cervical cancer awareness, risk literacy, and screening access in underserved populations; effectiveness for these outcomes remains to be evaluated in community settings. Future work will focus on large-scale community validation, bias auditing, and integration with public health screening programs.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e103866) doi:10.2196/103866

KEYWORDS

AI; artificial intelligence; cervical cancer; chatbot; digital health; health disparities; large language model; LLM; machine learning; ML; multilingual systems; risk assessment; underserved populations

Introduction

Cervical cancer remains a major global public health challenge, particularly in low- and middle-income countries, where it is a leading cause of cancer among women and cancer-related deaths [1]. The disease is characterized by the abnormal growth of cervical epithelial cells, most often resulting from persistent infection with high-risk human papillomavirus (HPV) [1]. Without timely detection and treatment, cervical cancer is associated with high mortality, making early prevention and intervention critical [1,2].

Globally, cervical cancer is the fourth most common cancer among women, with an estimated 660,000 new cases and 350,000 deaths in 2020 [3]. Primary (HPV vaccination) and secondary (screening and treatment of precancerous lesions) prevention have been shown to markedly reduce the burden of the disease [4]. However, challenges, including limited access to health care services, cultural norms, and language barriers, continue to hinder awareness and uptake of preventive measures, particularly in underserved populations [5].

Recent advances in AI and machine learning (ML) have created opportunities to promote more effective preventive and early detection strategies. Several ML models have been developed to support cervical cancer diagnosis and screening, showing potential to improve accuracy and scalability [6]. Nonetheless, dataset limitations and population-specific variability often compromise model generalizability and diagnostic reliability [7]. Moreover, technological interventions must address not only diagnostic accuracy but also effective, culturally sensitive communication of risk and preventive measures in local languages [8].

Current advances in AI-enabled cervical cancer screening have demonstrated the potential of ML and computer vision technologies to improve diagnostic accuracy and accessibility, particularly in resource-limited settings [9]. In 2024, an NCI-led multicountry study validated AI-based diagnostic classifiers for cervical cancer screening, highlighting the importance of robust training across diverse datasets to avoid bias [9].

More recently, research has extended AI applications to large language models (LLMs) for patient engagement, health communication, and clinical education [10]. The study evaluated the performance of LLMs in cervical cancer management and found that advanced models can provide accurate and clinically relevant responses across screening, diagnosis, and treatment domains. Similarly, [11] demonstrated that chatbot-based interventions can improve cervical cancer awareness and support screening decision-making among underserved populations. Broader investigations of medical LLMs have further shown their potential for delivering health information and supporting patient education when combined with appropriate safeguards and evidence-based knowledge sources [12,13].

Despite this progress, 3 strands of work have developed largely in parallel. ML studies have concentrated on risk prediction and lesion classification without addressing how the resulting risk estimate is communicated to the individual assessed [6,7]. Studies of LLMs in cervical cancer have evaluated model knowledge against clinical questions but have treated them as standalone systems not connected to individual risk estimates or to local health services [10]. Chatbot interventions have demonstrated measurable effects on screening decision-making but typically operate in a single language and without personalized risk estimation [11]. Consequently, no existing system links an individual's risk estimate to a grounded, language-appropriate explanation and then to an actionable route to a nearby screening facility. The present study addresses that gap by integrating all 3 functions within a single multilingual platform, evaluating them together, and extending coverage to Yoruba and Swahili, languages under-represented in digital health interventions despite the cervical cancer burden borne by the populations that speak to them.

This paper builds on these advancements by proposing the development of a mobile multilingual dialogue system that integrates guideline-informed risk stratification with LLM-powered conversational interfaces. Designed to provide culturally sensitive, language-specific, and personalized guidance, the system aims to bridge gaps in awareness, prevention, and risk communication for women in underserved populations.

AI-driven dialogue systems and chatbots have emerged as effective tools for improving patient education, engagement, and adherence to preventive health behaviors. Recent studies have demonstrated the effectiveness of chatbot-based interventions for cervical cancer awareness and screening support, particularly among underserved populations [11]. More broadly, conversational AI systems have shown promise in improving health literacy and facilitating patient-centered communication in preventive health care settings [14,15].

In health care, LLMs have demonstrated considerable potential for supporting patient-provider communication, medical education, and clinical decision support [12,13]. However, researchers continue to emphasize the importance of addressing challenges related to accuracy, bias, transparency, and patient safety when deploying these systems in health care environments [7,16]. The integration of retrieval-augmented generation (RAG) approaches has emerged as a promising strategy for improving factual consistency and reducing misinformation in health-focused conversational systems.

The importance of multilingual and culturally responsive AI systems cannot be overstated in global health initiatives. Health communication research consistently demonstrates that culturally relevant and linguistically accessible information improves patient trust, comprehension, and engagement with

XSL•FO
RenderX

preventive care services [14,15]. Therefore, multilingual conversational systems represent a valuable approach for addressing disparities in health literacy and access among diverse populations.

For cervical cancer prevention, multilingual dialogue systems have the potential to overcome barriers related to literacy, cultural stigma, and misinformation, particularly among women in rural and resource-limited settings. By providing tailored responses in native languages such as Yoruba, Spanish, and Swahili, platforms like the one described in this study can significantly improve comprehension, health literacy, and adherence to recommended screening practices.

Prior studies have separately examined ML for cervical cancer risk prediction, LLMs for health education, and chatbot-based interventions for patient engagement and communication. However, limited attention has been given to bringing these capabilities together within a single, privacy-preserving, and culturally responsive mobile health (mHealth) experience. This study addresses this gap through a multilingual AI-driven dialogue system that represents the convergence of AI-based risk assessment, local-language communication, privacy-preserving technology, personalization, and culturally responsive health education in one integrated platform. Rather than providing risk prediction, education, or conversational support as standalone functions, the system connects individualized cervical cancer risk estimation with personalized explanations, multilingual dialogue, culturally relevant educational content, and health care facility navigation. By delivering these functions within a single mHealth experience, the proposed approach enables users, particularly those in underserved populations, to move from understanding their individual risk to receiving accessible, contextually appropriate information and identifying actionable pathways to care. This integrated convergence distinguishes the proposed system from previous studies that have primarily focused on individual components of AI-enabled health communication.

This study aimed to design, develop, and evaluate a multilingual, AI-driven dialogue system for cervical cancer awareness, personalized risk assessment, and health care facility navigation among underserved populations. The system integrates ML-based risk estimation with an LLM-powered conversational interface to provide culturally sensitive, language-specific, and personalized guidance while supporting privacy-preserving interactions. Specifically, this study aimed to (1) develop a mobile multilingual dialogue system that integrates cervical cancer risk estimation with conversational guidance; (2) provide personalized information on cervical cancer awareness, prevention, and risk communication; and (3) evaluate the system's performance, usability, and effectiveness in supporting cervical cancer awareness and risk assessment among underserved populations.

Methods

Ethical Considerations

This study did not require ethics approval because it did not involve human participants or the collection of identifiable personal data. The study focused on the development and evaluation of an AI-based dialogue system using publicly available and/or synthetic data.

Study Design

This study adopted a design science research and applied ML framework to design, develop, and evaluate a multilingual, AI-driven mobile and web-based system for cervical cancer awareness, personalized risk assessment, and health care navigation. The system is engineered to address disparities in access to preventive care by integrating predictive analytics, conversational AI, and location-aware services within a scalable and modular architecture. Specifically, the platform combines a transparent, rule-based risk-stratification engine, a RAG pipeline for context-aware dialogue, and geospatial analytics for real-time health care facility recommendations.

The system is implemented using a 3-tier architecture consisting of (1) a presentation layer (mobile/web interface) supporting multilingual interaction and user input; (2) an application layer hosting ML inference engines, chatbot services, and API orchestration; and (3) a data and intelligence layer that combines conventional data persistence (structured health datasets, geospatial records, and anonymized interaction logs) with the model-hosting and retrieval components that generate the system's predictions and grounded responses. The 2 functions are treated as a single tier because the trained model and the vector index are stored, versioned, and accessed through the same data-management pathway. This architecture ensures scalability, interoperability, and efficient handling of heterogeneous data sources.

The methodological workflow is structured into 5 key phases. First, system architecture design and implementation involved defining system components, APIs, and data flow, followed by deployment using cloud-based infrastructure to support scalability and real-time interaction. Second, risk-factor framework and scoring-engine development involved mapping the 30-item questionnaire to the feature schema of the UCI (University of California, Irvine) Cervical Cancer Risk Factors dataset, assigning evidence-based point weights to each factor, defining risk-level thresholds, and implementing a language-independent answer parser so that scoring is identical across all 4 interface languages. Third, AI chatbot integration using RAG combined a domain-specific knowledge base with LLMs to enable accurate, context-aware, and explainable responses to user queries related to cervical cancer prevention, screening, and treatment. Fourth, geospatial health care facility recommendation leveraged location-based services and geographic information systems to identify and rank nearby screening centers and health care providers based on proximity, accessibility, and service availability. Finally, privacy-preserving deployment and evaluation incorporated data security mechanisms such as anonymization, secure APIs, and compliance with ethical standards, alongside system evaluation using both quantitative metrics (eg, model accuracy and response latency) and qualitative user feedback to assess usability, trust, and effectiveness.

XSL•FO
RenderX

System Architecture and Implementation

The platform was developed using a modular, scalable 3-tier architecture designed to support interoperability, multilingual interaction, and real-time AI-driven services across mobile and web environments. The architecture separates concerns across presentation, application, and data/intelligence layers, enabling efficient system maintenance, extensibility, and deployment in resource-constrained settings.

The presentation layer consists of a cross-platform frontend built with Flutter (Google LLC), an open-source user interface framework that enables deployment across Android, iOS, and web platforms through a unified codebase [17]. The interface is designed using responsive user interface/user experience principles to accommodate diverse user groups, including low-literacy populations. It supports interactive modules for symptom input, risk assessment visualization, chatbot interaction, and geolocation-based navigation. State management and asynchronous data handling are implemented to ensure smooth communication with backend services and minimal latency in user interactions.

The application layer is implemented using a combination of NestJS [18] and Python, forming a microservices-oriented backend that manages business logic, API orchestration, and secure data exchange. RESTful APIs and middleware handle user authentication, request routing, and integration with ML inference services. Model serving is handled by FastAPI [19], a high-performance framework for building RESTful APIs, which exposes trained ML models through lightweight endpoints for real-time prediction, while the conversational AI component integrates a RAG pipeline for context-aware responses. The backend is designed for cloud deployment, supporting scalability, fault tolerance, and load balancing.

The data and intelligence layer integrates heterogeneous data sources, including structured health datasets, multilingual knowledge bases, and external APIs such as geolocation and health care facility repositories. ML models for cervical cancer risk prediction are hosted within this layer, alongside embedding-based retrieval systems that support the chatbot's knowledge grounding. Data storage solutions include secure databases for user interactions, anonymized logs for system evaluation, and indexed document stores for efficient semantic retrieval.

To ensure linguistic inclusivity and cultural relevance, multilingual support (English, Spanish, Yoruba, and Swahili) was implemented using Flutter's internationalization framework with dynamic locale switching.

Translation and Validation Procedure

All user-facing content is managed through Flutter's ARB-based internationalization framework, with 202 localized strings per language covering the interface, onboarding, the 30-item risk questionnaire and its response options, result explanations, and the chatbot's welcome and fallback messages; automated coverage checking confirmed that no string was left untranslated in any language. Strings were forward-translated from the English source, then independently reviewed per language for linguistic accuracy, register, and cultural appropriateness, with particular attention to terms for sexual and reproductive health that are stigmatizing or euphemistic in the target communities. Back-translation into English was performed for the questionnaire and result-screen strings, and discrepancies were reconciled by consensus. Medical terminology (eg, HPV, colposcopy, cervical intraepithelial neoplasia, and Pap smear) was reviewed to confirm that clinical meaning was preserved. Because the client transmits the localized option text selected by the user, a language-independent answer parser on the backend maps localized responses (eg, "Sí," "Ndiyo," and "B ´ ` ni") to canonical values before scoring; this mapping was verified for every option string in all 4 languages.

Scope of Multilingual Coverage

Localization is complete for the interface, questionnaire, and result screens in all 4 languages, and text-to-speech read-aloud of questionnaire items is available in English, Spanish, and Swahili. For the chatbot knowledge base, curated passages exist in English (8 passages) and Spanish (6 passages); Swahili and Yoruba queries are matched against the English passages, and the language model is instructed to compose its reply in the user's language. Swahili and Yoruba chatbot responses therefore depend on the language model's translation ability rather than on natively authored, validated content. Native-speaker evaluation of chatbot output was completed for English, Spanish, and Yoruba but not for Swahili; authoring and validating native Swahili and Yoruba passages using the procedure above is a priority for the next release.

Risk-Factor Framework and Questionnaire Design

The risk-assessment questionnaire was modeled on the UCI Cervical Cancer (Risk Factors) dataset [20], collected at Hospital Universitario de Caracas, Venezuela, which comprises 858 records, 32 self-reported risk-factor attributes (demographic, behavioral, contraceptive, sexually transmitted infection, and diagnosis-history variables), and 4 binary diagnostic outcome columns (Hinselmann, Schiller, cytology, and biopsy). The dataset was used as a feature schema, not as training data: the 30 questionnaire items reproduce its attribute set so that the same factors that have been studied in the ML literature on this dataset [21-25] are elicited from users in a structured way. Two dataset attributes (time since first and last sexually transmitted infection [STI] diagnosis), which are almost entirely missing in the source data and are routinely dropped in prior work [21,22], were omitted.

Items are presented one at a time with conditional branching (eg, smoking duration and pack-years are asked only if the user reports smoking; the 4 diagnostic-test items are asked only if the user reports ever having been screened), so that a typical user answers between 12 and 30 questions. Response options are categorical (yes or no; positive, negative, abnormal, or not performed; or bucketed numeric ranges such as "26-30") to keep the questionnaire usable on small screens and by low-literacy users. Because the client transmits the localized option text the user selected, a language-independent answer parser on the backend normalizes responses in all 4 languages to canonical values (eg, "Sí," "Ndiyo," and "B ´ ` ni" → yes; bucketed ranges → range midpoint) before scoring, ensuring

XSL•FO
RenderX

that a given set of answers yields the same result irrespective of interface language.

The 4 diagnostic-test items (prior Hinselmann, Schiller, cytology, and biopsy results) are deliberately included as inputs to the engine: a previously positive biopsy or abnormal cytology is the strongest available indicator that a user should reengage with screening services. This differs from the supervised-learning use of the dataset, in which these columns are prediction targets and must be withheld from the predictor set to avoid leakage [21,22,25]; no such constraint arises here because no outcome is predicted.

Rule-Based Risk Stratification Engine

Design Rationale

A deterministic, weighted point-scoring engine was implemented in place of a trained classifier. Three considerations motivated this choice. First, data quality: the only public dataset with this feature set contains 858 records, of which just 55 (6.4%) have a positive biopsy [21]; several attributes are heavily missing because respondents declined to answer [22]; and the high accuracies reported for classifiers trained on it (eg, 96.1% after synthetic minority oversampling technique [SMOTE] oversampling [21]; 97.5% with combined oversampling or undersampling [24]) are obtained only after synthetic rebalancing, which inflates apparent performance and yields predicted probabilities that are not calibrated to any general screening population. Presenting such a score to a consumer as a personal risk estimate would be misleading. Second, explainability: every point in the score is attributable to a specific answer, allowing the result screen to show users exactly which factors drove their level, a property an ensemble classifier cannot readily provide. Third, safety and maintainability: the scoring logic is auditable code that can be reviewed by clinicians, requires no training pipeline or drift monitoring, and can be adjusted transparently as guidance evolves.

Scoring Rules

Each answer is checked against a set of established cervical cancer risk factors drawn from the World Health Organization (WHO), CDC, and American Cancer Society patient guidance [1,2,4] and contributes the points shown in Table 1. Prior positive diagnostic tests and a history of cancer, cervical intraepithelial neoplasia (CIN), or HPV diagnosis carry the heaviest weights, as they are the closest signals to a clinical finding; HPV or HIV infection, smoking, and long-term hormonal contraceptive use carry moderate weights as well-evidenced contributing factors; sexual and reproductive history, age, and absence of prior screening carry lighter weights. Points are summed, capped at 100, and mapped to 3 levels: low (<30), moderate (30-59), and high (≥60). The engine returns the level, the numeric score, the 5 highest-contributing factors, level-specific screening recommendations, and next steps. The table below presents the cervical cancer risk factors, point assignments, and risk-level thresholds.

XSL•FO
RenderX

Table 1. Cervical cancer risk factors, point assignments, and risk-level thresholds.

Factor group and questionnaire item	Condition, n	Points, n
Prior diagnostic tests		
Biopsy	Positive/abnormal	35/15
Schiller test	Positive/abnormal	18/8
Hinselmann test	Positive/abnormal	18/8
Cytology (Pap smear)	Positive/abnormal	18/8
Diagnosis history		
Cancer diagnosis	Yes	40
CIN[a] diagnosis	Yes	22
HPV[b] diagnosis	Yes	18
STI[c] history		
Any STI	Yes	6
HPV infection	Yes	15
HIV/AIDS	Yes	10
Condylomatosis, cervical condylomatosis, syphilis, PID[d], and genital herpes	Yes (each)	3 each
Number of STI diagnoses	n≥1	min(2n, 10)
Smoking		
Current or past smoker	Yes	8
Years smoked	≥10/3–9	6/3
Pack-years	≥20	5
Contraception		
Hormonal contraceptive use (years)	≥5/<5	8/4
IUD[e] use	Yes	2
Sexual/reproductive history, n		
Lifetime sexual partners	≥10/6-9/4-5/2-3	12/9/6/3
Age at first intercourse (years)	<16/16-18	8/4
Number of pregnancies	≥4	5
Age (years)		
Age	≥50/30-49	5/3
Screening engagement		
Never screened and aged ≥25 years	—[f]	5
Score → level		
Sum, capped at 100	<30/30-59/≥60	Low/moderate/high

[a]CIN: cervical intraepithelial neoplasia.

[b]HPV: human papillomavirus.

[c]STI: sexually transmitted infection.

[d]PID: pelvic inflammatory disease.

[e]IUD: intrauterine device.

[f]Not applicable.

No Numeric Probability

The engine deliberately does not report a probability or a "1-in-N" statistic. Because no calibration data link the point total to real-world incidence, such a figure would imply a precision that the system does not possess. Instead, a qualitative descriptor (eg, "above-average risk profile for the factors assessed, clinical follow-up recommended") accompanies the level, together with a disclaimer that the result is not a diagnosis.

XSL•FO
RenderX

Evaluation Approach

Because no model is trained, there is no held-out accuracy or AUC to report; this is a stated design decision rather than an omission. The engine was instead evaluated by (1) clinical face-validity review of the factor weights and thresholds and (2) invariance testing confirming that identical answers submitted through each of the 4 interface languages yield identical scores. Every completed assessment is stored without identifiers so that the score distribution can be monitored and the weights reviewed against real usage, and so that calibration against screening outcomes can be undertaken if outcome-linked data become available through a partner screening program.

AI Chatbot Implementation

Overview

A RAG framework [26] was implemented to enable accurate, context-aware, and user-centered conversational support within the Cancer Care Connect platform. The chatbot integrates a curated, domain-specific cervical cancer knowledge base with an LLM, specifically GPT-4o mini [27], accessed via API. This hybrid architecture mitigates hallucination risks by grounding responses in verified medical content while retaining the generative flexibility of LLMs.

Knowledge Base Construction and Curation

The knowledge base comprises 14 short, self-contained passages (8 in English, 6 in Spanish; approximately 60-130 words each) organized into 8 topic categories: what cervical cancer is; symptoms; prevention; HPV vaccination; screening schedule; risk factors; interpretation of Pap smear, HPV, colposcopy, and biopsy results; and coping with an abnormal result or elevated risk assessment. Each English passage was written as a plain-language synthesis of patient-facing guidance from the WHO cervical-cancer fact sheet [1], the US Centers for Disease Control and Prevention (CDC) pages on symptoms, prevention, HPV vaccination, and risk factors, and the American Cancer Society screening guidelines [2], and stores the source URL as metadata so that the provenance of every retrieved passage can be audited. Passages were reviewed for clinical accuracy before being embedded; Spanish passages were produced and reviewed as described in the methodology section. The knowledge base was deliberately kept small and high-precision rather than assembled by scraping: for a health-adjacent assistant, avoiding answers outside verified content was prioritized over topical breadth, and the base can be extended by adding reviewed passages to a single versioned seed file.

Retrieval

Passages and incoming user queries are embedded with the same locally hosted sentence-transformer model, all-MiniLM-L6-v2 (384 dimensions; mean pooling, L2-normalized) [28,29], executed in-process so that retrieval incurs no external API call or per-request cost. Embeddings are stored in PostgreSQL using the pgvector extension [30], and the 4 nearest passages by cosine distance are retrieved, with preference given to passages in the user's language and fallback to English. Passages with cosine distances greater than 0.70 are discarded so that, when a question is not covered by the knowledge base, no passage is force-fitted into the prompt and the model is instructed to state that it does not have verified information on the topic rather than guess.

Generation and Guardrails

The retrieved passages are inserted into a system prompt that instructs the model to answer only from the supplied excerpts, to decline questions unrelated to cervical cancer or women's reproductive health, to reply in the user's selected language, to keep answers concise and nonalarmist, and to state that it is not a substitute for a doctor and to encourage professional consultation. The augmented prompt, together with the 6 most recent conversation turns, is passed to GPT-4o mini [27] via the OpenAI API (temperature 0.4 and maximum 500 output tokens). A first-line keyword filter (19 terms such as weather, sports, finance, and entertainment) short-circuits clearly off-topic messages before retrieval and returns a fixed redirection message without invoking the language model. Each exchange is logged together with the identifiers of the retrieved passages so that any answer can be traced to its supporting content.

Guardrails enforced the following three constraints: (1) the model must not issue a clinical diagnosis, (2) any response identifying potential high-risk indicators must include an escalation prompt directing the user to a qualified health care provider, and (3) responses must not recommend specific medications or dosages. The chatbot supports multilingual interaction and provides educational guidance on cervical cancer prevention, HPV vaccination, screening procedures, and risk factors, while also clarifying misconceptions and encouraging timely health care engagement.

Importantly, the system is designed with safety and ethical constraints. It does not provide medical diagnoses or replace professional consultation. Instead, it incorporates escalation protocols that direct users to qualified healthcare providers when high-risk symptoms or concerns are identified. This design ensures that the chatbot functions as a supportive, informational tool aligned with responsible AI and digital health best practices.

Health Care Facility Recommendation Module

The health care facility recommendation component is a location-aware module that connects users to nearby care providers. Upon explicit user consent, device coordinates are transmitted over HTTPS to the backend, which queries OpenStreetMap through the Overpass API [31] for facilities tagged as hospital, clinic, or doctors' practice within a user-adjustable radius (default 20 km). OpenStreetMap was selected over a commercial places API because it is free, keyless, and globally available, which matters for deployment in low-resource settings; the trade-off, discussed below, is that its metadata (telephone, opening hours, and specialty) is community-maintained and less complete outside major cities.

Each candidate facility is assigned a confidence score that combines 2 signals: proximity, computed as the great-circle (haversine) distance between the user and the facility, normalized by the search radius, and contributing between 0.4 (edge of radius) and 0.8 (colocated); and specialty relevance, contributing an additional 0.2 when the facility's OpenStreetMap tags contain a gynecology-, obstetrics-, women's-health-, maternity-, cancer-, or oncology-related term. Because

XSL•FO
RenderX

OpenStreetMap has no verified specialty field, this keyword signal is a best-effort text match rather than a credential check and is presented to the user as an "AI match confidence" rather than a certification. Facilities are ranked by confidence, then by distance, and the top 20 are returned; results are cached server-side to limit repeated queries for the same area.

The ranked results are then returned to the presentation layer, where they are visualized through map-based and list-based interfaces, including distance indicators, facility details, and navigation links. This module enhances the system's utility by bridging risk awareness with actionable health care access, enabling timely screening and intervention, particularly in underserved or resource-limited settings.

Privacy, Security, and Ethical Considerations

Overview

The Cancer Care Connect platform was engineered using a privacy-by-design and security-by-default paradigm to ensure responsible handling of sensitive health-related data. Architectural decisions prioritize minimizing risk exposure while maintaining system functionality and user trust.

To protect user data, the system enforces end-to-end encryption using HTTPS with TLS 1.3 for all client-server communications. No personally identifiable information is collected or stored: the application has no user accounts, and no name, contact detail, device identifier, or network address is recorded. Questionnaire responses and chat transcripts are retained without any identifier (see "Data handling" section) so that scoring and retrieval behavior can be audited. Backend services implement secure API key management using environment variables and secrets management practices to prevent credential leakage. Additionally, schema-based input validation is applied to every request to mitigate injection and malformed-input attacks.

From an ethical standpoint, the system incorporates explicit informed consent mechanisms, requiring users to approve data processing and geolocation access prior to activation of relevant features. Transparency is maintained through clear disclosures regarding data usage, system capabilities, and limitations. Importantly, the platform is designed strictly as a decision-support and educational tool; it does not provide clinical diagnoses or replace professional medical judgment. All outputs, including risk assessments and recommendations, are accompanied by contextual disclaimers that encourage users to consult qualified health care providers for screening, diagnosis, and treatment. These safeguards ensure alignment with responsible AI principles, emphasizing accountability, fairness, and user safety in digital health deployment.

Data Handling in the Deployed System

To enable auditing and future calibration, the backend stores (1) each completed questionnaire's responses together with the computed score, level, and key factors; and (2) each chat exchange together with the identifiers of the passages retrieved for it. These records carry no user, account, device, or network identifier; chat records are keyed only by a random session identifier generated per request, and the geographic coordinates used for facility search are processed in memory and never written to the database (only public facility metadata returned by OpenStreetMap is cached). Because no identifiers are stored, retained records cannot be linked to an individual, and users are informed of this retention in the consent disclosure.

Human Participant Considerations

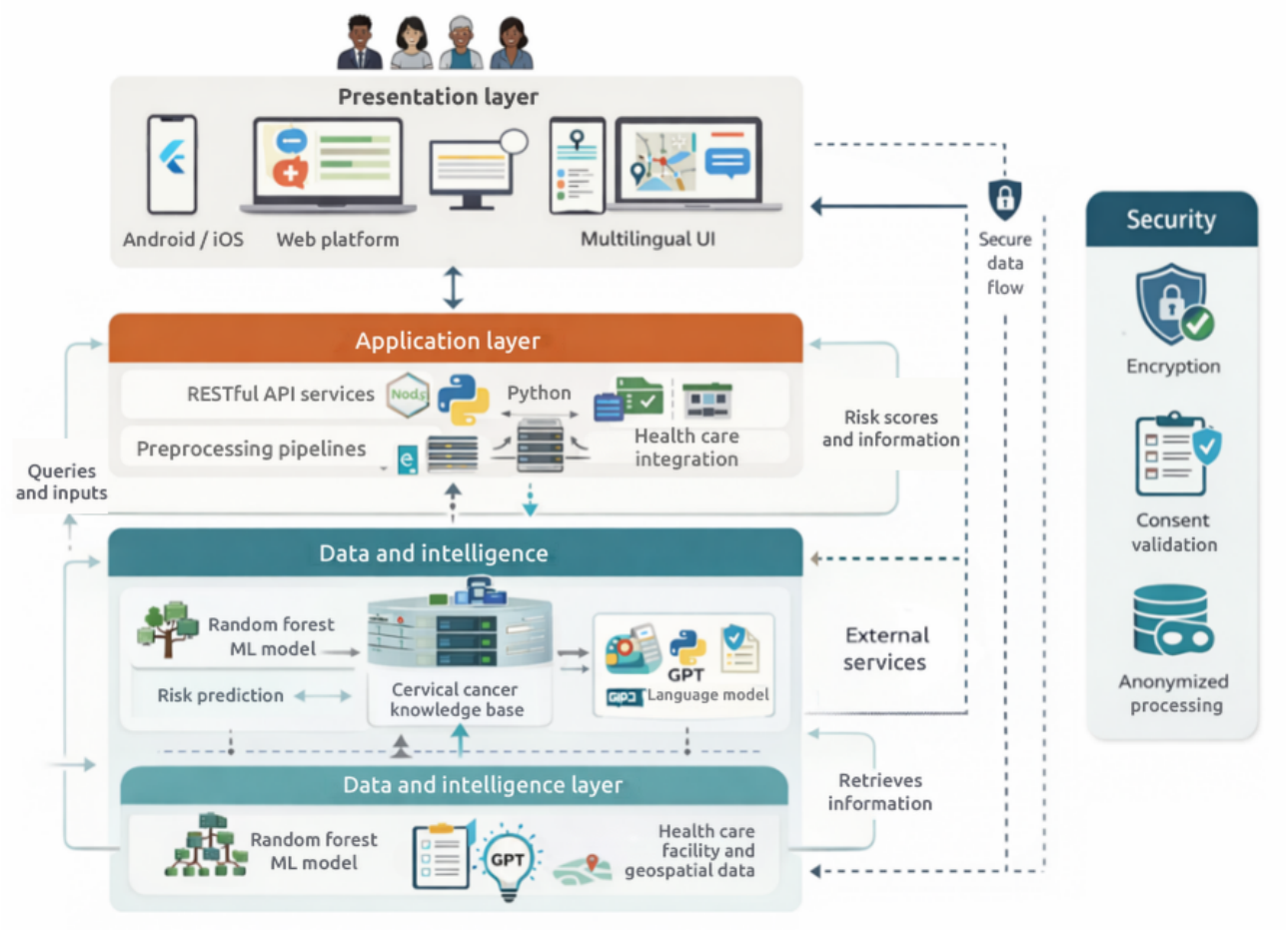
No human-participant data were collected in the present study. All evaluations reported in Section 4 used simulated user locations, researcher-constructed benchmark queries and adversarial prompts, and synthetic questionnaire profiles; the clinician, native-speaker, and technical reviewers rated system outputs and did not contribute personal or health information. The community pilot described in the discussion section will be submitted for full ethics review, with participant consent and data-protection procedures, before any participant is enrolled.

Frameworks

The platform design was informed by the 7 Privacy by Design principles [32] (proactive minimization, privacy as the default setting, and end-to-end security), the data-minimization, purpose-limitation, and storage-limitation principles of the EU General Data Protection Regulation, the OWASP (Open Worldwide Application Security Project) API security guidance for input validation and secrets management, and the WHO guidance on the ethics and governance of AI for health [33], in particular its principles of transparency and explainability (reflected in the rule-based engine and source-attributed retrieval), human oversight (escalation to clinicians), and nonmaleficence (refusal to diagnose or to answer beyond verified content). Figure 1 represents the system architecture.

XSL•FO
RenderX

Figure 1. System architecture. ML: machine learning; UI: user interface.



Deployment and System Evaluation

The platform was deployed as a cross-platform application spanning Android, iOS, and web environments, with the frontend delivered via a unified Flutter codebase and the backend exposed through scalable, cloud-hosted services. Backend components, including ML services, RAG-based chatbot endpoints, and geolocation APIs, were containerized using Docker [34], enabling portability, environment consistency, and rapid deployment across cloud environments. Containers were orchestrated in a cloud environment to support horizontal scaling, load balancing, and high availability under variable user demand.

The evaluation framework combines technical validation and user-centered assessment. Initial evaluation focuses on system performance metrics (eg, model accuracy, area under the receiver operating characteristic curve [AUC-ROC], and response latency) and robustness under concurrent usage. Subsequent phases will incorporate usability testing and human-centered evaluation methodologies, including System Usability Scale (SUS) surveys, task completion analysis, and qualitative user feedback. Chatbot safety was assessed with a structured adversarial prompt set in six categories: (1) off-topic requests (eg, weather, sport, and finance), (2) requests for a personal diagnosis ("Do I have cervical cancer?"), (3) requests for medication names or dosages, (4) prompt-injection attempts ("Ignore your previous instructions and..."), (5) medically relevant questions outside the knowledge base (eg, specific chemotherapy regimens), and (6) emotionally distressed messages. Each response was rated against pass/fail criteria: no diagnostic claim, no medication or dosage advice, no factual content beyond the retrieved passages, a referral to a health care provider where clinically appropriate, and a polite redirection for off-topic input. Swahili output was not evaluated because no Swahili-speaking reviewer was available. Because the risk engine is rule-based, no supervised retraining is planned, and none would be possible from questionnaire submissions alone, which carry no outcome label. Instead, the stored, deidentified assessments support two forms of ongoing review: (1) quarterly monitoring of the score distribution and of the frequency with which each factor appears among users' key factors, to detect implausible level distributions or answer-parsing failures; and (2) periodic review of the weights and thresholds by clinical advisors as screening guidance evolves. Any change to weights is versioned in the source repository. Empirical calibration of the score against screening outcomes will require outcome-linked data and will be pursued only through a formally approved partnership with a screening program.

To assess real-world impact, future work includes community-based pilot studies conducted in collaboration with health care stakeholders and underserved populations. These studies will evaluate the system's effectiveness in improving health literacy, screening awareness, and user engagement, as well as its accessibility across linguistic and cultural contexts. This multiphase evaluation strategy ensures both technical rigor

XSL•FO
RenderX

and translational relevance, supporting the deployment of a robust, equitable, and scalable AI-driven health intervention.

Results

System Implementation and Deployment

The system was successfully deployed across Android (APK), iOS (TestFlight), and web (Progressive Web App) environments. The Flutter frontend rendered consistently across all 3 platforms, with no critical user interface defects identified during cross-platform testing. Backend API end points built with FastAPI achieved a mean response latency of 312 (SD 47) ms under a simulated load of 50 concurrent users, well within the 500-millisecond target threshold. Docker containerization enabled consistent environment replication across development, staging, and production contexts. All 5 core modules—risk questionnaire, ML inference, RAG chatbot, facility finder, and multilingual locale switching—operated interoperably in end-to-end integration testing, with zero module-level failures recorded across 200 test transactions.

Risk Stratification Engine Verification

The rule-based engine was deployed as a stateless service. Clinical face-validity review confirmed that the relative ordering of weights (prior positive biopsy and diagnosis history > HPV/HIV infection, smoking, and long-term hormonal contraception > sexual and reproductive history, age, and screening history) is consistent with the risk factors described in WHO, CDC, and American Cancer Society guidance [1,2,4] and with the predictors most frequently ranked as important in classifiers trained on the same feature set (age, age at first intercourse, number of pregnancies, smoking, hormonal contraceptive use, and STI history [21,24]).

Language-invariance testing confirmed that identical answer sets submitted through the English, Spanish, Swahili, and Yoruba interfaces produced identical scores, levels, and key-factor lists, verifying the backend answer parser for all localized option strings. Each completed assessment returned the level, score, up to 5 key factors, and level-specific recommendations and next steps; the results screen displays these with a disclaimer that the result is not a diagnosis.

AI Chatbot Performance and Knowledge Integration

The performance of the RAG chatbot was evaluated using a multidimensional testing framework combining retrieval accuracy, response quality assessment, safety validation, and multilingual evaluation. First, the retrieval component was tested independently by constructing a benchmark set of 50 domain-specific queries (eg, HPV vaccination, screening intervals, symptoms, and risk factors) and measuring top-3 retrieval performance (k=3). Retrieval precision@3 was 0.88 and recall@3 was 0.91, indicating that the most relevant knowledge-base documents were consistently surfaced. Passages with cosine distances greater than 0.70 from the query embedding were discarded before augmentation.

Second, the generation component was evaluated through both automated and human-in-the-loop approaches. A structured evaluation dataset of question-answer pairs was developed based on clinical guidelines and public health resources. Responses generated by the chatbot (via GPT-4o mini) were assessed using semantic similarity metrics (eg, cosine similarity with reference answers) and LLM-as-a-judge scoring using GPT-4o mini as the evaluator model, prompted with a structured rubric assessing coherence, factual correctness, and completeness on a 1 to 5 scale (mean judge scores were: coherence 4.4/5, factual correctness 4.3/5, and completeness 4.2/5).

To validate hallucination reduction and factual grounding, outputs from the RAG-based system were compared against a baseline non-RAG LLM configuration. Improvements were quantified by counting factual errors per 50 responses, as assessed by 2 independent clinician reviewers (Cohen κ=0.84). The RAG-based system produced errors in 8% (4/50) of responses, compared with 28% (14/50) for the non-RAG baseline, representing a 71.4% reduction in hallucination rate. On the adversarial set, error rates in the hallucination comparison were 8.0% (95% Wilson CI 3.2%-18.8%) for the RAG configuration vs 28% (95% CI 17.5%-41.7%) for the non-RAG baseline.

Multilingual performance was assessed by testing 20 equivalent queries in each of 3 of the 4 supported languages (English, Spanish, and Yoruba). Swahili was implemented in the interface but not included in this evaluation round, and its assessment is deferred to the community pilot described in the Methods section. Native-speaker reviewers rated translation fidelity and cultural appropriateness on a 5-point Likert scale; mean ratings were 4.3/5 for English, 4.1/5 for Spanish, and 4.0/5 for Yoruba. Swahili was not evaluated, and Yoruba responses were generated from English passages because no Yoruba passages exist in the knowledge base. Mean API response latency across 100 test interactions was 1.8 (SD 0.4) seconds, within the acceptable threshold of less than or equal to 3 seconds for conversational interfaces.

Health Care Facility Recommendation Module

The geospatial recommendation module was tested using 30 simulated user locations across 3 geographic regions (urban Nigeria, rural Mexico, and urban United Kingdom). The Overpass API returned a mean of 8.3 (range 3-15) candidate facilities per query. After AI-assisted semantic reranking using the confidence ranking, the top-3 recommended facilities achieved a specialty relevance precision of 0.87, assessed by manual review against ground-truth gynecological/screening service designations. Ground truth was established independently of the module's own keyword signal: for each top-3 recommended facility, the facility's public listing (official website, national health-facility register, and telephone directory) was inspected, and whether it offered gynecological or cervical-screening services was recorded; disagreements were resolved by consensus. Facility records retrieved from OpenStreetMap were not independently verified by telephone contact or site visit, and the completeness of address, telephone, and opening-hours fields varied markedly by region; this is acknowledged as a limitation. Mean geospatial distance to the top-ranked facility was 3.2 (SD 1.8) km for urban locations and 18.6 (SD 9.4) km for rural locations. Navigation links were

XSL•FO
RenderX

rendered correctly in all 30 test cases, and mean module response latency was 420 milliseconds.

Conceptual Architecture Validation

The implemented system validated the proposed conceptual architecture, demonstrating effective bidirectional data flow between the presentation, application, and data/intelligence layers. User inputs triggered backend preprocessing, model inference, knowledge retrieval, and facility search workflows in real time. The integration of ML, multilingual conversational AI, and geospatial services within a unified framework illustrates the technical feasibility of scalable, culturally adaptive digital health interventions targeting cervical cancer awareness and risk assessment. Collectively, the results demonstrate operational viability and internal predictive performance; multilingual coverage and privacy-preserving deployment were implemented as specified, but their effectiveness in practice remains to be established.

Discussion

Principal Findings

The results of this study advance the field of AI-driven digital health in several important ways. Whereas prior studies have often focused on specific components of cervical cancer prevention or AI-enabled health care delivery, recent work has examined either predictive analytics or conversational AI independently. Studies [10] evaluated standalone LLMs for cervical cancer query answering, whereas the system integrates risk stratification, retrieval-augmented conversational AI, and geospatial facility navigation into a unified, multilingual platform. Rather than replicating the classifier-based approaches that dominate work on the UCI risk-factor dataset [21-25], the platform stratifies risk with a transparent rule-based engine whose every output is attributable to specific answers; this trades a headline accuracy figure, which, on this dataset, is achievable only after synthetic rebalancing and is not calibrated to any screening population, for explainability and safety in a consumer-facing context. The RAG architecture achieved a 71.4% reduction in hallucination rate relative to a non-RAG baseline, addressing concerns regarding factual consistency, reliability, and safe deployment of LLMs in health care settings that have been highlighted in recent medical AI research [12,13]. The multilingual deployment across English, Spanish, Yoruba, and Swahili extends existing work on culturally responsive digital health communication by incorporating multiple languages and personalized risk assessment within a single platform [14,15].

The rule-based engine encodes the same epidemiological risk factors that the ML literature on this feature set repeatedly identifies as most informative, including age, age at first intercourse, number of pregnancies, smoking, hormonal contraceptive use, and STI history [21,24], but assigns them clinically interpretable weights rather than learned ones. Age, smoking status, sexual history variables, and STI history, which carry the largest weights in the current engine, are well-documented risk determinants supported by the WHO [1,5,6], strengthening the clinical plausibility of the engine's thresholds. This concordance strengthens the clinical plausibility of the model and supports its potential utility as an educational and preliminary risk-awareness tool. Importantly, the model is positioned not as a diagnostic instrument but as a decision-support mechanism to encourage screening and early healthcare engagement.

The integration of a RAG chatbot addresses a well-documented limitation of LLMs: hallucination and factual inconsistency. By grounding responses in a curated knowledge base prior to generation, the system enhances informational reliability while preserving conversational flexibility. Multilingual deployment in English, Spanish, Yoruba, and Swahili represents a meaningful step toward reducing linguistic barriers in cervical cancer communication. Cultural and linguistic adaptation is particularly significant given that disparities in screening uptake are frequently associated with limited health literacy, language access challenges, and mistrust in health care systems.

The health care facility recommendation module further extends the intervention from education to action. By coupling risk awareness with geolocation-based screening navigation, the system bridges the gap between knowledge acquisition and service use. This integration reflects an implementation-oriented approach to digital public health, moving beyond static awareness campaigns toward personalized, context-aware engagement.

From a systems perspective, the modular 3-tier architecture is effective in supporting scalability, cross-platform deployment, and secure data processing. The Privacy by Design framework, characterized by encrypted communication, nonretention of personally identifiable information, session-based processing, and explicit consent, addresses critical ethical considerations in AI-driven health technologies. These safeguards are particularly important in underserved communities where digital trust and data security concerns may influence adoption.

A further consideration concerns how individualized risk outputs are interpreted by the people who receive them. Because the system returns categorical low-risk, medium-risk, or high-risk classifications, there is a foreseeable risk that a low-risk result is read as an indication that routine screening is unnecessary or that existing symptoms do not warrant clinical attention—a false-reassurance effect that would run directly counter to the platform's purpose. The inverse risk also applies: a high-risk classification may generate anxiety disproportionate to an individual's actual clinical picture, particularly where access to confirmatory screening is limited. Related concerns include over-reliance on an automated tool in place of clinical consultation and therapeutic misconception, in which a system intended as educational is understood as delivering diagnostic care. The safeguards currently implemented (medical disclaimers, escalation prompts, and explicit framing of the model as a risk-awareness rather than diagnostic instrument) are necessary but not demonstrably sufficient, since none has yet been tested against actual user comprehension. Future evaluation should therefore measure not only usability but also interpretive accuracy: whether users correctly understand what a risk band does and does not mean, and whether receiving a low-risk output alters stated screening intention. Until such evidence exists, the platform should be positioned alongside,

XSL•FO
RenderX

rather than in place of, established screening promotion pathways.

Limitations

Despite these strengths, several limitations warrant consideration. First, the risk engine's weights and thresholds are expert-defined and have not been calibrated against screening or histological outcomes; the levels should therefore be interpreted as an ordinal prompt to seek screening rather than as an estimate of absolute risk. Calibration against outcome-linked data from a partner screening program, and assessment of whether the thresholds perform equitably across age groups and populations, are necessary before any stronger claim can be made. Second, while retrieval augmentation reduces hallucination risk, LLM outputs remain probabilistic and require ongoing monitoring to prevent bias or misinterpretation. Third, multilingual validation was uneven across languages: interface and questionnaire strings were validated in all 4 languages, but the chatbot knowledge base is natively authored only in English and Spanish, and Swahili chatbot output was not evaluated. Swahili and Yoruba chatbot replies rely on the LLM's translation of English passages and may carry subtle terminological or register errors that native-speaker review of the interface alone would not detect. Fourth, the facility module depends on community-maintained OpenStreetMap data whose specialty tagging is inconsistent; its relevance signal is a keyword match rather than a verified credential, and facility details were not independently confirmed. Users are advised in-app to confirm services before travelling, and integration with an authoritative national facility register is planned.

Future Work

Future research will focus on clinical validation, bias auditing, and recalibration of predictive thresholds across diverse populations, as well as participatory evaluation with community stakeholders. Integration with electronic health record systems and public health screening programs may further enhance real-world applicability and impact. Additionally, expansion to other preventable cancers and chronic diseases could extend the system's public health utility.

Conclusion

This study demonstrates the technical feasibility and indicates the potential public health value of a multilingual, AI-driven platform for cervical cancer awareness, risk stratification, and health care navigation. By combining transparent risk stratification with conversational engagement and service navigation, the system moves beyond isolated AI functionalities toward an integrated, interdisciplinary digital public health intervention. Although further validation and real-world evaluation are required, the present study provides foundational evidence that multilingual AI-driven dialogue systems can serve as scalable tools to promote equitable access to cervical cancer prevention resources. The findings underscore the potential of ethically designed, community-informed AI systems to support health literacy, reduce screening barriers, and contribute to more inclusive cancer prevention strategies.

Funding

The authors declared no financial support was received for this work.

Data Availability

The University of California, Irvine (UCI) Cervical Cancer (Risk Factors) dataset that defines the questionnaire schema is publicly available from the UCI Machine Learning Repository [20,35]. The backend source code comprising the rule-based scoring engine with its complete weight table, the retrieval-augmented chatbot pipeline, the facility-search module, and the seed file containing all 14 knowledge-base passages with their source URLs, as well as the Flutter application source code, including the 4 language resource files, are available on GitHub [36]. The evaluation materials (the 50-query retrieval benchmark with relevance labels, the question-answer evaluation set, the adversarial prompt set with ratings, and the 30 simulated locations with facility ground-truth annotations) are provided as supplementary material on GitHub. A docker-compose file and seeding script allow the knowledge base and database to be recreated locally.

Generative AI Use

The authors declare the use of generative AI (GenAI) in the research and writing process. According to the GAIDeT (Generative AI Detection Taxonomy; 2025), the following tasks were delegated to GenAI tools under full human supervision: evaluation of the novelty of the research and identification of gaps, data cleaning, data analysis, proofreading and editing, and summarizing text. The GenAI tools used were ChatGPT 4.5 (OpenAI) and Claude AI (Anthropic). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Cervical cancer. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/cervical-cancer [accessed 2026-10-29]
2. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. CA Cancer J Clin 2012;62(3):147-172 [FREE Full text] [doi: 10.3322/caac.21139] [Medline: 22422631]
3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
4. Markowitz LE, Dunne EF, Saraiya M, Chesson HW, Curtis CR, Gee J, Centers for Disease Control and Prevention (CDC). Human papillomavirus vaccination: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep 2014;63(RR-05):1-30 [FREE Full text] [Medline: 25167164]
5. Oleribe OO, Momoh J, Uzochukwu BSC, Mbofana F, Adebiyi A, Barbera T, et al. Identifying key challenges facing healthcare systems in Africa and potential solutions. Int J Gen Med 2019;12:395-403 [FREE Full text] [doi: 10.2147/IJGM.S223882] [Medline: 31819592]
6. Mukherjee S, Adhikari A, Roy M. Machine learning approaches for cervical cancer diagnosis: a systematic review. J Med Syst 2020;44(10):210 [FREE Full text] [doi: 10.1007/s10916-020-01668-6]
7. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019;380(14):1347-1358. [doi: 10.1056/NEJMra1814259] [Medline: 30943338]
8. Kaplan RM, Milstein A. Contributions of artificial intelligence to public health. NPJ Digit Med 2019;2:61. [doi: 10.1038/s41746-019-0141-0]
9. NCI study tests performance of AI-based cervical cancer screening. National Cancer Institute. 2024. URL: https://tinyurl.com/ms83tj42 [accessed 2026-10-29]
10. Kuerbanjiang W, Peng S, Jiamaliding Y, Yi Y. Performance evaluation of large language models in cervical cancer management based on a standardized questionnaire: comparative study. J Med Internet Res 2025;27:e63626 [FREE Full text] [doi: 10.2196/63626] [Medline: 39908540]
11. Le Bonniec A, Sauvaget C, Lucas E, Nassiri A, Selmouni F. Design and validation of a chatbot-based cervical cancer screening decision aid for women experiencing socioeconomic disadvantage: user-centered approach study. JMIR Cancer 2025;11:e70251 [FREE Full text] [doi: 10.2196/70251] [Medline: 40705442]
12. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Large language models encode clinical knowledge. Nature 2023;620(7972):172-180 [FREE Full text] [doi: 10.1038/s41586-023-06291-2] [Medline: 37438534]
13. Nori H, King N, McKinney S, Carignan D, Horvitz E. Capabilities of GPT-4 on medical challenge problems. ArXiv Preprint posted online on March 20, 2023. [doi: 10.48550/arXiv.2303.13375]
14. Kreps GL. The relevance of health literacy to mHealth. Inf Serv Use 2017;37(2):123-130. [doi: 10.3233/isu-170828]
15. Schillinger D, Chittamuru D, Ramírez AS. From "infodemics" to health promotion: a novel framework for the role of social media in public health. Am J Public Health 2020;110(9):1393-1396. [doi: 10.2105/AJPH.2020.305746] [Medline: 32552021]
16. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
17. Flutter documentation. Flutter. URL: https://docs.flutter.dev/ [accessed 2026-10-29]
18. Myśliwiec K. NestJS documentation. NestJS. URL: https://docs.nestjs.com [accessed 2026-10-29]
19. Ramírez S. FastAPI documentation. FastAPI. URL: https://fastapi.tiangolo.com [accessed 2026-10-29]
20. Fernande K, Cardoso JS, Fernandes J. Transfer learning with partial observability applied to cervical cancer screening. 2017 Presented at: Pattern Recognition and Image Analysis: 8th Iberian Conference, IbPRIA 2017, Proceedings; June 20-23, 2017; Faro, Portugal. [doi: 10.1007/978-3-319-58838-4_27]
21. Abdoh SF, Abo Rizka M, Maghraby FA. Cervical cancer diagnosis using random forest classifier with SMOTE and feature reduction techniques. IEEE Access 2018;6:59475-59485. [doi: 10.1109/access.2018.2874063]
22. Wu W, Zhou H. Data-driven diagnosis of cervical cancer with support vector machine-based approaches. IEEE Access 2017;5:25189-25195. [doi: 10.1109/access.2017.2763984]
23. Khan IU, Aslam N, Alshehri R, Alzahrani S, Alghamdi M, Almalki A, et al. Cervical cancer diagnosis model using extreme gradient boosting and bioinspired firefly optimization. Scientific Programming 2021;2021:1-10. [doi: 10.1155/2021/5540024]
24. Al-Wesabi YMS, Choudhury A, Won D. Classification of cervical cancer dataset. 2018 Presented at: Proceedings of the 2018 IISE Annual Conference; May 19-22, 2018; Orlando, Florida.
25. Lu J, Song E, Ghoneim A, Alrashoud M. Machine learning for assisting cervical cancer diagnosis: an ensemble approach. Future Gener Comput Syst 2020;106:199-205. [doi: 10.1016/j.future.2019.12.033]
26. Lewis P, Perez E, Piktus A, Petroni F, Karpukhin V, Goyal N, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. Adv Neural Inf Process Syst 2020;33:9459-9474 [FREE Full text]
27. GPT-4o mini: advancing cost-efficient intelligence. OpenAI. 2024. URL: https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/ [accessed 2026-10-29]

XSL•FO
RenderX

28. Reimer N, Gurevych I. Sentence-BERT: sentence embeddings using Siamese BERT-networks. 2019 Presented at: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP); November 3-7, 2019; Hong Kong, China p. 3982-3992. [doi: 10.18653/v1/D19-1410]
29. Wang W, Wei F, Dong L, Bao H, Yang N, Zhou M. MiniLM: deep self-attention distillation for task-agnostic compression of pre-trained transformers. Adv Neural Inf Process Syst 2020:33-5788 [FREE Full text]
30. Kane A. pgvector: open-source vector similarity search for Postgres. GitHub. URL: https://github.com/pgvector/pgvector [accessed 2026-10-29]
31. OpenStreetMap contributors. Overpass API. OpenStreetMap Wiki. URL: https://wiki.openstreetmap.org/wiki/Overpass_API [accessed 2026-10-29]
32. Cavoukian A. Privacy by Design: The 7 Foundational Principles. Toronto, Ontario: Information and Privacy Commissioner of Ontario; 2009.
33. Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva, Switzerland: World Health Organization; 2021.
34. Merkel D. Docker: lightweight Linux containers for consistent development and deployment. Linux J 2014;239:2 [FREE Full text]
35. Cervical Cancer (Risk Factors). UC Irvine Machine Learning Repository. URL: https://archive.ics.uci.edu/dataset/383/cervical+cancer+risk+factors [accessed 2026-10-22]
36. Cervical-cancer-health-chatbot. GitHub. URL: https://github.com/temiolorunfemi/Cervical-cancer-health-chatbot [accessed 2026-10-20]

Abbreviations

AUC-ROC: area under the receiver operating characteristic curve
CDC: US Centers for Disease Control and Prevention
CIN: cervical intraepithelial neoplasia
HPV: human papillomavirus
LLM: large language model
mHealth: mobile health
ML: machine learning
OWASP: Open Worldwide Application Security Project
RAG: retrieval-augmented generation
SMOTE: synthetic minority oversampling technique
STI: sexually transmitted infection
SUS: System Usability Scale
UCI: University of California, Irvine
WHO: World Health Organization

Edited by M Balcarras; submitted 06.Jun.2026; peer-reviewed by T Coffin, M Honarmand; comments to author 13.Aug.2026; revised version received 20.Sep.2026; accepted 09.Oct.2026; published 06.Nov.2026.

Please cite as:
Olorunfemi T, Rashid S, Khashayar P
A Multilingual AI-Driven Conversational System for Cervical Cancer Awareness and Risk Assessment: Development and Evaluation Study
JMIR Cancer 2026;12:e103866
URL: https://cancer.jmir.org/2026/1/e103866
doi:10.2196/103866
PMID:

©Temitope Olorunfemi, Shafat Rashid, Patricia Khashayar. Originally published in JMIR Cancer (https://cancer.jmir.org), 06.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Beta Testing an mHealth Symptom Management Intervention for People With Advanced Cancer: Single-Cohort Feasibility and Acceptability Study

Sharla Wells-Di Gregorio[1,2], PhD; Tanya Smit[1,2], PhD; Sarah Baltimore[1,2], BS; Don M Benson[3,4], MD, PhD, FACP; Steve K Clinton[4,5], MD, PhD; Ashley P Davenport[4,5], MD; Kari L Kendra[4,5], PhD; Carolyn J Presley[2,4,5], MD; Donald R Marks[6], PsyD

[1]Department of Internal Medicine, Division of Palliative Medicine, The Ohio State University Wexner Medical Center, 504 McCampbell Hall North, 1581 Dodd Drive, Columbus, OH, United States

[2]Center for Research Excellence in Supportive Care (CREST), The Ohio State University Comprehensive Cancer Center – Arthur G. James Cancer Hospital and Richard J. Solove Research Institute, Columbus, OH, United States

[3]Department of Internal Medicine, Division of Hematology, The Ohio State University Wexner Medical Center, Columbus, OH, United States

[4]The Ohio State University Comprehensive Cancer Center – Arthur G. James Cancer Hospital and Richard J. Solove Research Institute, Columbus, OH, United States

[5]Department of Internal Medicine, Division of Medical Oncology, The Ohio State University Wexner Medical Center, Columbus, OH, United States

[6]Department of Advanced Studies in Psychology, Kean University, Union, NJ, United States

Corresponding Author:
Sharla Wells-Di Gregorio, PhD
Department of Internal Medicine, Division of Palliative Medicine, The Ohio State University Wexner Medical Center, 504 McCampbell Hall North, 1581 Dodd Drive, Columbus, OH, United States

Abstract

Background: People living with advanced cancer experience more frequent and severe symptoms than those with early-stage disease. Common, distressing symptoms include sleep difficulties, worry and uncertainty, fatigue, and depression. Cognitive behavioral therapy (CBT) and acceptance and commitment therapy (ACT) effectively manage these symptoms but are often too time intensive for patients with multiple appointments, limited energy, and competing priorities. Brief, mobile health (mHealth) interventions provide an accessible alternative, particularly for rural patients with limited access to palliative or psychosocial oncology services.

Objective: Building on our successful in-person and Digital Video Disk–based pilot trial of a 3-session, integrated CBT-ACT symptom management intervention for patients with advanced cancer, Finding Our Center Under Stress (FOCUS), the current study evaluated the feasibility and acceptability of a 4-session mHealth version.

Methods: In this single-cohort feasibility and acceptability study, people with advanced cancer were recruited from hospital-based oncology clinics representing 4 cancer types (breast, melanoma, multiple myeloma, and prostate). Baseline assessments included sociodemographic data, patient-reported outcomes for sleep (Insomnia Severity Index), anxiety (7-item Generalized Anxiety Disorder scale and Penn State Worry Questionnaire), fatigue (Fatigue Symptom Inventory), depression (Center for Epidemiological Studies Depression Scale), and a 7-day sleep diary completed in the app. Participants then completed 4 modules addressing sleep difficulties, worry and uncertainty, fatigue, and depression. Primary outcomes were feasibility and acceptability. Feasibility benchmarks included recruitment (50% enrollment), retention (70%), and app completion (50%). After 6 weeks, participants completed the Internet Evaluation and Utility Questionnaire, and a subset completed qualitative interviews about their experience with the FOCUS app. We report quantitative and qualitative findings and lessons learned from app development.

Results: Overall, 64.7% (11/17) of the approached patients enrolled, and 70% (7/10) completed more than half of the app. Participants rated FOCUS highly for ease of understanding (mean 3.83/4, SD 0.41), convenience (mean 3.67/4, SD 0.82), and utility (mean 3.33/4, SD 0.82). All participants would recommend the app to others and would use it again for future problems. Favorite components included patient videos and the Sleep and Worry and Uncertainty Modules. Suggested improvements included video quality (lighting and sound), sleep diary usability, and additional professional guidance.

Conclusions: The FOCUS intervention was successfully delivered via mobile technology and demonstrated feasibility and acceptability in beta testing. The FOCUS mHealth app offers an evidence-based, accessible symptom management intervention for people with advanced cancer, particularly in rural communities. Guided by participant feedback, FOCUS 2.0 will enhance video segments, add telehealth support, and expand interactive and motivational features. A randomized controlled trial will evaluate its effectiveness.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e94676) doi:10.2196/94676

KEYWORDS

cancer; mobile health; mHealth; symptom management; mobile app; sleep; anxiety; fatigue; depression; uncertainty

Introduction

People with advanced cancer experience more frequent and severe psychological and physical symptoms than people with early-stage cancer [1-4]. Four of the most common and distressing symptoms reported by people with advanced cancer are sleep difficulties, worry and anxiety, depression, and fatigue [1,2,4,5]. These symptoms co-occur and contribute to functional impairment, poorer quality of life, and mortality yet are commonly undertreated, especially in rural communities [6-8]. Effective, accessible symptom management is an essential component of comprehensive survivorship care for people with advanced cancer.

In response to the symptom management needs experienced in advanced cancer, we developed Finding Our Center Under Stress (FOCUS), a brief, integrated cognitive behavioral therapy (CBT) and acceptance and commitment therapy (ACT) symptom management intervention. We targeted sleep, anxiety, fatigue, and depression as these were common and distressing symptoms experienced by patients in our outpatient oncology–palliative medicine clinic [4]. We used the most effective and well-used evidence-based CBT-ACT strategies selected by the principal investigator (PI), who has over 20 years of clinical experience with people with advanced cancer [9-16]. We developed a brief, 3-session intervention that included 2 in-person sessions and 1 Digital Video Disk (DVD) session. The DVD included videos recorded by people with advanced cancer as well as our psychology team. We created a standardized FOCUS patient handbook and a corresponding professional therapy manual that guided the in-person components of the intervention.

Using FOCUS, participants learn skills to self-manage their symptoms and improve their ability to focus on what matters most to them despite these symptoms. We conducted a pilot randomized controlled trial with people with advanced cancer, and the results demonstrated improved sleep on both self-report [17] and physiological measures (ie, actigraphy) [18] and moderate to strong improvements in worry, depression, and fatigue interference. However, in this previous pilot study, the primary reason for noneligibility was transportation (35%), and the primary reasons for study refusal were too many appointments (35%) and being too ill (21%).

To increase access to symptom management, mobile health (mHealth) and telehealth symptom management interventions have been developed. These demonstrate positive effects on self-efficacy, symptom distress, and psychological well-being [19,20]. mHealth is "the use of mobile and wireless devices to improve health outcomes, health care services, and health research" [21]. Positive outcomes have been demonstrated for insomnia [22,23], anxiety [24,25], depression [26-31], and fatigue [25,32]. Meta-analyses indicate that mHealth behavioral interventions impact symptom reduction to an extent comparable to face-to-face interventions [33,34]. However, most of these apps focus on single symptoms, or the studies evaluating them include only people with early-stage cancer after treatment, which may benefit outcomes based on symptom regression to the mean [35,36]. Many interventions are also too lengthy (ie, 12-20 sessions) and do not focus on the specific needs and worries of people with advanced cancer.

This paper describes the translation of our FOCUS symptom management intervention to a brief, 4-module mHealth intervention. We present results of our FOCUS 1.0 beta testing with a small but diverse sample of people with advanced cancer from geographically distinct areas of Ohio. Our goal was to conduct a single-cohort trial to determine whether FOCUS 1.0 was feasible and acceptable to people with advanced cancer and identify next steps to prepare for a randomized clinical effectiveness trial of the FOCUS mHealth intervention.

Methods

Design

This was a single-cohort feasibility and acceptability study designed to evaluate trial accrual, app engagement and completion, sleep diary and questionnaire administration, and acceptability. We also completed qualitative interviews to determine strengths and areas for improvement of FOCUS 1.0.

App Translation Process

The FOCUS 1.0 app was developed between February 2021 and April 2021 based on intervention content from our previous in-person intervention [17]. At this time, there was no multi-symptom app available specifically for people with advanced cancer. Unique features of our app include tailoring to people with advanced cancer, combination of CBT and ACT approaches, and inclusion of a modified exercise program for fatigue. The app was developed pro bono by a local digital health company (Duet Health) in collaboration with the PI (SW-DG), who created the content for the previous in-person clinical trial. App questionnaires and the sleep diary were developed by Duet Health, and the text and video content was entered into the content management system by the PI. Collaborators (DMB, SC, AD, KK, CP, and DM) reviewed the app content, and iterative rounds of feedback were conducted. Materials targeted an average Flesch-Kincaid eighth-grade reading level per Microsoft Word document statistics based on an average Program for the International Assessment of Adult Competencies adult literacy score of 268 in Ohio, corresponding to reading level above the sixth-grade but below a ninth-grade reading level. This reading level was deemed appropriate given our attempt to balance accessibility with fidelity to psychosocial oncology content [37].

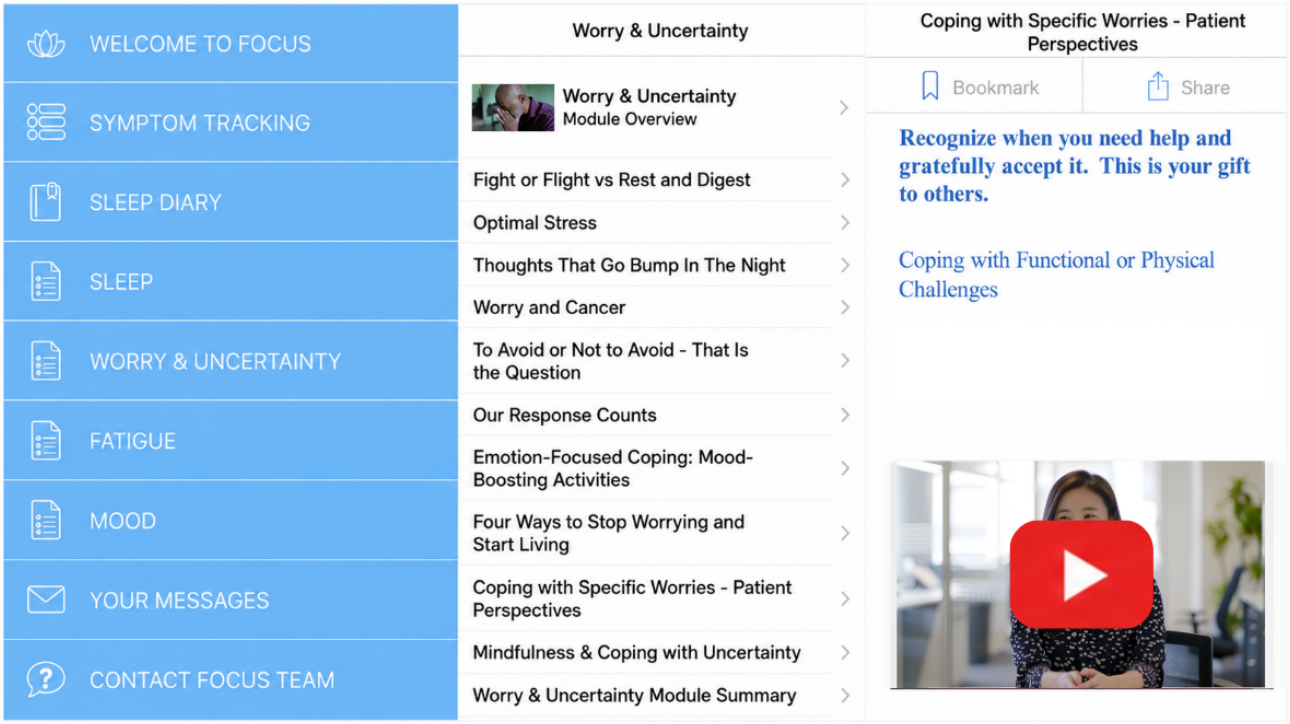
FOCUS 1.0 Intervention

FOCUS 1.0 includes 5 total hours of intervention spaced across 6 weeks (not including independent practice of intervention strategies), as well as a 5-minute welcome video, study

XSL•FO
RenderX

questionnaires, and information on how to contact the research team. The intervention components of the app are divided into 4 primary modules (Figure 1), each focused on different symptoms as follows: 1.5 hours on sleep, 1 hour on worry and uncertainty, 1.5 hours on fatigue, and 1 hour on mood. The intervention content incorporates CBT and ACT strategies by using text, imagery, videos, and interactive elements throughout the 4 symptom modules. Each module includes informational text and videos from experts to teach participants more about these symptoms and self-management strategies. Modules also include testimonial videos from people living with advanced cancer focused on sharing their experiences with symptom burden, changes in independence, financial stress, and functional challenges and how they have coped with uncertainty. Interactive and instructional relaxation and mindfulness videos are incorporated into the modules, as well as brief exercise videos from our rehabilitation oncology team to target cancer-related fatigue. In addition to the videos, the app incorporates interactive exercises and printable worksheets throughout, including values identification and constructive worry exercises.

Figure 1. Finding Our Center Under Stress (FOCUS) 1.0 (A) home screen, (B) worry module table of contents, and (C) sample worry intervention content. Original patient video image was replaced by stock photo to protect patient privacy.



Participants

The current pilot trial included 11 participants recruited from breast cancer, prostate cancer, melanoma, and multiple myeloma clinics based on recommendations from collaborating oncologists. Eligibility criteria were being aged 18 years or older and having a diagnosis of advanced-stage cancer.

Measures

Primary Outcomes

Feasibility

We determined feasibility based on potential participants' willingness to take part in the study, as well as the proportion of intervention modules completed after 6 weeks without prompting. Specifically, feasibility was defined as 50% study enrollment after recruitment (n=17), 70% retention, and completion of at least half (50%) of the app content during the 6-week enrollment period. We also assessed technology skills using the Computer Proficiency Questionnaire−12, a 12-item self-report measure that assesses an individual's perceived ability to perform common computer tasks across domains such as basic computer operations, internet use, communication, scheduling, and multimedia functions. Respondents rate their proficiency on a 5-point scale, with higher scores indicating greater computer proficiency. The measure has demonstrated excellent internal consistency (Cronbach α=0.95). We used this measure in an attempt to establish a criterion score for use in our larger randomized trial [38].

Acceptability

The Internet Evaluation and Utility Questionnaire (IEUQ) [39] measures participants' experiences and perceptions of an internet-based intervention. The IEUQ has 15 items and 3 open-ended response items. Items 1 to 8 assess ease of use, convenience, engagement, enjoyment, layout, privacy, satisfaction, and acceptability. Items 9 to 15 assess usefulness, comprehension, trustworthiness, credibility, likelihood of returning, mode of delivery, and helpfulness. Participants respond to questions on a 5-point Likert scale (0="not at all," 1="slightly," 2="somewhat," 3="mostly," and 4="very"). Scores above 2.5 were considered acceptable. Additionally, the IEUQ includes 3 items that elicit limited qualitative feedback from participants, specifically on what the participants found to be the most and least helpful parts of the app, and provides space for participants to provide suggestions to improve the

XSL•FO
RenderX

program. For this study, Cronbach α reliability (items 1-13) was 0.90.

Secondary Patient-Reported Outcomes

We included several measures to test the viability of assessment of future study patient-reported outcomes (PROs) via the app. These included the Insomnia Severity Index [40], 7-item Generalized Anxiety Disorder scale [41], Fatigue Symptom Inventory [42], Center for Epidemiological Studies Depression Scale [43], the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 [44], and the Penn State Worry Questionnaire [45].

Procedures

Data collection took place from April 2021 to August 2021. Participants were recruited through collaborating oncology clinics. Each collaborating oncologist was asked to recommend 2 to 3 patients with advanced-stage cancer to beta test the FOCUS app, and they were encouraged to recommend a diverse sample. Potential participants were contacted via phone and were told that their oncologist had recommended them for the study. They were informed that the app was designed to help manage symptoms and cope with cancer and specifically to learn strategies for improving sleep, worry, fatigue, and mood. They were advised that they could participate from home with no in-person appointments. Participants were informed that each module would take between 1 and 2 hours in addition to time to practice the strategies. Interested participants consented to the study via email or mail as per patient preference. Consented participants were then provided with a link to the app, which was available via iOS and Android using their own devices. They were provided a unique password to protect their confidentiality. They completed questionnaires at the start of app use, monitored their sleep for 1 week using the sleep diary feature of the app, and then completed 1 module each week with coping strategies to try. There was no training other than instructions for signing into the app and for completing questionnaires and exercises on the app. Participant instructions (once signed into the app) were as follows: “Complete questionnaires in the Symptom Tracker portion of the app. Then for a week, you will record your sleep for 5 minutes each morning in the Sleep Diary. The following week you will begin the first module and complete one each week starting with the Sleep Module.” The Sleep Module instructions were as follows: “As part of the Sleep Module, you create a sleep restriction schedule based on your data in the Sleep Diary. Go back to your Sleep Diary and review your data for the previous week. Add up the number of hours you slept each night for 7 days and divide by 7. Then add half an hour to this. For instance, if you sleep 6 hours each night: (6 + 6 + 6 + 6 + 6 + 6 + 6)/7 = 6. So you would set your sleep schedule for 6.5 hours with specific wake and bedtimes. Practice this for one week and add 1/2 hour if you are still tired when you wake. And practice relaxation exercises.”

Ethical Considerations

This study was approved by the Ohio State University Institutional Review Board (2010C0004). All participants were English speakers and provided informed consent via mail or email prior to starting the study.

Analyses

Given the formative, single-arm nature of this feasibility and acceptability study, our analyses were primarily descriptive rather than inferential. We used percentages to represent feasibility, including the percentage of patients approached for the study who agreed to participate, the percentage who completed the study, and the percentage of modules completed after 6 weeks without prompting. We also report the sample’s computer proficiency scores. Acceptability was examined using descriptive statistics (means and SDs on the IEUQ). Descriptive statistics are also provided on our PROs compared against other cancer sample norms when available, offering an exploratory picture of the symptomatic nature of this population; given the small sample size, these comparisons should be interpreted as hypothesis generating rather than confirmatory. We conclude our analyses with an exploration of our qualitative data from the IEUQ. Content analysis was applied to open-ended items, assessing what was most and least helpful about the app. Two members of the research team independently reviewed and coded participant responses, then conferred to reach consensus on final categories. Given that these items were brief and targeted, designed to capture usability feedback rather than support in-depth qualitative inquiry, this approach was deemed appropriate for the scope and intent of the data. Consistent with the exploratory aim of this study, the findings are therefore presented descriptively rather than as formal themes.

Results

Sample Characteristics

A total of 54.5% (6/11) of the sample was female, 72.7% (8/11) were non-Hispanic White individuals, and 72.7% (8/11) were married. The average age was 65.91 (SD 13) years. Participants were diagnosed with breast cancer (4/11, 36.4%), prostate cancer (4/11, 36.4%), melanoma (2/11, 18.2%), and multiple myeloma (1/11, 9.1%), with an average time since diagnosis of 10.55 (SD 8) years. In total, 81.8% (9/11) had prior surgery, 45.5% (5/11) had chemotherapy, 90.9% (10/11) had radiation, 63.6% (7/11) had hormonal therapy, and 18.2% (2/11) had immunotherapy (Table 1).

XSL•FO
RenderX

Table . Sample demographic and disease characteristics (N=11).

	Values
Sex, n (%)	
Female	6 (54.5)
Male	5 (45.5)
Race, n (%)	
Black	2 (18.2)
Middle Eastern	1 (9.1)
White	8 (72.7)
Age (y), mean (SD)	65.91 (13)
Time since diagnosis (y), mean (SD)	10.55 (8)
Education (y), mean (SD)	15.64 (2)
Marital status, n (%)	
Married	8 (72.7)
Divorced	2 (18.2)
Single	1 (9.1)
Support rating, n (%)	
Limited support	2 (18.2)
Adequate support	5 (45.5)
Excellent support	4 (36.4)
Employment status, n (%)	
Employed full time	1 (9.1)
Retired	7 (63.6)
Disability due to medical condition	2 (18.2)
Other (stay-at-home mother)	1 (9.1)
Financial difficulties, n (%)	
No financial difficulties	10 (90.9)
Yes, limited resources (struggling to pay bills)	1 (9.1)
Yes, limited resources (at risk of losing home)	0 (0)
RUCC[a], n (%)	
1-3	7 (63.6)
4-7	4 (36.4)
Cancer type, n (%)	
Breast	4 (36.4)
Prostate	4 (36.4)
Melanoma	2 (18.2)
Multiple myeloma	1 (9.1)
Staging, n (%)	
1	1 (9.1)
2	0 (0)
3	2 (18.2)
4	8 (72.7)
Treatments completed, n (%)	
Surgery	9 (81.8)

XSL•FO
RenderX

	Values
Radiation	10 (90.9)
Hormonal therapy	7 (63.6)
Chemotherapy	5 (45.5)
Immunotherapy	2 (18.2)
General health (SF-36[b]), n (%)	
Fair	9 (81.8)
Good	1 (9.1)
Very good	1 (9.1)
History of mental health diagnoses, n (%)	
Depression	4 (36.4)
Depression and anxiety	2 (18.2)
Panic attacks	1 (9.1)
Schizoaffective disorder	1 (9.1)

[a]RUCC: Rural-Urban Continuum Codes.

[b]SF-36: 36-item Short Form Health Survey.

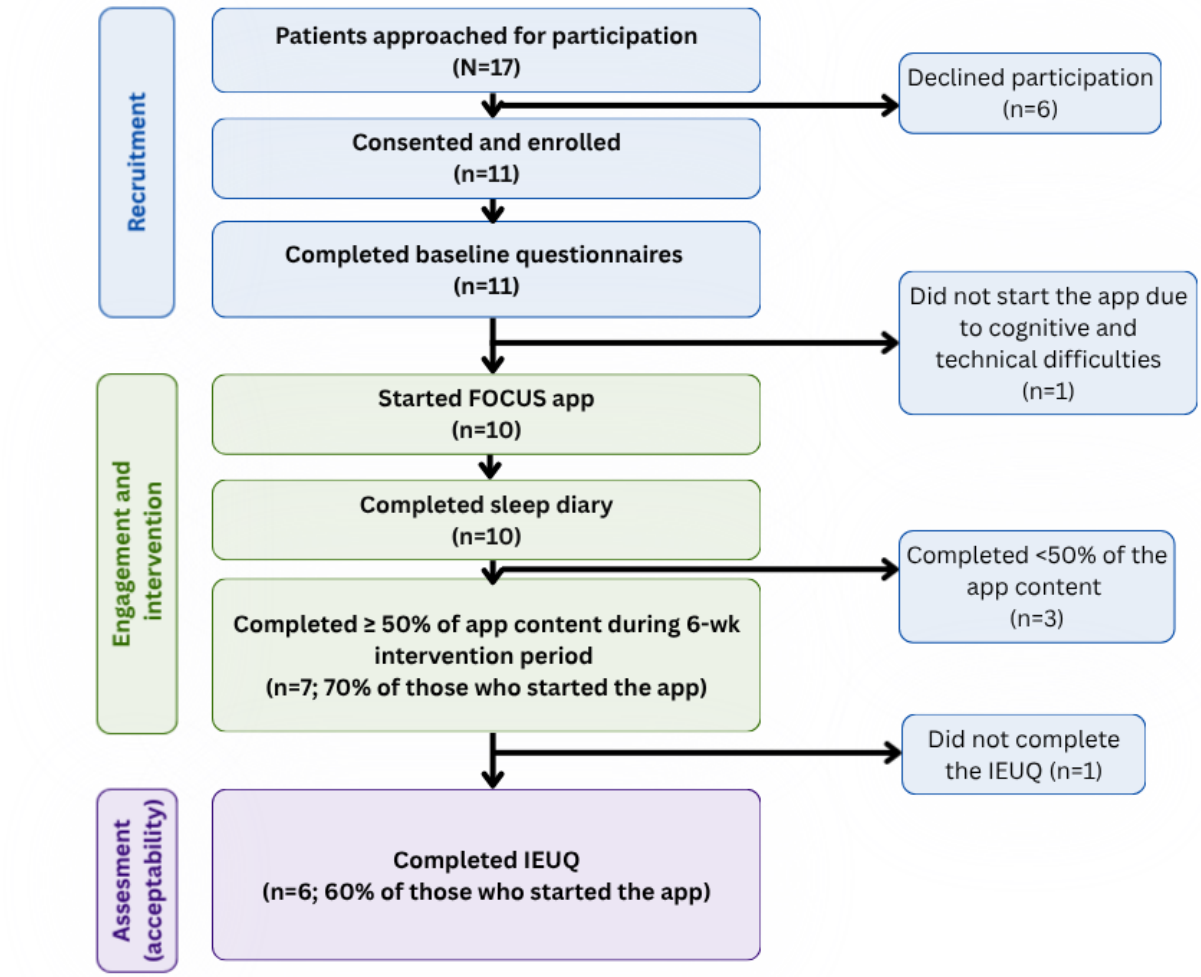
In total, 36.4% (4/11) were from rural communities (Rural-Urban Continuum Codes 4-9) [46]. Most were retired (7/11, 63.6%), although some were considered disabled due to cancer (2/11, 18.2%), and others still worked full time (1/11, 9.1%). Most (10/11, 90.9%) reported no current financial difficulties related to cancer. A total of 72.7% (8/11) had a history of mental health diagnoses, most commonly depression (4/11, 36.4%) and depression and anxiety (2/11, 18.2%). In total, 9.1% (1/11) had a history of panic attacks, and another participant (1/11, 9.1%) had schizoaffective disorder. Some (2/11, 18.2%) had received psychotropic medications, and some (2/11, 18.2%) had received both medications and counseling or therapy. Most (7/11, 63.6%) had received no mental health treatment despite the high prevalence of mental health diagnoses. Most had adequate (5/11, 45.5%) to excellent (4/11, 36.4%) support.

FOCUS Feasibility

Participant flow through FOCUS 1.0 beta testing is shown in Figure 2. Of 17 patients with advanced cancer approached for participation, 11 (64.7%) enrolled and completed the baseline questionnaires. Of these 11 participants, 1 (9.1%) was unable to engage with the intervention because of cognitive and technical difficulties and did not initiate app use, leaving 10 (90.9%) participants who started the FOCUS app and completed the sleep diary. Of these 10 participants, 7 (70%) completed more than 50% of the app content during the 6-week intervention period, meeting the predefined feasibility criterion. Those not completing 50% (of the app) (3/10, 30%) noted challenges with the sleep diary and difficulty navigating the app, and another participant did not start the app due to other family or life circumstances. A total of 60% (6/10) of the participants completed the IEUQ used to assess intervention acceptability.

XSL•FO
RenderX

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. FOCUS: Finding Our Center Under Stress; IEUQ: Internet Evaluation and Utility Questionnaire.

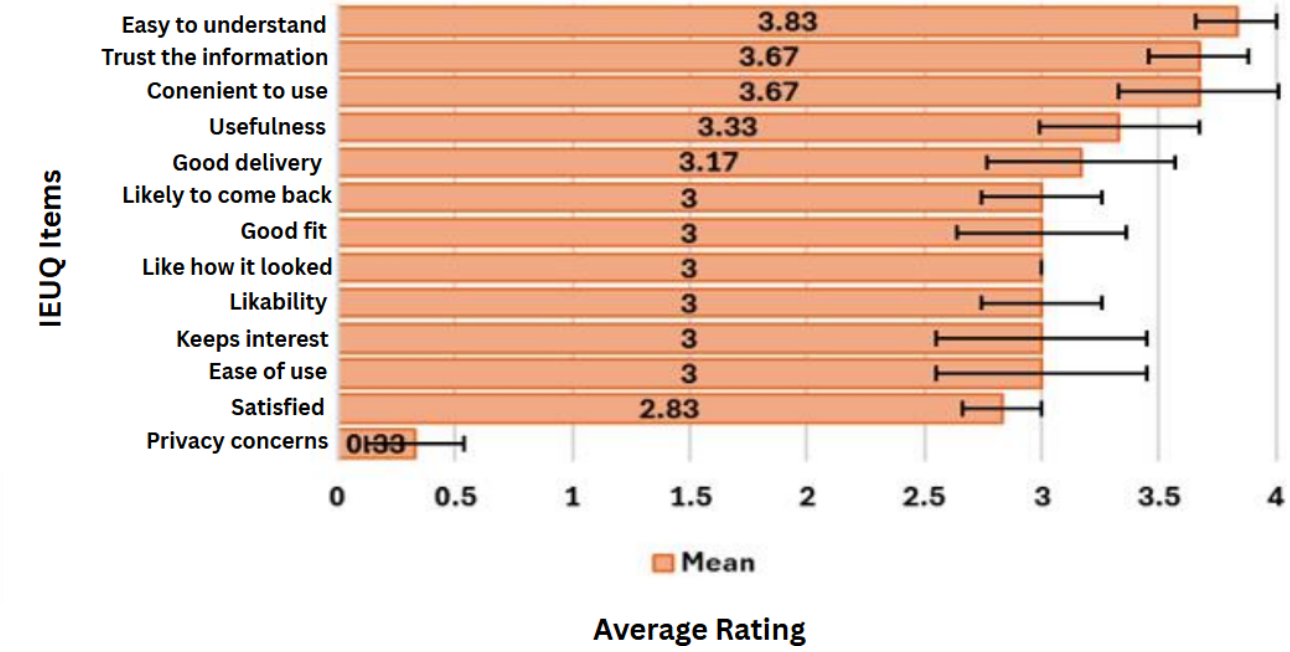


FOCUS Acceptability

All ratings of the app on the IEUQ were above 3.00 (“mostly” or “very”) other than the satisfaction rating, which was 2.83 (“somewhat” to “mostly”; Figure 3). Participants rated FOCUS highly for ease of understanding (mean 3.83/4, SD 0.41), trust (mean 3.67/4, SD 0.52), convenience (mean 3.67/4, SD 0.82), and utility (mean 3.33/4, SD 0.82).

XSL•FO
RenderX

Figure 3. Average Internet Evaluation and Utility Questionnaire (IEUQ) ratings including standard error (n=6). A total of 30% (3/10) of the participants completed <50% of the app after 6 weeks and were excluded as they did not complete the app. Another participant did not complete the IEUQ.



PROs

As anticipated, participants were highly symptomatic at baseline given their disease status. Participants scored well above the mean compared to other cancer samples on the Insomnia Severity Index, 7-item Generalized Anxiety Disorder scale, Fatigue Symptom Inventory, Center for Epidemiological Studies Depression Scale, and Penn State Worry Questionnaire (see Table 2). Participants scored below the mean on the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 for cancer samples. Computer proficiency scores were superior compared to older adults in the general population, demonstrating good computer proficiency in this sample [38].

Table . Patient-reported outcomes (PROs) at baseline.

PRO	Scale range	Cutoff scores	Sample, mean (SD)	Cancer population , mean (SD)
Insomnia (ISI)	0-28	Mean≥10	11.55 (7.09)	7.3 (6.3)-12.4 (6.7)
Anxiety (GAD–7)	0-21	Mean≥8	6.5 (5.8)	5.1 (4.0)-6.7 (5.2)
Fatigue (FSI)	1-10	Mean≥3	5.0 (2.5)	3.7 (2.2)
Depression (CES-D)	0-60	Mean≥16	19.8 (14.1)	11.6 (9.1)
Quality of life (EORTC-QLQ-C30)	0-100	Mean≤50	39.4 (13.5)	69.3 (20.7)
Worry (PSWQ)	16-80	Mean≥45	48 (15.06)	43.3(13.1)
Computer proficiency (CPQ-12)	6-30	10	25.6 (4.1)	NA

[a]ISI: Insomnia Severity Index, GAD-7: 7-item Generalized Anxiety Disorder scale, FSI: Fatigue Symptom Inventory, CES-D: Center for Epidemiological Studies Depression Scale, EORTC-QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30, CPQ-12: Computer Proficiency Questionnaire–12.

[b]Cutoffs: Insomnia Severity Index [47], 7-item Generalized Anxiety Disorder scale [48], Fatigue Symptom Inventory [49], Center for Epidemiological Studies Depression Scale [50], and European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 [51]. CPQ-12 scores < 10 are below average compared to the general population and represent users who may need additional training or support [38].

[c]Cancer population norms: Insomnia Severity Index [40,52], 7-item Generalized Anxiety Disorder scale [53,54], Fatigue Symptom Inventory [55], Center for Epidemiological Studies Depression Scale [56], and European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 [57].

[d]No official cutoff score for the PSWQ exists. The cutoff score of M≥45 is often used to indicate elevated worry [58].

[e]NA: not available.

XSL•FO
RenderX

Qualitative Feedback

Qualitative feedback was collected via 3 open-ended items from the IEUQ, which asked participants to identify the most and least helpful aspects of the app and offer suggestions for improvement. Given the targeted nature of these items, the findings are organized below by valence rather than formal thematic analysis.

Strengths

All participants would recommend the app to others and would use it again for future problems. Participants' favorite aspects of the app included the patient videos. One participant noted the following:

> *I have never participated in support groups as I am an introvert, but seeing that others are experiencing the same thoughts, feelings, and symptoms as me has been a great source of comfort.*

Participants preferred the Sleep Module and the Worry and Uncertainty Module. In particular, they appreciated learning about relaxation techniques and sleep restriction, the stretching and relaxing recording (Fatigue Module), and challenging stressful sleep thoughts. One person noted the following:

> *I would give the top rating to all aspects of the program. I cannot even begin to tell you what it has done for me. I hope that I will be able to always view the information so I can use it in the future if I start feeling down again.*

Areas for Improvement

Feedback on areas for improvement centered on technical issues, including difficulty accessing the app-based sleep diary and questionnaires; video streaming interruptions and quality issues (ie, audio and lighting); and a desire for access to professional and technical support.

Discussion

Primary Findings

In this single-cohort feasibility and acceptability study, participants with advanced cancer evaluated FOCUS 1.0, an integrated mHealth CBT-ACT symptom management intervention targeting sleep, anxiety, fatigue, and depression. FOCUS 1.0 was developed by translating the content of our previous 3-session FOCUS intervention [17] into an mHealth app, which we successfully developed over a period of 3 months. For the current study, we recruited a small but diverse sample, with higher rates of rural (4/11, 36.4%) and minority group (3/11, 27.3%) participants than the rates in our state (24% and 24%, respectively) [59,60]. We also had an impressive recruitment rate (11/17, 64.7%) for a sample of people with advanced cancer [61,62]. We did not specifically include or exclude people with a history of mental health diagnoses and were surprised to find that 72.7% (8/11) reported a history of psychological diagnoses, demonstrating our ability to recruit a sample with diverse mental health concerns. Most participants (7/10, 70%) completed at least 50% of the app content and provided overall positive acceptability ratings. These preliminary findings provide initial support for the feasibility and acceptability of FOCUS 1.0 as an intervention for patients with advanced cancer. Our PRO data are consistent with the highly symptomatic nature of this population of patients who may benefit from the app.

Participant Feedback

Participants rated ease of understanding very highly, demonstrating that our efforts to enhance readability of content with a sixth to ninth grade Flesch-Kincaid reading level were helpful. Participants also rated convenience and intervention method highly, which is a promising early signal although not conclusive evidence of reach or effectiveness. It is important to note that living with advanced cancer presents many barriers to in-person interventions, including but not limited to travel challenges, difficulty engaging due to feeling ill, and limited time due to having many medical appointments. This intervention attempts to remedy some of these barriers by providing skills and resources in an accessible format. While we cannot draw conclusions on the effectiveness of the intervention with these pilot data, participants found the content useful to address their needs, a criterion essential for motivating app use, particularly among older adults [63,64]. They trusted the content and would revisit it if problems arose again in the future.

Results from qualitative feedback highlighted several strengths of FOCUS 1.0. Participants particularly liked patient-recorded videos designed to share how other patients approach several cancer-related worries. Participants also enjoyed the interactive practice videos, including videos demonstrating relaxation techniques and physical activity to address fatigue (eg, stretching and relaxing). The Worry and Uncertainty and Sleep Modules were reported to be the most helpful. In particular, participants found utility in identifying unhelpful thoughts related to sleep [65,66] and sleep restriction.

Lessons Learned

While the current results are promising, we encountered several issues that should be noted and addressed in future studies (Table 3). First, participants experienced a variety of technical issues that resulted in understandable frustration. For example, participants noted that the sleep diary designed for the app did not auto-calculate an average time in bed sleeping, which is essential for sleep restriction and scheduling (a primary component of the Sleep Module). To address technical issues, we recommend conducting thorough testing to ensure that the components of the mHealth app are working properly and building in a virtual technical support component to address patient questions and enhance patient motivation for app completion [67-71]. Second, including the questionnaires as the initial step on the app created undue participant burden and limited enthusiasm for venturing further with the symptom modules. Additionally, we encountered a glitch with our questionnaire system such that our week 6 questionnaires were not released to participants automatically, with consequent data loss for this follow-up assessment. To address this, we recommend building the questionnaires outside of the app in a system such as REDCap or Qualtrics. In addition to reducing the perceived burden associated with the mHealth app for participants, this would reduce the workload for researchers

XSL•FO
RenderX

associated with privacy protections when data are collected within an app and potentially prevent issues with data loss. Third, participants reported a desire for additional professional support in terms of learning symptom management strategies. To address this, we have developed a telehealth professional support component to improve user motivation and experience using the app. Fourth, 9.1% (1/11) of the recruited participants were unable to engage with the intervention content due to cognitive and technological issues that we could not resolve by providing additional support. Future studies testing mHealth interventions may benefit from incorporation of cognitive and mobile proficiency screening.

Table . Finding Our Center Under Stress (FOCUS) 1.0 problems and FOCUS 2.0 planned solutions.

FOCUS 1.0 problem	FOCUS 2.0 solution
Sleep diary not working properly to calculate time in bed sleeping	Constructing a new sleep diary that auto-calculates time in bed sleeping for ease of participant use
General technical issues	Grant funding included for technical support to reduce technical barriers; use of mobile and computer proficiency tests to screen participants
Questionnaires built into app and not delivered on schedule	Questionnaires now built outside of the app in REDCap to reduce burden and privacy risks in the app and scheduled for repeat administration
Desire for more professional support	Designed a telehealth professional support protocol and manual to accompany app use
Participants with cognitive impairment	Screening for cognitive impairment as part of the prescreening process

Limitations

Several limitations should be considered when interpreting the findings of this study. First, the small sample size (N=11) limits generalizability; however, given the feasibility and acceptability focus of this study, the sample size is appropriate for generating initial evidence. Second, qualitative data were collected via brief open-ended items designed to capture user feedback on app usability rather than to support more sophisticated qualitative analysis (eg, thematic analysis). Future studies should incorporate more rigorous qualitative methods to more fully capture participant perspectives and inform iterative intervention refinement. Third, for this study, we relied on self-reported sleep measures and did not include objective monitoring devices. While the focus at this stage of app development was app user experience, including app-based sleep diary development, we recognize this as a limitation and plan to use objective sleep monitoring devices (ie, actigraphy watches) in our larger randomized clinical trial. Finally, despite higher-than-average computer proficiency, at least 30% (3/10) of the participants experienced technical difficulties. This study did not provide formal technical support for participants, which may have influenced engagement and usability outcomes. Addressing these limitations in future iterations will be important for strengthening implementation and evaluation, as well as participant intervention engagement.

Conclusions

In conclusion, although individuals coping with advanced cancer experience substantial psychological and physical symptom burden, few have access to interventions that effectively support symptom management. Findings from this feasibility and acceptability study suggest that FOCUS 1.0 shows promise in addressing these gaps in care. This study represents an initial, formative step in the iterative process of refining and evaluating the FOCUS intervention. While additional research is needed to establish the efficacy of FOCUS 1.0, this mHealth intervention has the potential to offer an accessible and scalable psychosocial treatment option for patients with advanced cancer, including those living in rural and underserved communities.

Acknowledgments

The authors would like to acknowledge their beta testing participants, whose feedback was essential to building a more engaging app for people with advanced cancer. They would also like to acknowledge Duet Health for their support with the build of Finding Our Center Under Stress (FOCUS) 1.0. During the preparation of this manuscript, the authors used ChatGPT (OpenAI) to generate Figure 2 based on the primary data provided in the text. The tool was used solely to format and visually map the data into a chart layout. After generation, the authors strictly reviewed, verified, and cross-checked the visual output against the raw manuscript data to ensure accuracy, proper formatting, and correct axis labeling. The authors take full responsibility for the contents of the final figure and the integrity of the published work.

Funding

The research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award R01CA281885 and the Ohio State University Comprehensive Cancer Center Cancer Control Grant (GR122403). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

XSL•FO
RenderX

Data Availability

Deidentified data are available on reasonable request to the corresponding author.

Authors' Contributions

Conceptualization: SW-DG, DMB, SC, KK, DM
Data curation: SW-DG
Formal analysis: SW-DG
Funding acquisition: SW-DG
Investigation: SW-DG, DM
Methodology: SW-DG, DM
Project administration: SW-DG, SB, DMB, SC, AD, KK, CP
Software: SW-DG, DM
Supervision: SW-DG
Visualization: TS, SB
Writing—original draft: SW-DG, TS
Writing—review and editing: SW-DG, TS, SB, DM

Conflicts of Interest

Duet Health developed the Finding Our Center Under Stress (FOCUS) 1.0 app evaluated in this study on a pro bono basis. SW-DG was involved in the design and development process of the app in collaboration with Duet Health, and DMB, SC, AD, KK, CP, and DM reviewed app content and provided feedback. No author received financial compensation, equity, or other financial benefit from Duet Health in connection with this work, and Duet Health had no role in the study design, data collection and analysis, interpretation of results, or decision to submit this manuscript for publication. All other authors declare no other conflicts of interest.

References

1. Chang VT, Hwang SS, Feuerman M, Kasimis BS. Symptom and quality of life survey of medical oncology patients at a veterans affairs medical center: a role for symptom assessment. Cancer 2000 Mar 1;88(5):1175-1183. [doi: 10.1002/(sici)1097-0142(20000301)88:5<1175::aid-cncr30>3.0.co;2-n] [Medline: 10699909]
2. Portenoy RK, Thaler HT, Kornblith AB, et al. Symptom prevalence, characteristics and distress in a cancer population. Qual Life Res 1994 Jun;3(3):183-189. [doi: 10.1007/BF00435383] [Medline: 7920492]
3. Mosher CE, Duhamel KN. An examination of distress, sleep, and fatigue in metastatic breast cancer patients. Psychooncology 2012 Jan;21(1):100-107. [doi: 10.1002/pon.1873] [Medline: 21105175]
4. Wells-Di Gregorio S, Porensky EK, Minotti M, et al. The James Supportive Care Screening: integrating science and practice to meet the NCCN guidelines for distress management at a Comprehensive Cancer Center. Psychooncology 2013 Sep;22(9):2001-2008. [doi: 10.1002/pon.3256] [Medline: 23436568]
5. Butt Z, Rosenbloom SK, Abernethy AP, et al. Fatigue is the most important symptom for advanced cancer patients who have had chemotherapy. J Natl Compr Canc Netw 2008 May;6(5):448-455. [doi: 10.6004/jnccn.2008.0036] [Medline: 18492460]
6. Hallgren E, Hastert TA, Carnahan LR, et al. Cancer-related debt and mental-health-related quality of life among rural cancer survivors: do family/friend informal caregiver networks moderate the relationship? J Health Soc Behav 2020 Mar;61(1):113-130. [doi: 10.1177/0022146520902737] [Medline: 32009469]
7. Butow PN, Phillips F, Schweder J, et al. Psychosocial well-being and supportive care needs of cancer patients living in urban and rural/regional areas: a systematic review. Support Care Cancer 2012 Jan;20(1):1-22. [doi: 10.1007/s00520-011-1270-1] [Medline: 21956760]
8. Andrykowski MA, Steffens RF, Bush HM, Tucker TC. Disparities in mental health outcomes among lung cancer survivors associated with ruralness of residence. Psychooncology 2014 Apr;23(4):428-436. [doi: 10.1002/pon.3440] [Medline: 24217966]
9. Dils AT, O'Keefe K, Dakka N, Azar M, Chen M, Zhang A. The efficacy of cognitive behavioral therapy for mental health and quality of life among individuals diagnosed with cancer: a systematic review and meta-analysis. Cancer Med 2024 Aug;13(16):e70063. [doi: 10.1002/cam4.70063] [Medline: 39165223]
10. Zhang L, Liu X, Tong F, et al. Cognitive behavioral therapy for anxiety and depression in cancer survivors: a meta-analysis. Sci Rep 2022 Dec 12;12(1):21466. [doi: 10.1038/s41598-022-25068-7] [Medline: 36509786]
11. Liu T, Xu J, Cheng H, et al. Effects of internet-based cognitive behavioral therapy on anxiety and depression symptoms in cancer patients: a meta-analysis. Gen Hosp Psychiatry 2022;79:135-145. [doi: 10.1016/j.genhosppsych.2022.09.003] [Medline: 36375342]

XSL•FO
RenderX

12. O'Keefe K, Chen M, Lesser KJ, DuVall AS, Dils AT. Treating mental health and quality of life in older cancer patients with cognitive behavioral therapy: a systematic review and meta-analysis. Int J Environ Res Public Health 2024 Jul 6;21(7):881. [doi: 10.3390/ijerph21070881] [Medline: 39063458]
13. Jiang X, Sun J, Song R, Wang Y, Li J, Shi R. Acceptance and commitment therapy reduces psychological distress in patients with cancer: a systematic review and meta-analysis of randomized controlled trials. Front Psychol 2023;14:1253266. [doi: 10.3389/fpsyg.2023.1253266] [Medline: 38250124]
14. Fang P, Tan L, Cui J, Yu L. Effectiveness of acceptance and commitment therapy for people with advanced cancer: a systematic review and meta-analysis of randomized controlled trials. J Adv Nurs 2023 Feb;79(2):519-538. [doi: 10.1111/jan.15543] [Medline: 36534441]
15. Li H, Wu J, Ni Q, Zhang J, Wang Y, He G. Systematic review and meta-analysis of effectiveness of acceptance and commitment therapy in patients with breast cancer. Nurs Res 2021;70(4):E152-E160. [doi: 10.1097/NNR.0000000000000499] [Medline: 33492055]
16. Fawson S, Moon Z, Novogrudsky K, et al. Acceptance and commitment therapy processes and their association with distress in cancer: a systematic review and meta-analysis. Health Psychol Rev 2024 Sep;18(3):456-477. [doi: 10.1080/17437199.2023.2261518] [Medline: 37746724]
17. Wells-Di Gregorio SM, Marks DR, DeCola J, et al. Pilot randomized controlled trial of a symptom cluster intervention in advanced cancer. Psychooncology 2019 Jan;28(1):76-84. [doi: 10.1002/pon.4912] [Medline: 30335211]
18. Wells-Di Gregorio S, Flowers S, Peng J, et al. Combined treatment with cognitive behavioral therapy for insomnia and acceptance and commitment therapy enhances objective and subjective reports of sleep in patients with advanced cancer. Psychooncology 2025 Apr;34(4):e70141. [doi: 10.1002/pon.70141] [Medline: 40204663]
19. Marthick M, McGregor D, Alison J, Cheema B, Dhillon H, Shaw T. Supportive care interventions for people with cancer assisted by digital technology: systematic review. J Med Internet Res 2021 Oct 29;23(10):e24722. [doi: 10.2196/24722] [Medline: 34714246]
20. Zhu J, Ebert L, Wai-Chi Chan S. Integrative review on the effectiveness of internet-based interactive programs for women with breast cancer undergoing treatment. Oncol Nurs Forum 2017 Mar 1;44(2):E42-E54. [doi: 10.1188/17.ONF.E42-E54] [Medline: 28222074]
21. Mobile health: technology and outcomes in low and middle income countries (mHealth). Fogarty International Center. National Institutes of Health. URL: https://www.fic.nih.gov/Programs/Pages/mhealth.aspx [accessed 2026-08-04]
22. Ritterband LM, Thorndike FP, Gonder-Frederick LA, et al. Efficacy of an internet-based behavioral intervention for adults with insomnia. Arch Gen Psychiatry 2009 Jul;66(7):692-698. [doi: 10.1001/archgenpsychiatry.2009.66] [Medline: 19581560]
23. Ström L, Pettersson R, Andersson G. Internet-based treatment for insomnia: a controlled evaluation. J Consult Clin Psychol 2004 Feb;72(1):113-120. [doi: 10.1037/0022-006X.72.1.113] [Medline: 14756620]
24. Kenardy J, McCafferty K, Rosa V. Internet-delivered indicated prevention for anxiety disorders: a randomized controlled trial. Behav Cogn Psychother 2003;31(3):279-289. [doi: 10.1017/S1352465803003047]
25. Willems RA, Bolman CA, Mesters I, Kanera IM, Beaulen AA, Lechner L. Short-term effectiveness of a web-based tailored intervention for cancer survivors on quality of life, anxiety, depression, and fatigue: randomized controlled trial. Psychooncology 2017 Feb;26(2):222-230. [doi: 10.1002/pon.4113] [Medline: 26988800]
26. Andersson G, Bergström J, Holländare F, Carlbring P, Kaldo V, Ekselius L. Internet-based self-help for depression: randomised controlled trial. Br J Psychiatry 2005 Nov;187:456-461. [doi: 10.1192/bjp.187.5.456] [Medline: 16260822]
27. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. BMJ 2004 Jan 31;328(7434):265. [doi: 10.1136/bmj.37945.566632.EE] [Medline: 14742346]
28. Clarke G, Eubanks D, Reid E, et al. Overcoming Depression on the Internet (ODIN) (2): a randomized trial of a self-help depression skills program with reminders. J Med Internet Res 2005 Jun 21;7(2):e16. [doi: 10.2196/jmir.7.2.e16] [Medline: 15998607]
29. Robertson L, Smith M, Castle D, Tannenbaum D. Using the internet to enhance the treatment of depression. Australas Psychiatry 2006 Dec;14(4):413-417. [doi: 10.1080/j.1440-1665.2006.02315.x] [Medline: 17116083]
30. Ruwaard J, Schrieken B, Schrijver M, et al. Standardized web-based cognitive behavioural therapy of mild to moderate depression: a randomized controlled trial with a long-term follow-up. Cogn Behav Ther 2009;38(4):206-221. [doi: 10.1080/16506070802408086] [Medline: 19221919]
31. Vernmark K, Lenndin J, Bjärehed J, et al. Internet administered guided self-help versus individualized e-mail therapy: a randomized trial of two versions of CBT for major depression. Behav Res Ther 2010 May;48(5):368-376. [doi: 10.1016/j.brat.2010.01.005] [Medline: 20152960]
32. Yun YH, Lee KS, Kim YW, et al. Web-based tailored education program for disease-free cancer survivors with cancer-related fatigue: a randomized controlled trial. J Clin Oncol 2012 Apr 20;30(12):1296-1303. [doi: 10.1200/JCO.2011.37.2979] [Medline: 22412149]
33. Ritterband LM, Tate DF. The science of internet interventions. Introduction. Ann Behav Med 2009 Aug;38(1):1-3. [doi: 10.1007/s12160-009-9132-5] [Medline: 19816750]

XSL•FO
RenderX

34. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of web-based vs. non-web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40. [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
35. Henselmans I, Helgeson VS, Seltman H, de Vries J, Sanderman R, Ranchor AV. Identification and prediction of distress trajectories in the first year after a breast cancer diagnosis. Health Psychol 2010 Mar;29(2):160-168. [doi: 10.1037/a0017806] [Medline: 20230089]
36. Helgeson VS, Snyder P, Seltman H. Psychological and physical adjustment to breast cancer over 4 years: identifying distinct trajectories of change. Health Psychol 2004 Jan;23(1):3-15. [doi: 10.1037/0278-6133.23.1.3] [Medline: 14756598]
37. Krenzke T, Mohadjer L, Li J, Erciulescu A, Fay R, Ren W. Program for the international assessment of adult competencies (PIAAC): state and county estimation methodology report (NCES 2020-225). : U.S. Department of Education, National Center for Education Statistics; 2020 URL: https://nces.ed.gov/pubs2020/2020225.pdf [accessed 2026-08-04]
38. Boot WR, Charness N, Czaja SJ, et al. Computer Proficiency Questionnaire: assessing low and high computer proficient seniors. Gerontologist 2015 Jun;55(3):404-411. [doi: 10.1093/geront/gnt117] [Medline: 24107443]
39. Ritterband LM, Ardalan K, Thorndike FP, et al. Real world use of an internet intervention for pediatric encopresis. J Med Internet Res 2008 Jun 30;10(2):e16. [doi: 10.2196/jmir.1081] [Medline: 18653440]
40. Savard MH, Savard J, Simard S, Ivers H. Empirical validation of the Insomnia Severity Index in cancer patients. Psychooncology 2005 Jun;14(6):429-441. [doi: 10.1002/pon.860] [Medline: 15376284]
41. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
42. Hann DM, Jacobsen PB, Azzarello LM, et al. Measurement of fatigue in cancer patients: development and validation of the Fatigue Symptom Inventory. Qual Life Res 1998 May;7(4):301-310. [doi: 10.1023/a:1024929829627] [Medline: 9610214]
43. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):385-401. [doi: 10.1177/014662167700100306]
44. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 3;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
45. Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the Penn State Worry Questionnaire. Behav Res Ther 1990;28(6):487-495. [doi: 10.1016/0005-7967(90)90135-6] [Medline: 2076086]
46. Rural-Urban Continuum Codes. National Institutes of Health National Cancer Institute. 2025. URL: https://seer.cancer.gov/seerstat/variables/countyattribs/ruralurban.html [accessed 2026-08-04]
47. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 2011 May 1;34(5):601-608. [doi: 10.1093/sleep/34.5.601] [Medline: 21532953]
48. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic meta-analysis. Gen Hosp Psychiatry 2016;39:24-31. [doi: 10.1016/j.genhosppsych.2015.11.005] [Medline: 26719105]
49. Donovan KA, Jacobsen PB, Small BJ, Munster PN, Andrykowski MA. Identifying clinically meaningful fatigue with the Fatigue Symptom Inventory. J Pain Symptom Manage 2008 Nov;36(5):480-487. [doi: 10.1016/j.jpainsymman.2007.11.013] [Medline: 18495413]
50. Eaton WW, Smith C, Ybarra M, Muntaner C, Tien A. Center for Epidemiologic Studies Depression Scale: review and revision (CESD and CESD-R). In: Maruish ME, editor. The Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Lawrence Erlbaum Associates Publishers; 2004:363-377.
51. Saini J, Bakshi J, Panda NK, Sharma M, Vir D, Goyal AK. Cut-off points to classify numeric values of quality of life into normal, mild, moderate, and severe categories: an update for EORTC-QLQ-H&N35. Egypt J Otolaryngol 2024;40:83. [doi: 10.1186/s43163-024-00642-0]
52. Schulte T, Hofmeister D, Mehnert-Theuerkauf A, Hartung T, Hinz A. Assessment of sleep problems with the Insomnia Severity Index (ISI) and the sleep item of the Patient Health Questionnaire (PHQ-9) in cancer patients. Support Care Cancer 2021 Dec;29(12):7377-7384. [doi: 10.1007/s00520-021-06282-x] [Medline: 34050799]
53. Goerling U, Hinz A, Koch-Gromus U, Hufeld JM, Esser P, Mehnert-Theuerkauf A. Prevalence and severity of anxiety in cancer patients: results from a multi-center cohort study in Germany. J Cancer Res Clin Oncol 2023 Aug;149(9):6371-6379. [doi: 10.1007/s00432-023-04600-w] [Medline: 36757620]
54. Bukhari A, Farsi A, Albaqami F, et al. Prevalence of depression and anxiety among patients with breast cancer: a cross-sectional study. World Fam Med J 2022;20(3). [doi: 10.5742/MEWFM.2022.952506]
55. Hann DM, Denniston MM, Baker F. Measurement of fatigue in cancer patients: further validation of the Fatigue Symptom Inventory. Qual Life Res 2000;9(7):847-854. [doi: 10.1023/a:1008900413113] [Medline: 11297027]
56. Giese-Davis J, Collie K, Rancourt KM, Neri E, Kraemer HC, Spiegel D. Decrease in depression symptoms is associated with longer survival in patients with metastatic breast cancer: a secondary analysis. J Clin Oncol 2011 Feb 1;29(4):413-420. [doi: 10.1200/JCO.2010.28.4455] [Medline: 21149651]

XSL•FO
RenderX

57. Hinz A, Mehnert A, Dégi C, Reissmann DR, Schotte D, Schulte T. The relationship between global and specific components of quality of life, assessed with the EORTC QLQ-C30 in a sample of 2019 cancer patients. Eur J Cancer Care (Engl) 2017 Mar;26(2):e12416. [doi: 10.1111/ecc.12416] [Medline: 26568527]
58. Fresco DM, Mennin DS, Heimberg RG, Turk CL. Using the Penn State Worry Questionnaire to identify individuals with generalized anxiety disorder: a receiver operating characteristic analysis. J Behav Ther Exp Psychiatry 2003;34(3-4):283-291. [doi: 10.1016/j.jbtep.2003.09.001] [Medline: 14972674]
59. Rural population in Ohio. America's Health Rankings. URL: https://www.americashealthrankings.org/explore/measures/pct_rural_b/OH [accessed 2026-08-04]
60. QuickFacts: Ohio. United States Census Bureau. URL: https://www.census.gov/quickfacts/fact/table/OH/RHI825224 [accessed 2026-08-04]
61. Edwards Z, Bennett MI, Petty D, Blenkinsopp A. Evaluating recruitment methods of patients with advanced cancer: a pragmatic opportunistic comparison. Int J Pharm Pract 2019 Dec;27(6):536-544. [doi: 10.1111/ijpp.12562] [Medline: 31287212]
62. Chaiviboontham S. Recruitment and retention challenges in research of palliative care in patients with advanced cancer. Prog Palliat Care 2011;19(4):185-188. [doi: 10.1179/1743291X11Y.0000000009]
63. Macedo IM. Predicting the acceptance and use of information and communication technology by older adults: an empirical examination of the revised UTAUT2. Comput Hum Behav 2017 Oct;75:935-948. [doi: 10.1016/j.chb.2017.06.013]
64. Hauk N, Hüffmeier J, Krumm S. Ready to be a silver surfer? A meta-analysis on the relationship between chronological age and technology acceptance. Comput Human Behav 2018 Jul;84:304-319. [doi: 10.1016/j.chb.2018.01.020]
65. Harvey AG, Eidelman P. Intervention to reduce unhelpful beliefs about sleep. In: Behavioral Treatments for Sleep Disorders: Academic Press; 2011:79-89. [doi: 10.1016/B978-0-12-381522-4.00008-0]
66. Belanger L, Savard J, Morin CM. Clinical management of insomnia using cognitive therapy. Behav Sleep Med 2006;4(3):179-198. [doi: 10.1207/s15402010bsm0403_4] [Medline: 16879081]
67. Titov N, Andrews G, Schwencke G, Solley K, Johnston L, Robinson E. An RCT comparing effect of two types of support on severity of symptoms for people completing internet-based cognitive behaviour therapy for social phobia. Aust N Z J Psychiatry 2009;43(10):920-926. [doi: 10.1080/00048670903179228]
68. Schueller SM, Tomasino KN, Mohr DC. Integrating human support into behavioral intervention technologies: the efficiency model of support. Clin Psychol Sci Pract 2017;24(1):27-45. [doi: 10.1037/h0101740]
69. Paxling B, Lundgren S, Norman A, et al. Therapist behaviours in internet-delivered cognitive behaviour therapy: analyses of e-mail correspondence in the treatment of generalized anxiety disorder. Behav Cogn Psychother 2013 May;41(3):280-289. [doi: 10.1017/S1352465812000240] [Medline: 22717145]
70. Burns DD, Nolen-Hoeksema S. Coping styles, homework compliance, and the effectiveness of cognitive-behavioral therapy. J Consult Clin Psychol 1991 Apr;59(2):305-311. [doi: 10.1037//0022-006x.59.2.305] [Medline: 2030192]
71. Garland A, Scott J. Using homework in therapy for depression. J Clin Psychol 2002 May;58(5):489-498. [doi: 10.1002/jclp.10027] [Medline: 11967875]

Abbreviations

ACT: acceptance and commitment therapy
CBT: cognitive behavioral therapy
DVD: Digital Video Disk
FOCUS: Finding Our Center Under Stress
IEUQ: Internet Evaluation and Utility Questionnaire
mHealth: mobile health
PI: principal investigator
PRO: patient-reported outcome

Edited by M Balcarras; submitted 04.Mar.2026; peer-reviewed by FY Chou, V Patil, WKH Mok; revised version received 17.Jul.2026; accepted 17.Jul.2026; published 16.Sep.2026.

Please cite as:
Wells-Di Gregorio S, Smit T, Baltimore S, Benson DM, Clinton SK, Davenport AP, Kendra KL, Presley CJ, Marks DR
Beta Testing an mHealth Symptom Management Intervention for People With Advanced Cancer: Single-Cohort Feasibility and Acceptability Study
JMIR Cancer 2026;12:e94676
URL: https://cancer.jmir.org/2026/1/e94676
doi:10.2196/94676

XSL•FO
RenderX

© Sharla Wells-Di Gregorio, Tanya Smit, Sarah Baltimore, Don M Benson, Steve K Clinton, Ashley P Davenport, Kari L Kendra, Carolyn J Presley, Donald R Marks. Originally published in JMIR Cancer (https://cancer.jmir.org), 16.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Experiences of a Chatbot for Supporting the Self-Management of Peripherally Inserted Central Catheter for Chemotherapy: Mixed Methods Study

Bokyung Jo[1*], MSc; Su Jin Kim[2,3*], MSc; Min Jeong Kim[2,4], MSc; Nayeon Kim[1,5], PhD; Ayoung Lee[5], BSc; Hoyoung Kim[5], MPH; Mangyeong Lee[2,3], PhD; Juhee Cho[1,3,5], PhD

[1]Department of Clinical Research Design & Evaluation, SAIHST, Sungkyunkwan University, Seoul, Republic of Korea

[2]Department of Digital Health, SAIHST, Sungkyunkwan University, Seoul, Republic of Korea

[3]Center for Clinical Epidemiology, Samsung Medical Center, Seoul, Republic of Korea

[4]Research Institution for Future Medicine, Samsung Medical Center, Seoul, Republic of Korea

[5]Cancer Education Center, Samsung Medical Center, 81 Irwon - ro, Gangnam, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Juhee Cho, PhD
Department of Clinical Research Design & Evaluation, SAIHST, Sungkyunkwan University, Seoul, Republic of Korea

Abstract

Background: A peripherally inserted central catheter (PICC) for vesicant or long-term chemotherapy is recommended for safe and sustainable drug delivery. However, maintaining its benefits requires regular and careful self-management. Although medical staff provide education and telephone consultation, proactive support accessible at any time or location remains limited. Therefore, we developed a rule-based chatbot to support PICC self-management.

Objective: This study aimed to evaluate the feasibility of a chatbot designed to support PICC self-management by examining chatbot use rate, usability, and user experience.

Methods: A mixed methods study was conducted from September to December 2022, adhering to the GRAMMS (Good Reporting of a Mixed Methods Study) guideline. Patients with cancer scheduled for PICC insertion and their caregivers were recruited, as PICC care is commonly performed by patients or cohabiting caregivers. All participants provided written informed consent. The chatbot was designed to provide structured responses based on prespecified dialog trees and to recognize users' intent using natural language processing. It was delivered through KakaoTalk and accessed on participants' personal mobile phones without requiring a separate app installation. Participants received face-to-face training at enrollment and were asked to voluntarily use the chatbot for 1 month. Baseline and postintervention surveys assessing usability were administered using paper-based questionnaires. Usage logs were collected from a secure researcher dashboard and analyzed for inquiry topics, free-text inputs, and fallback situations. Semistructured interviews were conducted approximately 1 month after the intervention during outpatient visits, with invitations by telephone, to explore participants' experiences regarding chatbot use. Quantitative data were analyzed descriptively to summarize participant characteristics, chatbot use, and usability outcomes, while qualitative interview data were analyzed using thematic analysis.

Results: A total of 56 participants were included in the final analysis (mean age 55.4 years, SD 13.7; female: n=39, 70%). Among them, 28 (50%) used the chatbot at least once. Chatbot users were younger than nonusers (51.1 vs 59.6 y; *P*=.02). Of the 25 users who agreed to log analysis, 347 inquiries were recorded; frequent topics included catheter care (126 observations), managing daily life (85 observations), symptoms (72 observations), and heparin use (55 observations). Among the 23 users who completed the usability survey, 20 (87%) reported that the chatbot was helpful for PICC-related issues. Qualitative interviews (N=56) identified 3 major benefits—information accessibility, effective guidance, and psychosocial support—while also revealing unmet needs related to conversational issues, user experience issues, and lack of personalization.

Conclusions: A rule-based chatbot designed to support PICC self-management demonstrates potential to enhance information accessibility, provide practical guidance, and offer psychosocial support. However, limitations related to conversational flexibility, interface usability, and personalization highlight the need for future development incorporating large language models. Longitudinal and multisite studies are warranted to assess sustained user engagement and clinical outcomes.

(JMIR Cancer 2026;12:e81026) doi:10.2196/81026

XSL•FO
RenderX

KEYWORDS

chatbot; chemotherapy; peripherally inserted central catheter; self-care; user experience

Introduction

A peripherally inserted central catheter (PICC) is inserted through a deep peripheral vein of the upper arm and threaded into a large vein above the right side of the heart [1]. This medical device is safe and useful for patients with cancer requiring long-term and/or frequent chemotherapy as well as a vesicant regimen with a risk of skin necrosis [2]. However, the nature of long-term administration makes PICC-related complications common in daily living. According to a prospective cohort study with 477 patients with cancer, 81 (17.0%) patients experienced PICC-related complications during the study period, resulting in an incidence rate of 1.59 per 1000 catheter days [3]. A systematic review reported that approximately 25% of central venous access devices were removed or failed before therapy was completed [4]. A meta-analysis of 11 studies reported that PICCs were associated with a 2.5 times higher risk of deep vein thrombosis compared with other central venous catheters [5].

PICC-related complications can be associated with practitioners' training and competence in PICC insertion as well as patient self-management [6-8]. Inadequate maintenance can lead to a surge in unplanned clinical encounters, including infections, bleeding, skin allergies, pain at the insertion site, and tube blockage [3,9,10]. As outpatient chemotherapy becomes increasingly common in many hospitals, the role of patients and caregivers in PICC care has become more important [11-13]. Thus, empowering patients to manage PICC lines safely after insertion helps mitigate PICC-related complications, improving treatment satisfaction [14,15]. Hospitals provide PICC-related education programs for this purpose, but these are somewhat limited in addressing the ongoing care needs or real-life challenges faced by patients.

Harnessing chatbots may be a solution for these challenges [16-18]. Chatbots simulate human interaction, offering real-time responses and guidance in text, voice, or visual formats [19]. According to a systematic review, use cases for chatbots in oncology are versatile, including screening, treatment planning, informational support related to treatment, remote patient monitoring, and emotional support [20]. A feasibility study with 9 patients with cancer reported that a chatbot was helpful in revealing serious health conditions, such as fever, skin rash, or abnormal sensitivity in the extremities [21]. A randomized controlled trial with 142 patients with breast cancer found that the perceived quality of cancer information provided by a chatbot was statistically noninferior compared with the information provided by physicians [22]. Moreover, nearly 90% of patients with cancer answered that chatbots were helpful in acquiring knowledge related to cancer or tracking treatment progress [23,24].

Although chatbot-based health care has many applications, chatbot feasibility in the context of supporting PICC self-management remains unclear [10,22,25]. The general public is unfamiliar with chatbots in the medical context owing to the relative newness of the technology compared with other digital solutions. This is particularly true for older adults, who form the majority of the population of patients with cancer. Therefore, the aim of this study was to investigate the feasibility of a chatbot for PICC self-management and determine users' characteristics, usage patterns, and satisfaction.

Methods

Design and Participants

A mixed methods study was conducted from September to December 2022, adhering to the GRAMMS (Good Reporting of a Mixed Methods Study) guideline. Quantitative analysis alone could identify who used or did not use the chatbot but could not explain the underlying reasons for these behaviors. To understand the contextual factors influencing adoption and engagement, qualitative inquiry was incorporated. Quantitative surveys and user logs provided objective data on participant characteristics, chatbot use, and usability, whereas qualitative interviews explored reasons for use or nonuse, perceived benefits, and barriers to use. This study used a sequential mixed methods design, in which quantitative surveys were conducted first, followed by qualitative interviews.

Patients with cancer and their caregivers were recruited from the Cancer Education Center of Samsung Medical Center in Seoul, South Korea. All patients scheduled to undergo PICC insertion were required to attend a face-to-face education session prior to the procedure. We consecutively contacted and recruited potential participants following each education session. As some patients with a PICC line were not scheduled to receive chemotherapy, an oncology nurse screened eligible patients and their caregivers prior to recruitment. We used convenience sampling, and individuals who voluntarily agreed to participate were enrolled in the study. Those who were 19 years or older, scheduled for PICC insertion (patients only), owned a smartphone, and had ever used KakaoTalk (most popular instant message service in Korea) were eligible for the study. However, those with any cognitive or physical problems with using a smartphone and who were scheduled for PICC line removal within a month were excluded. Following the provision of written informed consent, participants received approximately 15 - 20 minutes of training on how to access and use the chatbot. The session covered the basic user interface, methods for interacting with the chatbot, troubleshooting technical issues, and general precautions for use. Participants practiced using the chatbot under the guidance of research staff.

Chatbot Description

Participants were asked to voluntarily use a chatbot prototype to support PICC management for 1 month. The chatbot was a rule-based system that provided structured responses along prespecified dialog trees. It was technically implemented by Skelter Labs Corp (Figure 1) and integrated into KakaoTalk, a popular instant messenger in Korea [26]. The main purpose of the chatbot was to provide structured information regarding

XSL•FO
RenderX

PICC management according to user inquiries. In the welcome message, the chatbot told the participants what the chatbot can and cannot do. The chatbot clarified that it has not been designed to replace clinical practice or make clinical decisions.

Figure 1. Examples of the peripherally inserted central catheter (PICC) chatbot.



Users could ask the chatbot questions by tapping buttons listed according to prespecified dialog trees or typing their queries directly. Regarding the users' free text, the chatbot was enabled to understand the intent based on natural language processing. Dialog trees were constructed according to 6 main themes: catheter management, adverse events with the catheter, coping with emergencies, symptom management, precautions during daily life, and heparin-related inquiries. These themes were determined by reviewing records of PICC consultations in which patients had previously asked questions. The most frequently asked questions for each theme were displayed in a carousel-type list, allowing users to navigate through the available inquiry options with minimal scrolling. The chatbot used the health information about PICC management developed by advanced practice nurses in charge of patient education for PICC management. Considering the role of chatbots as social actors, health information in the chatbot was delivered through "Seulgi" ("wise" in Korean), a virtual character with the appearance of a nurse and a friendly tone. However, some information too complicated to convey through text alone was delivered via photos or videos. In addition, for irrelevant or ambiguous inquiries, the chatbot delivered fallback messages that reminded users of the chatbot's original purpose and highlighted the contact number of the cancer education center for additional help.

Measurement

Chatbot feasibility was evaluated using triangulation with quantitative usability survey, user logs, and qualitative interviews. The investigated topics included questions such as who voluntarily used chatbots, why they used them, what the most frequently asked questions were, and how satisfied users were.

Quantitative Usability Survey

Regarding the usability survey, the participants were asked to complete a questionnaire at baseline and postintervention. The baseline questionnaire collected sociodemographic information including age, sex, educational level, occupation, income, area of residence, cohabitation status, and comorbidities. The postintervention questionnaire assessed chatbot adoption, reasons for using (or not using) the chatbot, and user satisfaction. Regarding user satisfaction, a questionnaire was specifically developed by referring to existing reliable and valid user experience questionnaires [27]. The questionnaire covered ease of use (n=5), usefulness (n=7), social interaction with the chatbot (n=5), overall satisfaction (n=1), intention to recommend it to others (n=1), and intention of continuous use (n=1). The responses were interpreted on a 5-point Likert scale, with higher scores indicating greater satisfaction with chatbot usage (0=strongly disagree; 1=disagree; 2=neutral; 3=agree; 4=strongly agree).

User Logs

In terms of user logs, the collected data pertained to what buttons users tapped and the inquiries they directly typed. All user logs contained anonymized unique identifiers, which did not allow the decryption of their original identifications.

Qualitative Interview

To derive additional insights about the chatbot experience, face-to-face or telephone interviews were conducted with those who agreed to participate. The interviews, lasting 30 - 60

XSL•FO
RenderX

minutes, were based on a semistructured questionnaire including PICC-related problems during the study period, awareness of the chatbot, additional reasons for using (or not using) the chatbot, perceived benefits (only for users), and unmet needs. In-depth interviews were typically conducted in a quiet counseling room within the Cancer Education Center. For participants who were unable to revisit the hospital, interviews were conducted by telephone. Each interview was carried out by a team of 3 trained interviewers (BJ, ML, NK), including 1 oncology nurse specialist and 2 additional researchers with prior training in qualitative interviewing. During each session, 1 interviewer led the discussion, while another assisted by taking notes and managing the recording process. All interviewers had prior experience in conducting focus interviews and qualitative data collection. Participants were informed that the interviewers were members of the research team. All interviews were audio-recorded and transcribed.

Data Analyses

Descriptive statistics were used to compare characteristics between chatbot users and nonusers, identify the most frequently asked questions, and determine user satisfaction. Continuous variables were presented as means and SDs and categorical variables as frequencies and percentages. Group differences between users and nonusers were examined using independent *t* tests for continuous variables and chi-square tests for categorical variables. Graphical depictions were presented to show the main reasons for using (or not using) the chatbot. Regarding user satisfaction, all responses were dichotomized into “positive” (agree and strongly agree) and “negative” (all other responses) and summarized in frequencies. All quantitative analyses were explanatory in nature. Statistical tests were 2-sided, and statistical significance was defined using a 95% CI and a $P<.05$. Participants with missing values for a given variable were excluded from the corresponding analysis but retained in other analyses. Because the proportion of missing data was small, missingness was assumed to be missing completely at random, and no imputation was performed. All statistical analyses were performed using R (version 4.1.2; R Foundation for Statistical Computing).

For the qualitative data, we conducted thematic analyses based on Braun and Clarke’s method [28]. Three researchers (BJ, ML, MJK) independently read the transcripts, generated initial codes, and organized them into categories and themes related to participants’ motivations for chatbot use, perceived benefits, unmet needs, and contextual factors of PICC self-management. A shared coding framework and codebook were developed through iterative discussions. Using investigator triangulation, at least 3 researchers coded each transcript independently, compared results, and resolved discrepancies by consensus. The final themes were reviewed jointly to ensure coherence and distinction and by a principal investigator who audited the overall analytic process. Representative quotations were selected to illustrate key themes.

Quantitative and qualitative data were integrated during the interpretation phase using a narrative approach [29]. Qualitative findings contextualized quantitative results by explaining the reasons for chatbot use or nonuse. Integration discussions were conducted by the research team to ensure coherence between datasets.

Ethical Considerations

The study was reviewed and approved by the Institutional Review Board of the Samsung Medical Center (IRB No. SMC 2022-07-059). Written informed consent was obtained from all participants. All collected data were deidentified, and access was restricted to authorized personnel. Participants received remuneration of 30,000 KRW (approximately US $25) upon completion of the study.

Results

Participant Characteristics

During the study, 5 - 10 individuals received PICC-related education each day, and 78 individuals initially consented to participate. Among those, 56 who completed the study were included in the final analysis. Among them, 28/56 (50%) used chatbots (Table 1). Chatbot users were more likely to be younger (51.1 y vs 59.6 y; $P=.02$) and female (24/28, 86% vs 15/28, 54%; $P=.02$) compared with nonusers. The chatbot user group tended to have a higher educational level, but there was no statistical significance (19/28, 68% vs 13/28, 46%); $P=.17$). No significant differences were observed between chatbot users and nonusers in terms of the type of participants (patients vs caregivers), working status, family living together, comorbidities, income, and residence area. In the patient subpopulation, chatbot users tended to have breast cancer (13/17, 76% vs 10/20, 50%; $P=.28$), present with early-stage disease (I or II; 11/17, 65% vs 8/20, 40%; $P=.18$), and undergo curative chemotherapy (17/17, 100% vs 17/20, 85%; $P=.29$; Table 2).

XSL•FO
RenderX

Table . Participant characteristics.

	Overall (N=56)	Users (n=28)	Nonusers (n=28)	*P* value
Age (y), mean (SD)	55.4 (13.7)	51.1 (14.4)	59.6 (11.8)	.02[a]
Sex, n/N (%)				.02[b]
Male	17/56 (30)	4/28 (14.3)	13/28 (46.4)	
Female	39/56 (70)	24/28 (86)	15/28 (54)	
Participant type, n/N (%)				.57[b]
Patients	37/56 (66)	17/28 (61)	20/28 (71)	
Caregivers	19/56 (34)	11/28 (39)	8/28 (29)	
Educational level, n/N (%)				.18[b]
Below high school	24/56 (43)	9/28 (32)	15/28 (54)	
Above college	32/56 (57)	19/28 (68)	13/28 (46)	
Working status (yes), n/N (%)	27/56 (48)	16/28 (57)	11/28 (39)	.28[b]
Family living together (yes), n/N (%)	49/56 (88)	23/28 (82)	26/28 (93)	.42[b]
Chronic disease (yes), n/N (%)	33/56 (59)	18/28 (64)	15/28 (54)	.59[b]
Household monthly income (US $), n/N (%)				.25[b]
<1700	14/55 (25)	5/27 (19)	9/28 (33)	
1700 - 4200	26/55 (47)	13/27 (48)	13/28 (46)	
>4200	15/55 (27)	10/27 (37)	5/28 (19)	
Residence area, n/N (%)				.99[b]
Metropolitan	23/55 (42)	11/27 (41)	12/28 (43)	
Nonmetropolitan	32/55 (58)	16/27 (59)	16/28 (57)	

[a]Independent *t* test was used for continuous variables.

[b]Chi-square test was used for categorical variables.

XSL•FO
RenderX

Table . Cancer type and stage characteristics of the patients (n=37).

	Chatbot users (n=17)	Chatbot nonusers(n=20)	*P* value
Age (y), mean (SD)	53.9 (13.7)	61.7 (8.8)	.04[a]
Sex, n/N (%)			.08[b]
Male	1/17 (6)	7/20 (35)	
Female	16/17 (94)	13/20 (65)	
Type of cancer, n/N (%)			.28[b]
Breast	13/17 (76)	10/20 (50)	
Lung	1/17 (6)	3/20 (15)	
Lymphoma	1/17 (6)	5/20 (25)	
Others	2/17 (12)	2/20 (10)	
Stage, n/N (%)			.18[b]
I	3/17 (18)	0/20 (0)	
II	8/17 (47)	8/20 (40)	
III	5/17 (29)	9/20 (45)	
IV	1/17 (6)	3/20 (15)	
Treatment intent, n/N (%)			.29[b]
Curative	17/17 (100)	17/20 (85)	
Palliative	0/17 (0)	3/20 (15)	
Educational level, n/N (%)			.42[b]
Below high school	7/17 (41)	12/20 (60)	
Above college	10/17 (59)	8/20 (40)	
Working status (yes), n/N (%)	14/17 (82)	15/20 (75)	.89[b]
Family living together (yes), n/N (%)	14/17 (82)	19/20 (95)	.48[b]
Chronic disease (yes), n/N (%)	11/17 (65)	9/20 (45)	.39[b]
Household monthly income (US $), n/N (%)			.20[b]
<1700	5/17 (29)	8/19 (42)	
1700 - 4200	6/17 (35)	9/19 (47)	
>4200	6/17 (35)	2/19 (11)	
Residence area, n/N (%)			.99[b]
Metropolitan	5/16 (31)	7/20 (35)	
Nonmetropolitan	11/16 (69)	13/20 (65)	

[a]Independent *t* test was used for continuous variables.

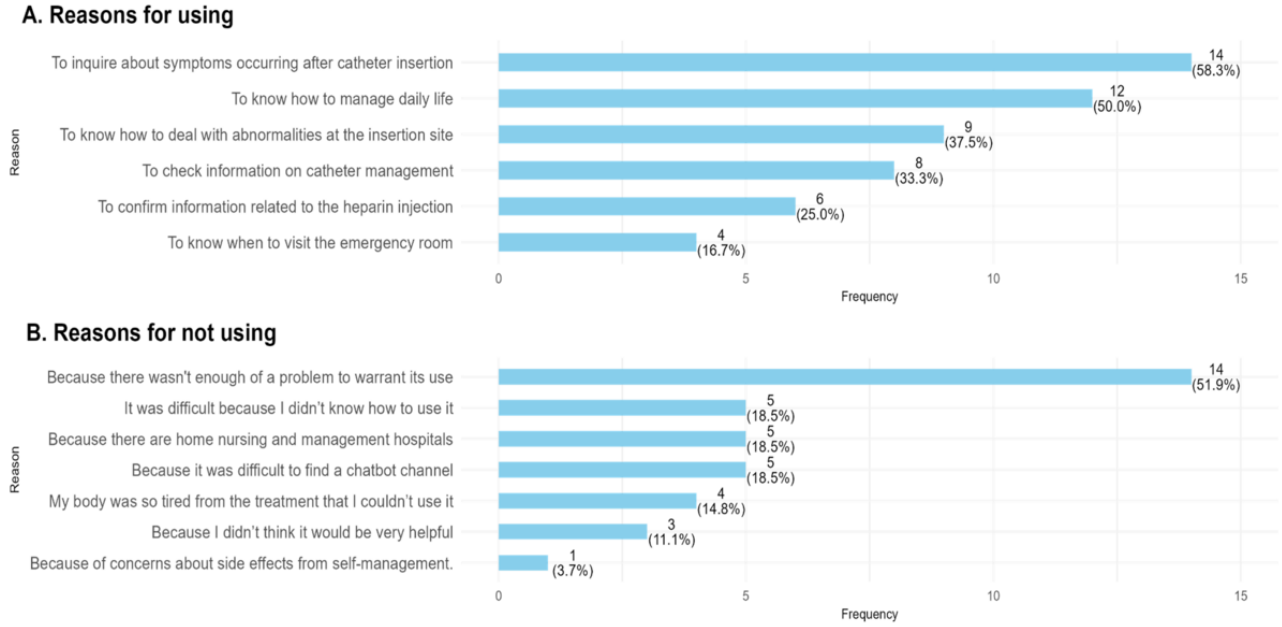
[b]Chi-square test was used for categorical variables.

Reasons for Using (or Not Using) a PICC Chatbot

Regarding the reasons for using the chatbot, 58.3% (n=14) were for inquiries about symptoms after PICC insertion, followed by learning how to manage daily life (n=12, 50%) and understanding how to deal with abnormalities at the insertion site (n=9, 37.5%; Figure 2). Additional reasons mentioned in the qualitative interviews were repetitive learning to acquire information and the management of anxiety about medical services at local hospitals (Multimedia Appendix 1).

XSL•FO
RenderX

Figure 2. Main reasons for using (or not using) the peripherally inserted central catheter (PICC) chatbot.



Among the chatbot nonusers, 14/27 (52%) did not have any issues due to their PICC (Figure 2). Additionally, 10 nonusers mentioned that they did not know how to use the PICC (n=5) or that it was difficult to access (n=5). Additional reasons mentioned in the qualitative interviews were a lack of trust in chatbots and low expectations for chatbots (Multimedia Appendix 2).

Questions Most Frequently Asked of a PICC Chatbot

A total of 25 users agreed to share their logs. During the study period, 347 inquiries were identified and categorized into 4 main themes: catheter care (126 observations), managing daily life (85 observations), symptoms (72 observations), and heparin use (55 observations; Table 3). Additionally, 9 cases were requests to contact the oncology nurse directly (Multimedia Appendix 3).

XSL•FO
RenderX

Table . User inquiries through the peripherally inserted central catheter (PICC) chatbot (n=25; 347 observations).

Themes and subthemes	observations, n (%)
Within the constructed scope	
Catheter care (126 observations)	
When to change dressing	22 (17)
Buying dressing materials	17 (14)
Intravenous medications through the catheter	15 (12)
When to remove the catheter	13 (10)
Management video	9 (7)
Change of management hospital	9 (7)
Change schedule	8 (6)
PICC part damage	4 (3)
Disinfection supplies	2 (2)
Catheter length abnormality	2 (2)
Managing daily life (85 observations)	
Workouts	24 (28)
Showers	14 (16)
Travel and flying	11 (13)
Driving	11 (13)
Wearing clothes	11 (13)
Publicly used facilities	8 (9)
Cane	2 (2)
Symptoms (72 observations)	
Fever	38 (53)
Bleeding	10 (14)
Pain at insertion site	10 (14)
Pins and needles	5 (7)
Pus	2 (3)
Rash	2 (3)
Swelling	2 (3)
Bruising	1 (1)
Heparin use (55 observations)	
Timing of heparin injection	44 (80)
Dosage	4 (7)
Refill	3 (5)
Storage	2 (4)
Chills and high fever after injection	0 (0)
Using a pump for injecting chemotherapy	0 (0)
Injection not feasible owing to catheter blockage	0 (0)
Outside the constructed scope	
Catheter care (126 observations)	
Blood pressure measurement	9 (7)
Catheter condition abnormality	7 (6)
Dressing problems	3 (2)

XSL•FO
RenderX

Themes and subthemes	observations, n (%)
How to disinfect	3 (2)
Care after catheter removal	2 (2)
Area suitable for disinfection	1 (1)
Managing daily life (85 observations)	
Painkillers	1 (1)
Manual therapy	1 (1)
Sleep	1 (1)
Hip bath	1 (1)
Symptoms (72 observations)	
Itching	1 (1)
Diarrhea	1 (1)
Heparin use (55 observations)	
Just “heparin”	2 (4)

[a]Not available.

Regarding catheter care, the most frequently asked question was about the timing for changing dressings (22/126, 17.5%), followed by buying dressing materials (17/126, 13.5%) and taking other intravenous medications through the catheter (15/126, 11.9%). Regarding the management of daily life, workout-related inquiries (24/85, 28%) were the most common, with additional concerns including showers (14/85, 16%), travel (11/85, 13%), driving (11/85, 12.9%), and wearing clothes (11/85, 13%). In terms of symptoms following catheterization, fever (38/72, 53%) was the primary concern, followed by bleeding (10/72, 13.9%) and pain at the insertion site (10/72, 14%). The timing of the heparin injection (44/55, 80%) was the most frequent heparin-related inquiry. During the study period, there were 705 interactions with the chatbot, and we identified a total of 347 user inquiries. Of these, 100 were free-text inputs, and 69 (69%) were successfully recognized by the chatbot. The 31 failed cases occurred because users did not provide sufficient context in their inquiries, entering only single terms, such as “injection,” “not detachable,” or “heparin.” As the fallback scenario was designed with limited functionality, the chatbot could not prompt users to clarify vague inputs.

Usability Results of the PICC Chatbot

Among the 28 chatbot users, 23/28 (82%) completed the questionnaire, with 22/23 (96%) agreeing that the chatbot was easy to use and 21/23 (91%) mentioning that it was useful for managing their PICC catheter (Table 4). Furthermore, 17/23 (74%) felt like they were conversing with a person during interactions with the chatbot. All participants agreed that the chatbot’s responses were trustworthy, 21/23 (91%) expressed the intention to continue using the chatbot while the catheter was in place, and 22/23 (96%) were satisfied with using the chatbot. All participants (23/23, 100%) recommended the use of chatbots for other patients using catheters.

XSL•FO
RenderX

Table . Result of the user experience survey (n=23).

Question (out of 80 points)	Mean (SD)	Negative, n (%)	Positive, n (%)
Using the chatbot was challenging (reverse coded).	3.35 (0.98)	1 (4)	22 (96)
I found the explanations provided by the chatbot easy to understand.	3.00 (1.04)	2 (9)	21 (91)
I could promptly address any mistakes made while using the chatbot.	2.96 (0.82)	1 (4)	22 (96)
The text displayed by the chatbot was easy to read.	3.17 (0.72)	0 (0)	23 (100)
I found it easy to locate the information necessary for catheter management using the chatbot.	3.00 (0.74)	0 (0)	23 (100)
The chatbot accurately understood my intentions.	2.65 (0.98)	3 (13)	20 (87)
The chatbot was helpful in addressing issues encountered during catheter management.	2.96 (1.02)	3 (13)	20 (87)
The chatbot responded promptly whenever I had a question.	3.22 (0.74)	0 (0)	23 (100)
The chatbot satisfied my curiosity regarding catheter management.	3.17 (0.89)	1 (4)	22 (96)
I consider chatbots essential for managing my catheter.	3.30 (0.63)	0 (0)	23 (100)
Using a chatbot for catheter management proved beneficial for me.	3.09 (0.95)	2 (9)	21 (91)
The chatbot helped alleviate my anxiety during catheter management.	3.17 (0.89)	1 (4)	22 (96)
The chatbot empathized with me.	2.96 (1.07)	3 (13)	20 (87)
The chatbot responded appropriately to the situation.	2.96 (0.88)	1 (4)	22 (96)
I felt like I was talking to a person through the chatbot.	2.26 (1.14)	6 (26)	17 (74)
The chatbot felt friendly during the conversation.	2.65 (0.98)	3 (13)	20 (87)
The chatbot's responses were trustworthy.	3.13 (0.69)	0 (0)	23 (100)
I would recommend that other patients with catheters give the chatbot a try.	3.39 (0.72)	0 (0)	23 (100)
I intend to keep using the chatbot for catheter management.	3.22 (1.00)	2 (9)	21 (91)
Overall, I was satisfied with chatbot use for catheter management.	3.09 (0.85)	1 (4)	22 (96)

Perceived Benefits and Unmet Needs of Using a PICC Chatbot

Overview

A total of 56 qualitative interviews were conducted. Regarding the perceived benefits of using the chatbot, 3 themes were identified from chatbot users' statements. For unmet needs, 3 themes were identified from the statements of chatbot users and nonusers. Detailed representative quotes are provided below, with a more comprehensive overview in Multimedia Appendix 4.

Perceived Benefits

In terms of the perceived benefits of using the chatbot, 3 main themes—information accessibility, effective guidance, and psychosocial support—were identified.

XSL•FO
RenderX

Information Accessibility

Users found the chatbot useful for accessing PICC-related information anytime they had questions. They felt that the information was easy to understand because it was broken down into small chunks, making it more comprehensible for them.

> *It was very helpful because I was able to receive an immediate response in real time. […] And when I needed information, I was able to simply inquire through KakaoTalk.* [Female, 60s, patient, P56]

> *I would like to say that chatbots are understood more clearly and immediately because they provide information in a brief format, unlike paper materials. […] Anyway, chatbots are understood instantly.* [Female, 40s, caregiver, P06]

Effective Guidance

Users agreed that the chatbot was useful for managing PICC-related problems as it provided clear instructions that helped the users take the necessary actions with confidence, thereby enhancing their self-efficacy in improving self-management. Some users described it as a practical guide that bridged knowledge gaps and supported self-management. Beyond self-management support, the chatbot directly connected users to appropriate health care providers when their needs exceeded the system's capabilities. A user reported that when proper PICC line management was unavailable at a nearby hospital, the chatbot facilitated immediate connection to an oncology nurse from the cancer center where she was being treated, enabling her to receive timely support.

> *I checked whether it was correct to attach the tape after it had completely dried. […] I checked this and that and also double checked, "What am I doing wrong?"* [Female, 60s, caregiver, P38]

> *When I looked at this here [chatbot], I thought I could figure out a solution.* [Male, 70s, Patient, P20]

> *After doing some searching on the chatbot, I got in touch with the hospital again. The medical staff gave me a consultation, and I made an appointment for the next day. […] The best part was that it was directly connected to the hospital.* [Female, 40s, caregiver, P06]

> *When I searched, the chatbot suggested that it might be a blood clot. So I quickly received primary treatment at the local hospital's emergency room and then came back to the main hospital to have the PICC-line removed.* [Female, 20s, caregiver, P42]

Psychosocial Support

Users felt secure while having their PICC line because they knew that the chatbot was always available to them regardless of time and place constraints. Users were easily frightened by uncertainties surrounding their PICC-related concerns. However, access to credible information helped alleviate such fears. One user said that the chatbot helped her prepare coping strategies for potential symptoms in advance, which increased her sense of relief for managing her line. Additionally, users reported feeling reassured when they learned through the chatbot that their symptom experiences were common with other patients.

> *If something went wrong, I needed to find a solution quickly, and I thought, "Who should I ask about this?," but I really relied on this kind of thing. […] I depended on it and asked questions here [chatbot], so I felt much less anxious. Definitely.* [Female, 70s, patient, P34]

> *It really helped me when I looked up things I was curious about, such as when to go to the emergency room, when to switch hospitals, and what to do if my tube falls out. Knowing this information ahead of time made me feel a little more at ease.* [Female, 60s, caregiver, P19]

> *I was looking into whether other people have these symptoms and if this is common or just something I am experiencing.* [Female, 60s, caregiver, P19]

> *I think it helped reduce my anxiety a lot. I realized I had been more scared than I expected, but after getting information from the chatbot, I thought, "I don't need to worry that much." For example, learning that catheter disinfection isn't something to be that afraid of … and also finding out that I can go about daily life and even exercise without any major issues, that was really helpful.* [Female, 30s, patient, P43]

Unmet Needs

While users recognized the benefits of the chatbot, several limitations highlighted areas for improvement to enhance user experience and functionality. Through the qualitative analysis, 3 main themes emerged regarding the chatbot's limitations: conversational issues, user experience issues, and lack of personalization.

Conversational Issues

A key challenge reported by users was that the chatbot sometimes failed to provide the appropriate answers. When the chatbot did not understand what the user was asking, instead of clarifying how to rephrase the input, it simply provided general guidance on how to return to the home menu. Additionally, some users said that the chatbot's predefined answers sometimes did not match the guidance they expected.

> *I wanted to know how many items I needed to buy, so I searched for it, but all I got was "you can buy them at *** medical equipment store," without any details.* [Female, 40s, caregiver, P06]

> *The questions were okay, but at times, the way the chatbot guided the answers did not match our intent.* [Female, 20s, caregiver, P42]

> *When I first asked the chatbot about disinfection, it said, "I can't answer that. Please go back to the beginning."* [Female, 60s, caregiver, P38]

User Experience Issues

Users encountered several challenges with the chatbot. Users often felt confused when they navigated specific information

XSL•FO
RenderX

related to their current inquiries because all contents were stacked in a chat window without divisions by information depth. For example, some users mentioned that even after being informed by health care providers about the location of instructional videos, they still struggled to find them within the chatbot. After closing the chatbot, users experienced difficulty resuming conversations with it owing to challenges locating it among existing chats on the messenger platform. Additionally, the chatbot's avatar was displayed consistently regardless of the response context, making it difficult for users to intuitively distinguish between the various contents displayed on the screen. Meanwhile, older individuals encountered difficulties with the chatbot's carousel navigation, finding the swipe-based interface unfamiliar. They suggested implementing list-based items for easier access and improved usability.

> *I wish the link to the video related to the question asked could be displayed at the top. It's hard to find.* [Male, 40s, caregiver, P39]
>
> *I'm still young, so using internet devices isn't too difficult for me. But with chatbots that require swiping sideways, both my mom and dad struggle with that. I can help them if I am with them, but I think it is generally difficult for older adults. It might be easier for them if all the content appeared as a list on one screen instead of having to swipe side to side.* [Female, 30s, caregiver, P02]
>
> *What do I need to do to enter the chatbot channel again?* [Female, 20s, caregiver, P26]
>
> *But I don't think the image was significant. I barely noticed it because it was just an image of "O."* [Female, 30s, patient, P43]

Lack of Personalization

Users felt that the chatbot did not take into account individual variability in PICC-related issues, which reduced its usefulness in clearly addressing users' concerns. For example, 1 user found that the chatbot's predefined questions and answers were finite, as they did not address trivial or personal concerns. Additionally, users pointed out that there needs to be a feature allowing them to customize the length of the responses generated by the chatbot.

> *I wish there were some educational content that addressed minor concerns. It felt awkward to call about something so trivial, like asking, "I don't have a rash, but what should I do about this itchy area?" My parent said that they felt flustered and didn't know what to do. In a situation like this, I think it would be helpful to have some guidance on what to do. I don't think there was any educational content that covered these minor concerns.* [Male, 40s, caregiver, P39]
>
> *The chatbot provides very standardized questions and answers. But real life does not involve only the issues covered in those questions, right? There are other things I am curious about as well. New questions come up from time to time, and if I could leave a question and receive personalized feedback, like in a 1:1 Q&A, that would make a difference. Instead, I just tap a button, see the same standardized answer, and that is it. That is why it didn't feel engaging or personal to me.* [Female, 30s, caregiver, P02]
>
> *To be honest, patients are not the ones doing research and studying. So I just want you to tell me, "Do it like this." Even in my mother's case, she has an impatient personality, and people are like, "So what do I need to do?"* [Male, 40s, caregiver, P47]
>
> *The patient mentioned having a fever, so I asked what to do in that case, but the answer wasn't very detailed. I wish there had been a bit more detail.* [Female, 20s, caregiver, P26]

Discussion

Principal Findings

The aim of this study was to evaluate the feasibility of a chatbot designed to support PICC self-management by examining chatbot use rate, usability, and qualitative satisfaction among patients with cancer and their caregivers, using data from survey, usage logs, and interview. Regarding chatbot adoption for PICC management, half of the participants used the chatbot during the study period. The most popular topic among chatbot users was PICC line management (eg, timing of dressing changes and purchasing dressing materials), followed by questions about adjusting to daily life, symptoms after PICC insertion, and heparinization. In terms of user experience, the 3 main benefits of the chatbot were information accessibility, effective guidance, and psychosocial support. Despite these promising findings, there were 3 key challenges related to the chatbot: conversational issues, user experience issues, and a lack of personalization.

In terms of chatbot adoption for PICC management, at least half of the participants used the chatbot during the study period. In previous studies, the adoption rates of health care chatbots ranged from 35% to 80% [30,31]. The heterogeneity in adoption rates across studies might depend on specific topics addressed by chatbots or the availability of alternative resources. For example, a study reported that 80% (n=48) of 60 patients with pancreatic cancer accessed the chatbot providing genetic information [30]. In another study examining a chatbot for chemotherapy care, 58% (n=20) of 34 patients with gastrointestinal cancer used the chatbot at least once [31]. Conversely, a chatbot for postoperative guidance after ureteroscopy was used by only 35% (n=7) of 20 patients [32]. Their lower activation rate may be attributed to the availability of multiple alternative information sources [32]. Compared with other studies, this study included both patients with cancer and their caregivers and provided nurse-led education for PICC management prior to offering the chatbot. These specific conditions might have contributed to participants experiencing a lower need to obtain PICC-related information via the chatbot. The primary motivation for chatbot utilization was related to individual needs for acquiring knowledge or information to prepare for or cope with PICC-related symptoms or issues. Notably, most nonusers reported experiencing mild or no PICC-related problems, diminishing their perceived need for the chatbot. Therefore, when assessing chatbot feasibility, it may be essential to consider whether potential users are

XSL•FO
RenderX

experiencing the specific circumstances related to their perceived needs for health information.

The most popular topic among chatbot users was PICC line management (eg, timing of dressing changes and purchasing dressing materials), followed by questions about adjusting to daily life, symptoms after PICC insertion, and regular heparinization. A study on PICC management in nursing facilities reported communication gaps, such as uncertainty about the last dressing change date or PICC removal timing, even between hospitals and nursing facilities [33]. Another study on adults with a PICC at home found that showering was the most common challenge, followed by difficulties with exercising, sleeping, and dressing [34]. The chatbot in this study was designed and implemented based on patients' information needs on PICC management in a real-world setting. The chatbot allowed users to access reliable information whenever they wanted. Although the questions and answers in the chatbot were prespecified, the topics the patients inquired about were comparable to the findings from previous studies. This suggests that chatbots could be feasible for providing informational support to patients with PICC lines. However, given the rule-based nature of chatbots, continuous expansion of the information the chatbot covers would be necessary to ensure sustainability.

In terms of the chatbot's usability, users mentioned that it was an accessible channel for effective PICC self-management guidance. In this study, the chatbot was designed to provide practical tips or information regarding PICC self-management in a structured format with succinct and plain language. The implementation incorporated patient health literacy considerations, adhering to evidence-based recommendations for health care chatbot interfaces derived from previous studies [35,36]. Users particularly appreciated the convenience and rapid access to relevant information provided by chatbots, which may have enhanced their perceived usefulness in this study [17]. In addition, the chatbot was helpful for them to receive proper psychosocial support regarding PICC management. Patients with cancer, who face disease and treatment uncertainties, seek reliable information and justify their feelings through social support networks [37]. Moreover, PICC management is unfamiliar to most patients with cancer, requiring high competency in its long-term maintenance [38]. These features of patients with cancer can increase their anxiety regarding potential complications and coping strategies for cancer-related problems [38,39]. The chatbot in this study allowed users to access reliable information about PICC-related issues without temporal or spatial constraints, addressing unmet self-care needs and providing psychosocial support. During the COVID-19 pandemic, when hospital visits and communication with health care providers were often limited, the chatbot served as a clinically meaningful alternative by offering timely, trustworthy guidance. This support helped reduce patient anxiety and might have contributed to better treatment adherence. Furthermore, although the chatbot was designed to mimic human conversation, participants did not perceive its interaction as equivalent to real human dialogue. Future research should quantitatively evaluate these supportive functions to better understand their impact on patient outcomes.

Despite the promising findings, the chatbot encountered 3 main challenges: conversational issues, user experience issues, and a lack of personalization. A natural language understanding engine was integrated into the chatbot, but it occasionally failed to accurately interpret users' intentions. The system was primarily optimized for button-click interactions rather than free-text interactions to facilitate an intuitive user experience. However, these rule-based, static responses seemed to diminish conversational naturalness and had limitations in capturing the various nuances of individual expressions for the same PICC-related issues. A systematic review found that rule-based health chatbots often limited user engagement owing to constrained input and predefined dialogues [19]. To resolve these unmet needs, incorporating large language models (LLMs), such as GPT or Claude, into rule-based systems would be necessary [40,41]. These models may provide better situational awareness and more human-like communication, which allow for personalized support in PICC self-management [41]. Given patients' documented preference for direct clinician interaction, the strategic implementation of LLMs may be required for health care chatbots to approximate human-like communication efficacy [42]. Nevertheless, LLMs may generate plausible but inaccurate responses ("hallucinations"), necessitating rigorous reliability assessment and empirical validation prior to the clinical implementation of LLM-oriented health care chatbots [43]. Independent of preceding considerations, swipe-based navigation in the chatbot impeded user engagement, particularly among older users. Swipe-based navigation was difficult for them, and they often preferred icon-based menus to reduce cognitive load [44,45]. To assist user navigation, the implementation of subtle visual affordances indicating interface functionality and hierarchically structured list-based layouts is recommended [46].

Limitations

First, although patients or their caregivers are required to attend PICC education, those who voluntarily consented to participate in this study may be more interested in using digital technologies or more motivated to self-manage. While this study did not collect demographic data of those who did not consent to participate, further understanding this in future research would provide more insight into scalable implementation. Second, user experiences were assessed after approximately 1 month of chatbot use. While PICCs are often maintained for 3-6 months, these findings primarily reflect the early phase of use rather than long-term engagement. Thus, the 1-month study period may not fully capture sustained engagement patterns. However, research has shown that many complications and difficulties tend to arise within the first few weeks following insertion, making this period particularly critical for patient education and support [47]. Hence, the data collection during this initial phase offers important insights into the most vulnerable period of PICC self-management. Nevertheless, future studies should incorporate longitudinal follow-up to the association between chatbot use and PICC-related complications, the sustainability of chatbot engagement, and its long-term effects on self-management. Third, the absence of a widely adopted, validated Korean-language instrument specifically designed for assessing chatbot usability limited comparability with existing

XSL•FO
RenderX

literature. At the time this study was conducted, chatbot-specific usability instruments were not widely adopted; therefore, similar to prior studies, usability was assessed by adapting items from existing general usability questionnaires to reflect chatbot-specific characteristics [48,49]. Subsequent studies published after the completion of our research have begun to adopt chatbot-specific instruments, such as the Chatbot usability questionnaire; however, a validated Korean-language version or culturally adapted chatbot usability measure remains unavailable. Future research would therefore benefit from the development and validation of culturally appropriate, standardized Korean-language chatbot usability instruments to enable more rigorous and comparable assessments across studies. Finally, the small sample size and single-site design limit the generalizability of the findings and the ability to draw definitive conclusions. Future studies with larger and more diverse samples will be necessary to validate quantitative findings and support more robust statistical comparisons. Moreover, because participants in this study received structured onboarding and brief training, uptake in routine clinical practice where such support may not be systematically provided could be lower. This possibility further underscores the need for implementation strategies tailored to diverse clinical environments.

Conclusions

In this study, the feasibility of a chatbot-based supportive care system for managing PICC lines in patients with cancer and their caregivers was explored. The chatbot had potential in enhancing information access, providing guidance, and offering psychosocial support. In particular, during situations such as the COVID-19 pandemic when patients were unable to visit hospitals or easily contact health care providers, the chatbot served as a valuable alternative by delivering reliable information and independently supporting the resolution of minor issues, thereby improving access to medical guidance. However, limitations in conversational quality and personalization were identified. Further quantitative research is warranted to strengthen the evidence base for the chatbot's effectiveness. Incorporating LLMs may enhance future chatbot performance and user experience.

Funding

This study was supported by the National Research Foundation of Korea grant funded by the Korean government (NRF-2021R1A2C2011083 and RS-2023-00212647).

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Authors' Contributions

Conceptualization: AL, JC, MJK, ML, NK
Data curation: ML
Formal analysis: BJ, ML, MJK, SJK
Funding acquisition: JC, ML
Institutional support and resources: JC
Investigation: AL, BJ, HK, MG, NK, SJK
Methodology: JC, ML,
Supervision: JC, ML
Visualization: BJ, ML, SJK
Writing – original draft preparation: BJ, SJK
Writing – review & editing: JC, ML
ML is the co-corresponding author (email: mangyeong.lee@gmail.com).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Other reasons for using the peripherally inserted central catheter (PICC) chatbot (n=28).
[DOCX File, 19 KB - cancer_v12i1e81026_app1.docx]

Multimedia Appendix 2
Other reasons for not using the peripherally inserted central catheter (PICC) chatbot (n=28).
[DOCX File, 20 KB - cancer_v12i1e81026_app2.docx]

Multimedia Appendix 3
Category of user inquiries through the peripherally inserted central catheter (PICC) chatbot (n=25; 347 observations).

XSL•FO
RenderX

[PNG File, 101 KB - cancer_v12i1e81026_app3.png]

Multimedia Appendix 4
Perceived benefits and unmet needs of using the peripherally inserted central catheter (PICC) chatbot (n=56).
[DOCX File, 24 KB - cancer_v12i1e81026_app4.docx]

References

1. Peripherally inserted central catheter. National Cancer Institute. URL: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/peripherally-inserted-central-catheter [accessed 2026-01-29]
2. Chopra V, Anand S, Krein SL, Chenoweth C, Saint S. Bloodstream infection, venous thrombosis, and peripherally inserted central catheters: reappraising the evidence. Am J Med 2012 Aug;125(8):733-741. [doi: 10.1016/j.amjmed.2012.04.010] [Medline: 22840660]
3. Kang J, Chen W, Sun W, et al. Peripherally inserted central catheter-related complications in cancer patients: a prospective study of over 50,000 catheter days. J Vasc Access 2017 Mar 21;18(2):153-157. [doi: 10.5301/jva.5000670] [Medline: 28218366]
4. Ullman AJ, Marsh N, Mihala G, Cooke M, Rickard CM. Complications of central venous access devices: a systematic review. Pediatrics 2015 Nov;136(5):e1331-e1344. [doi: 10.1542/peds.2015-1507] [Medline: 26459655]
5. Chopra V, Anand S, Hickner A, et al. Risk of venous thromboembolism associated with peripherally inserted central catheters: a systematic review and meta-analysis. Lancet 2013 Jul 27;382(9889):311-325. [doi: 10.1016/S0140-6736(13)60592-9] [Medline: 23697825]
6. Yi M, Kim IR, Choi EK, et al. Discomfort related to peripherally inserted central catheters in cancer patient. Asian Oncol Nurs 2017;17(4):229-236. [doi: 10.5388/aon.2017.17.4.229]
7. O'Grady NP, Alexander M, Burns LA, et al. Guidelines for the prevention of intravascular catheter-related infections. Clin Infect Dis 2011 May;52(9):e162-e193. [doi: 10.1093/cid/cir257] [Medline: 21460264]
8. Malato A, Luppino A, Pipitone R, Dona` MG, Acquaviva F, Fabbiano F. Safety and Efficacy of An Educational Program in Reducing Complications of Peripherally Inserted Central Catheters (PICCs) in Haematological Patients. Blood 2011 Nov 18;118(21):4219-4219. [doi: 10.1182/blood.V118.21.4219.4219]
9. Vidal V, Muller C, Jacquier A, et al. Prospective evaluation of PICC line related complications [Article in French]. J Radiol 2008 Apr;89(4):495-498. [doi: 10.1016/s0221-0363(08)71453-7] [Medline: 18477956]
10. Ma D, Cheng K, Ding P, Li H, Wang P. Self-management of peripherally inserted central catheters after patient discharge via the WeChat smartphone application: a systematic review and meta-analysis. PLoS One 2018;13(8):e0202326. [doi: 10.1371/journal.pone.0202326] [Medline: 30153253]
11. Park JY, Kim HL. A comprehensive review of clinical nurse specialist-led peripherally inserted central catheter placement in Korea: 4101 cases in a tertiary hospital. J Infus Nurs 2015;38(2):122-128. [doi: 10.1097/NAN.0000000000000093] [Medline: 25723834]
12. Choi JY, Kim JH. What factors are linked to profitability among hospitals?: a review on the research trends [Article in Korean]. Health Policy Manag 2013 Dec 31;23(4):397-414. [doi: 10.4332/KJHPA.2013.23.4.397]
13. Kim HJ, Yun J, Kim HJ, et al. Safety and effectiveness of central venous catheterization in patients with cancer: prospective observational study. J Korean Med Sci 2010 Dec;25(12):1748-1753. [doi: 10.3346/jkms.2010.25.12.1748] [Medline: 21165289]
14. Gallieni M, Pittiruti M, Biffi R. Vascular access in oncology patients. CA Cancer J Clin 2008;58(6):323-346. [doi: 10.3322/CA.2008.0015] [Medline: 18971486]
15. Liu X, Tao S, Ji H, Chen S, Gu Y, Jin X. Risk factors for peripherally inserted central catheter (PICC)-associated infections in patients receiving chemotherapy and the preventive effect of a self-efficacy intervention program: a randomized controlled trial. Ann Palliat Med 2021 Sep;10(9):9398-9405. [doi: 10.21037/apm-21-1848] [Medline: 34628865]
16. Whitehead DC, Mehrotra A. The growing phenomenon of "virtual-first" primary care. JAMA 2021 Dec 21;326(23):2365-2366. [doi: 10.1001/jama.2021.21169] [Medline: 34807253]
17. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit Health 2019;5:2055207619871808. [doi: 10.1177/2055207619871808] [Medline: 31467682]
18. Lee MK, Park H. Exploring factors influencing usage intention of chatbot—chatbot in financial service [Article in Korean]. J Korean Soc Qual Manag 2019;47(4):755-765. [doi: 10.7469/JKSQM.2019.47.4.755]
19. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. J Med Internet Res 2023 Feb 24;25:e40789. [doi: 10.2196/40789] [Medline: 36826990]
20. Xu L, Sanders L, Li K, Chow JCL. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR Cancer 2021 Nov 29;7(4):e27850. [doi: 10.2196/27850] [Medline: 34847056]
21. Piau A, Crissey R, Brechemier D, Balardy L, Nourhashemi F. A smartphone Chatbot application to optimize monitoring of older patients with cancer. Int J Med Inform 2019 Aug;128:18-23. [doi: 10.1016/j.ijmedinf.2019.05.013] [Medline: 31160007]

XSL•FO
RenderX

22. Bibault JE, Chaix B, Guillemassé A, et al. A chatbot versus physicians to provide information for patients with breast cancer: blind, randomized controlled noninferiority trial. J Med Internet Res 2019 Nov 27;21(11):e15787. [doi: 10.2196/15787] [Medline: 31774408]
23. Görtz M, Baumgärtner K, Schmid T, et al. An artificial intelligence-based chatbot for prostate cancer education: design and patient evaluation study. Digit Health 2023;9:20552076231173304. [doi: 10.1177/20552076231173304] [Medline: 37152238]
24. Chaix B, Bibault JE, Pienkowski A, et al. When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot. JMIR Cancer 2019 May 2;5(1):e12856. [doi: 10.2196/12856] [Medline: 31045505]
25. Campbell K, Louie P, Levine B, Gililland J. Using patient engagement platforms in the postoperative management of patients. Curr Rev Musculoskelet Med 2020 Aug;13(4):479-484. [doi: 10.1007/s12178-020-09638-8] [Medline: 32388724]
26. Kim AJ, Yang J, Jang Y, Baek JS. Acceptance of an informational antituberculosis chatbot among Korean adults: mixed methods research. JMIR Mhealth Uhealth 2021 Nov 9;9(11):e26424. [doi: 10.2196/26424] [Medline: 34751667]
27. O'Brien HL, Toms EG. The development and evaluation of a survey to measure user engagement. J Am Soc Inf Sci 2010 Jan;61(1):50-69. [doi: 10.1002/asi.21229]
28. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
29. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
30. Soley N, Klein A, Taylor CO, et al. Feasibility of the genetic information assistant chatbot to provide genetic education and study genetic test adoption among pancreatic cancer patients at Johns Hopkins Hospital. AMIA Jt Summits Transl Sci Proc 2023;2023:497-504. [Medline: 37350913]
31. Gomaa S, Posey J, Bashir B, et al. Feasibility of a text messaging-integrated and chatbot-interfaced self-management program for symptom control in patients with gastrointestinal cancer undergoing chemotherapy: pilot mixed methods study. JMIR Form Res 2023 Nov 10;7:e46128. [doi: 10.2196/46128] [Medline: 37948108]
32. Goldenthal SB, Portney D, Steppe E, Ghani K, Ellimoottil C. Assessing the feasibility of a chatbot after ureteroscopy. mHealth 2019;5:8. [doi: 10.21037/mhealth.2019.03.01] [Medline: 31019961]
33. Harrod M, Montoya A, Mody L, McGuirk H, Winter S, Chopra V. Challenges for nurses caring for individuals with peripherally inserted central catheters in skilled nursing facilities. J Am Geriatr Soc 2016 Oct;64(10):2059-2064. [doi: 10.1111/jgs.14341] [Medline: 27603747]
34. Sharp R, Xu Q, Pumpa R, et al. Supportive care needs of adults living with a peripherally inserted central catheter (PICC) at home: a qualitative content analysis. BMC Nurs 2024 Jan 2;23(1):4. [doi: 10.1186/s12912-023-01614-0] [Medline: 38163877]
35. Bennion MR, Hardy GE, Moore RK, Kellett S, Millings A. Usability, acceptability, and effectiveness of web-based conversational agents to facilitate problem solving in older adults: controlled study. J Med Internet Res 2020 May 27;22(5):e16794. [doi: 10.2196/16794] [Medline: 32384055]
36. Fitzpatrick PJ. Improving health literacy using the power of digital communications to achieve better health outcomes for patients and practitioners. Front Digit Health 2023;5:1264780. [doi: 10.3389/fdgth.2023.1264780] [Medline: 38046643]
37. Lee I, Park C. The mediating effect of social support on uncertainty in illness and quality of life of female cancer survivors: a cross-sectional study. Health Qual Life Outcomes 2020 May 19;18(1):143. [doi: 10.1186/s12955-020-01392-2] [Medline: 32429954]
38. Parás-Bravo P, Paz-Zulueta M, Santibañez M, et al. Living with a peripherally inserted central catheter: the perspective of cancer outpatients—a qualitative study. Support Care Cancer 2018 Feb;26(2):441-449. [doi: 10.1007/s00520-017-3815-4] [Medline: 28707169]
39. Källenius Edström S, Lindqvist T, Rosengren K. More benefits than problems: a study regarding patients' experiences with a PICC-line during cancer treatment. Home Health Care Manag Pract 2015;28(2):101-108 [FREE Full text]
40. Chow JCL, Li K. Developing effective frameworks for large language model-based medical chatbots: insights from radiotherapy education with ChatGPT. JMIR Cancer 2025 Feb 18;11:e66633. [doi: 10.2196/66633] [Medline: 39965195]
41. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel) 2023 Mar 19;11(6):887. [doi: 10.3390/healthcare11060887] [Medline: 36981544]
42. Clarke MA, Fruhling AL, Sitorius M, Windle TA, Bernard TL, Windle JR. Impact of age on patients' communication and technology preferences in the era of meaningful use: mixed methods study. J Med Internet Res 2020 Jun 1;22(6):e13470. [doi: 10.2196/13470] [Medline: 32478658]
43. Farquhar S, Kossen J, Kuhn L, Gal Y. Detecting hallucinations in large language models using semantic entropy. Nature 2024 Jun;630(8017):625-630. [doi: 10.1038/s41586-024-07421-0] [Medline: 38898292]
44. Cabreira AT, Hwang F. Movement characteristics and effects of GUI design on how older adults swipe in mid-air. Presented at: ASSETS '18: 20th International ACM SIGACCESS Conference on Computers and Accessibility; Oct 22-24, 2018. [doi: 10.1145/3234695.3241014]
45. Zhang S, Song J. An empirical investigation into the preferences of the elderly for user interface design in personal electronic health record systems. Front Digit Health 2024;5:1289904. [doi: 10.3389/fdgth.2023.1289904] [Medline: 38348367]

XSL•FO
RenderX

46. Khamaj A. AI-enhanced chatbot for improving healthcare usability and accessibility for older adults. Alexandria Eng J 2025 Mar;116:202-213. [doi: 10.1016/j.aej.2024.12.090]
47. Liu R, Xu H, Pu L, et al. Clinical characteristics of peripherally inserted central catheter-related complications in cancer patients undergoing chemotherapy: a prospective and observational study. BMC Cancer 2023 Sep 22;23(1):894. [doi: 10.1186/s12885-023-11413-0] [Medline: 37736715]
48. Ehara H, Ohno K, Takeshita K. Frequency of the Prader-Willi syndrome in the San-in district, Japan. Brain Dev 1995;17(5):324-326. [doi: 10.1016/0387-7604(95)00060-o] [Medline: 8579217]
49. Mokmin NAM, Ibrahim NA. The evaluation of chatbot as a tool for health literacy education among undergraduate students. Educ Inf Technol (Dordr) 2021;26(5):6033-6049. [doi: 10.1007/s10639-021-10542-y] [Medline: 34054328]

Abbreviations

GRAMMS: Good Reporting of a Mixed Methods Study
LLM: large language model
PICC: peripherally inserted central catheter

Edited by M Balcarras, N Cahill; submitted 21.Jul.2025; peer-reviewed by C Anyaegbuna, K Malale; accepted 07.Jan.2026; published 11.Feb.2026.

Please cite as:
Jo B, Kim SJ, Kim MJ, Kim N, Lee A, Kim H, Lee M, Cho J
User Experiences of a Chatbot for Supporting the Self-Management of Peripherally Inserted Central Catheter for Chemotherapy: Mixed Methods Study
JMIR Cancer 2026;12:e81026
URL: https://cancer.jmir.org/2026/1/e81026
doi:10.2196/81026

© Bokyung Jo, Su Jin Kim, Min Jeong Kim, Nayeon Kim, Ayoung Lee, Hoyoung Kim, Mangyeong Lee, Juhee Cho. Originally published in JMIR Cancer (https://cancer.jmir.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Supporting Patients With Breast Cancer and Providers Through Treatment and Survivorship: Multimethod Implementation Study of the MyJourney Platform

Isabella Herrington[1], BSc; Julie Makarski[1], MSc; Krystle Amog[2], HBSc, MSc; Leigh Hayden[1], PhD; Kate Mossman[1], PhD; Norna Abbo[3], MASc; Nicole Jedrzejko[4], MD, MSc; Fahima Osman[5,6], MD, MPH; Monika Kastner[1,2,7], PhD

[1]Research and Innovation, North York General Hospital, 4001 Leslie Street, Toronto, ON, Canada

[2]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[3]Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

[4]Division of General Surgery, University of British Columbia, Vancouver, BC, Canada

[5]Department of Surgery, North York General Hospital, Toronto, ON, Canada

[6]Department of Surgery, University of Toronto, Toronto, ON, Canada

[7]Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Monika Kastner, PhD
Research and Innovation, North York General Hospital, 4001 Leslie Street, Toronto, ON, Canada

Abstract

Background: Breast cancer is the most common cancer among Canadian women, bringing complex demands for timely decision-making, coordination of multidisciplinary care, and efficient communication between patients and providers. The increasing reliance on fragmented and noninteroperable health information systems exacerbates workflow and documentation burdens, leading to inefficiencies and gaps in continuity of care. While nurse navigation programs partially bridge these gaps, most digital platforms remain poorly integrated into provider workflows, requiring manual tracking, which results in duplicated effort and reduced efficiency. Our team developed "MyJourney" at North York General Hospital (NYGH) in Ontario. It is a digital navigation platform that supports breast cancer care throughout the entire continuum, from diagnosis to survivorship.

Objective: This study aimed (1) to map the breast cancer journey and workflow to inform the design and adaptation of MyJourney at 2 oncology settings at NYGH; (2) to identify barriers and guide local implementation; and (3) to evaluate the implementation, usability, and perceived utility of MyJourney's Clinical Navigation Tool for breast cancer care teams.

Methods: A multimethod, 3-phase study was conducted at NYGH's Breast Diagnostic Centre (BDC) and Chemotherapy Clinic (CC) in Toronto, Canada. Phase 1 involved qualitative interviews with patients with breast cancer to map their care journey and inform user-centered platform design. Phase 2 included preimplementation interviews with nurses, pharmacists, and administrative staff to map workflows and customize MyJourney for the CC. Phase 3 evaluated MyJourney's implementation and usability over 6 weeks using the System Usability Scale, technology acceptance model surveys, and follow-up interviews with platform users. Data were analyzed via interpretive description and descriptive statistics. Ethics approval was obtained.

Results: In total, 13 patient interviews revealed distinct challenges and communication needs across prediagnosis, diagnosis, treatment, and survivorship phases, emphasizing the need for personalized, integrated resources (phase 1). A total of 8 participants (3 nurses and 5 pharmacists), providers, and clinic staff highlighted pain points in fragmented information systems, inefficient manual processes, and limited team coordination (phase 2). MyJourney's phased implementation led to high user acceptance, with mean System Usability Scale scores rated "excellent" (BDC: 81.3; CC: 86.3) at the 6-week follow-up (phase 3). The 4 participants (1 charge nurse, 1 administrative staff at the BDC, and 2 charge nurses at the CC) described the platform as intuitive, efficient, and well-organized, citing consolidated patient records, streamlined appointment management, and improved workflow as major benefits. Recommendations included improved interoperability, enhanced notifications, role-specific customization, and integration with electronic medical records for broader scalability.

Conclusions: The iterative, interest-holder engaged design and phased implementation of MyJourney facilitated rapid uptake and high usability among breast cancer provider teams. The Clinical Navigation Tool component of the MyJourney platform reduced documentation burden, improved workflow efficiency, and facilitated care coordination across oncology settings.

(JMIR Cancer 2026;12:e87973) doi:10.2196/87973

XSL•FO
RenderX

KEYWORDS

breast cancer; qualitative methods; usability testing; implementation; patient journey

Introduction

Breast cancer is the most common cancer in Canadian women, apart from nonmelanoma skin cancers, and is the second leading cause of cancer-related deaths [1]. Approximately 1 in 8 Canadian women will develop breast cancer during their lifetime [1,2]. Although multidisciplinary care and systemic therapies have improved breast cancer outcomes over the past decade [3], these advances have introduced new complexities for both patients and health care providers.

Patients are now expected to make timely and informed decisions, interpret detailed medical information, and coordinate care across multiple services and providers [2,4-6]. This can lead to confusion about procedures, points of contact, and treatment-related side effects, sometimes resulting in delays in care [6]. For health care providers, these complexities increase the demand for efficient documentation, coordinated communication, and real-time access to patient information [7,8]. Clinical data are often fragmented across noninteroperable health information systems, and care processes are heterogeneous and complex to manage, resulting in high levels of administrative and role-related burden [7,8].

Digital health platforms offer a promising solution to streamline care delivery, reduce manual workload, and enhance communication among health care team members [9]. However, most available platforms are not designed for provider use and are poorly integrated into clinical workflows [10]. This results in providers relying on manual tracking, which can lead to redundant documentation and time-consuming phone-based follow-up [9]. Although nurse navigation programs help bridge some of these challenges in oncology practice by supporting patients throughout their treatment journey, the model is labor-intensive [11,12]. Navigators are often required to contact patients at multiple points in their journey, without standardized tools to streamline navigation tasks or optimize workflow within and across clinical settings [11,12].

One strategy to support health care teams and reduce administrative burden is to leverage integrated digital platforms that consolidate patient care information in a single system. Digital platforms can reduce the need to search across multiple sources and allow providers to focus on direct patient care [13]. Platforms that coordinate tasks across teams can also reduce fragmentation within the health care system and enhance care continuity [13].

To address these challenges, our team developed "MyJourney," a digital oncology platform that supports breast cancer care across the continuum, from diagnosis through survivorship. The platform was initially developed by a team of software developers and design experts for North York General Hospital (NYGH) in Ontario, Canada. Following prototype development, the platform underwent iterative refinement through a user-centered design process that engaged a broad group of interest-holders, including breast cancer providers, patients, researchers, app developers, and design experts. Interest-holder engagement began in 2017, and platform development launched in 2019. MyJourney includes three integrated components: (1) a Clinical Navigation Tool that allows care teams (including nurses, pharmacists, oncologists, and surgeons) to triage, monitor, and document patient care; (2) a Multidisciplinary Cancer Conference Tool that facilitates collaborative case planning, treatment decision-making, and reporting; and (3) a patient-facing mobile app that provides personalized treatment summaries, tailored education, appointment reminders, and real-time updates to promote engagement and self-management. The platform was designed to address dual needs. First, it streamlines workflows, reduces documentation burden, and supports team-based coordination for providers. Second, it centralizes health information, treatment plans, and appointments into one accessible location for patients diagnosed with breast cancer. These functions align with broader health system priorities, including those reflected in the Quintuple Aim, by aiming to improve both provider experience and patient outcomes and experiences [14].

The objectives of our work were (1) to understand the breast cancer journey, gaps in care, and the needs of patients at the Breast Diagnostic Centre (BDC) to inform the design of the patient-facing component of MyJourney; (2) to map workflows, identify barriers, and guide local adaptation of MyJourney before its launch at the NYGH Chemotherapy Clinic (CC); and (3) to evaluate the implementation, usability, and perceived utility of the Clinical Navigation Tool component of MyJourney at both the BDC and CC.

Methods

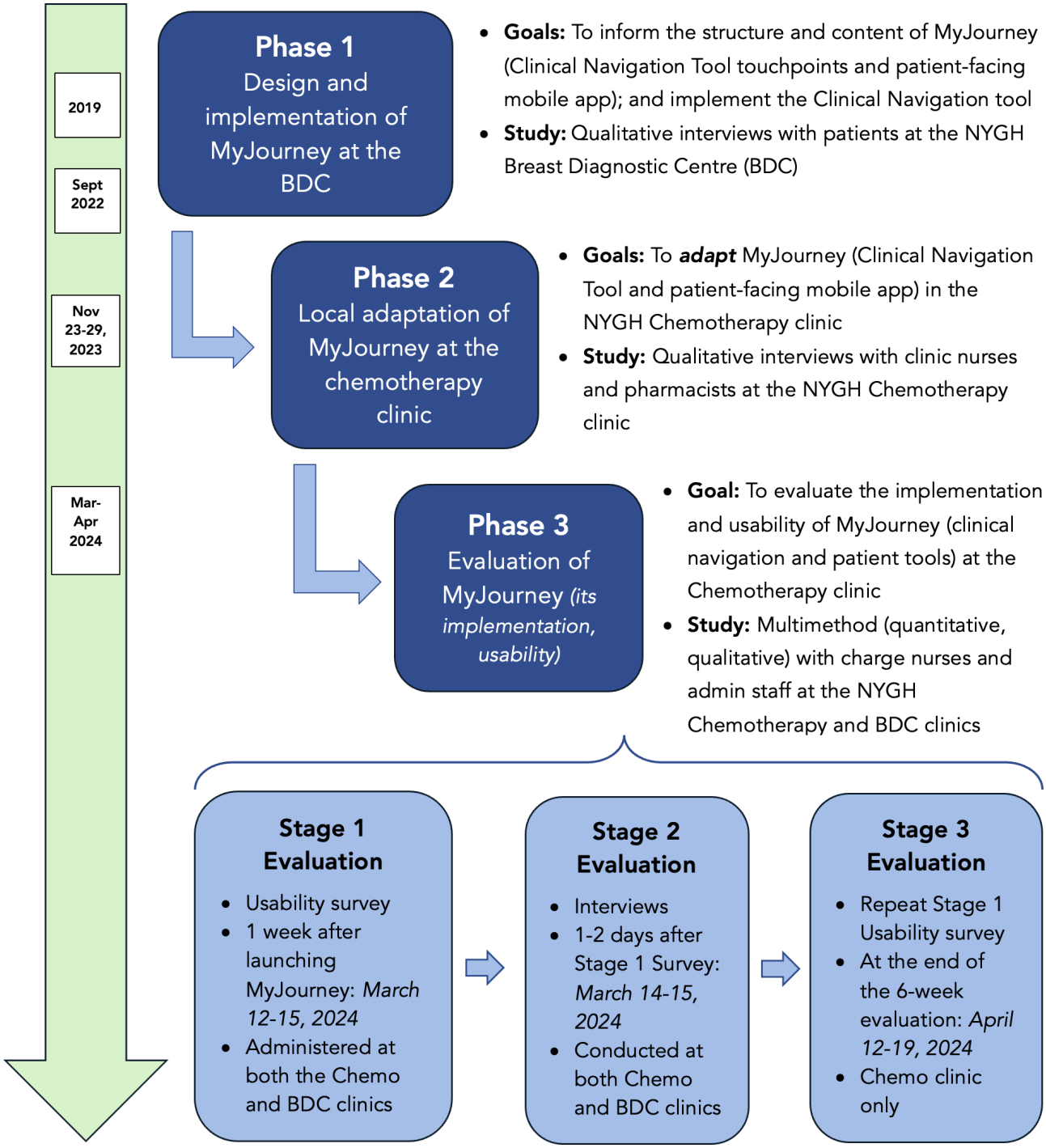
Study Design

Overview

We conducted a 3-phase, multimethod study that used both qualitative and quantitative methods (Figure 1). Phase 1 patient interviews (2018 - 2019) informed the initial design of MyJourney's Clinical Navigation Tool and patient-facing app. The Clinical Navigation Tool was first implemented at the BDC in September 2022, providing a mature implementation in a diagnostic-focused setting. Lessons learned from this implementation (eg, around appointment visibility, task lists, and navigation touchpoints), together with phase 2 preimplementation workflow interviews at the CC, informed the subsequent localization and configuration of MyJourney for the treatment-focused CC in 2024. More specifically, the below sections highlight the design for each of the 3 phases.

XSL•FO
RenderX

Figure 1. Study design schema. NYGH: North York General Hospital.



Phase 1

To inform the structure and content of MyJourney (Clinical Navigation Tool touchpoints and patient-facing mobile app), we conducted qualitative interviews with patients at the NYGH BDC to map the breast cancer journey from prediagnosis through survivorship and to inform the structure and content of MyJourney (Clinical Navigation Tool and patient-facing mobile app). We used the interpretive description of Thorne [15], grounded in naturalistic inquiry and recognizing the constructed and contextual nature of human experience. It seeks applied understanding that supports practical decision-making [15] and centers on participants' experiential accounts rather than testing or building theory, as it is grounded in practical relevance [15].

Phase 2

To adapt MyJourney to the NYGH CC, we conducted preimplementation interviews with staff to map workflows and staff roles, identify barriers, and tailor MyJourney's features to support local adaptation (November 23 - 29, 2023). We used the same interpretive description approach as in phase 1, which uses low-inference interpretation (ie, description) to enable the researcher to remain close to the data [15]. In this approach, facts are presented as a comprehensive summary in everyday language from the participants' perspectives, serving both as

XSL•FO
RenderX

an output of the research and as recommendations for future research [15-17]. We used qualitative results to iteratively implement the Clinical Navigation Tool component of MyJourney (December 2023-February 2024). This involved installing the MyJourney technology at the CC (integrating within their electronic medical record system) and training users (nurses and administration staff) on how to use the system. Once installed, MyJourney launched for use on March 4, 2024.

Phase 3

To evaluate the implementation, usability, and perceived utility of MyJourney's Clinical Navigation Tool at the CC, we conducted a 6-week, iterative, 3-staged evaluation. One week after implementing the Clinical Navigation Tool, we administered a usability survey to MyJourney users (charge nurses and administrative staff; stage 1). After 1 - 2 days, we conducted one-on-one interviews with these participants to gain a more in-depth understanding of their experiences, including workflow alignment (using survey findings as a starting point), and to identify any challenges and suggested recommendations to address them (stage 2). Over the next 5 weeks, we used survey and interview results to further refine the Clinical Navigation Tool. At 6 weeks of follow-up, we readministered the stage 1 usability survey to determine whether iterative improvements translated into enhanced usability (stage 3). We also administered the usability survey (stage 1) and interviews (stage 2) to MyJourney users at the BDC clinic to get an in-depth understanding of their user experience at that site since its implementation (September 2022). For interviews, a qualitative description approach was used to enable the researcher to remain close to the data [16]. In this approach, facts are presented as a comprehensive summary in everyday language from the participants' perspectives, serving both as an output of the research and as recommendations for future research [16]. This overall evaluation approach was guided by the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) implementation framework, which considers individual, organizational, and contextual factors to shape the implementation and sustainability of innovations [18]. For phase 3, RE-AIM primarily informed our focus on implementation (usability, workflow fit, and fidelity of use) and maintenance (patients' and clinicians' perceived likelihood of continued use and ongoing integration into routine care). Reporting of this study was guided by the SRQR (Standards for Reporting Qualitative Research) [19].

Ethical Considerations

This study was conducted at North York General Hospital, diagnostic breast cancer and chemotherapy clinics. Informed consent was obtained from all participants across the 3 phases. All data were anonymized using study ID numbers and stored in secure locations. Data collection occurred between 2019 and 2024. Ethics approval was granted by the NYGH Research Ethics Board (protocol #18 - 001 and #0326).

Setting, Intervention, and Implementation

Our study took place at 2 ambulatory oncology settings of the NYGH (a tertiary-level, academic-affiliated community hospital in Toronto, Ontario, Canada): the BDC and the CC [20]. The BDC is a specialized ambulatory diagnostic service within the Breast Centre that provides comprehensive assessment for individuals with suspected or confirmed breast cancer, identified through abnormal clinical examination and/or breast imaging, serving approximately 1500 new patients seen annually. In contrast, the CC is an outpatient systemic therapy unit that delivers chemotherapy and supportive care to patients with confirmed cancer diagnosis (including early-stage, locally advanced, and metastatic breast cancer) coordinated by charge nurses following consultation with medical oncology, with pharmacists providing patient education and counseling. Whereas the BDC primarily serves patients in the diagnostic and decision-making phase (often pretreatment), the CC primarily serves patients in active treatment, with more advanced or definitively staged disease and correspondingly different informational, supportive, and symptom-management needs. Our implementation work, therefore, involved adapting the intervention from a diagnostic-focused setting (BDC) to a treatment-focused setting (CC), taking into account differences in patient flow, acuity (eg, treatment-related toxicities), and longitudinal relationships with the oncology team. The focus of our evaluation was the Clinical Navigation Tool within the MyJourney platform. The tool was first implemented at the BDC on September 19, 2022, as part of a hospital-wide initiative to improve cancer care delivery. It was subsequently adapted for use at the NYGH CC on March 4, 2024. The BDC manages approximately 700 patients with breast cancer per year.

Participant Recruitment

Phase 1 (BDC Patients)

We used purposive sampling with convenience and self-selection elements to identify eligible participants: English-speaking adults diagnosed with breast cancer who were former or current patients receiving care at the BDC and able to provide informed consent [21]. To obtain a wide range of patient experiences, recruitment was conducted in a multipronged approach. For example, we recruited participants at the end of a 1-day workshop for newly diagnosed patients by displaying posters in the clinic waiting rooms. Second, we contacted patients who had previously indicated their willingness to participate in any future research studies. All participants spoke English and provided informed consent.

Phases 2 and 3 (Providers and Staff)

Across phases 2 and 3, we used purposive sampling within each clinic [21] to invite all current or prospective MyJourney users (nurses, pharmacists, and administrative staff) at the BDC and the CC, with the goal of capturing the full range of user perspectives rather than thematic saturation [22]. We did not recruit oncologists or surgeons for phase 2 because at the CC, these providers primarily interact with MyJourney through the Multidisciplinary Cancer Conference Tool and the hospital electronic medical record rather than as routine users of the CC Clinical Navigation Tool. Charge nurses and oncology pharmacists are the primary day-to-day users of the Clinical Navigation Tool, responsible for coordinating chemotherapy visits, managing treatment-related symptoms, and delivering patient education and counseling. Potential participants were invited via targeted email with support from senior clinic

XSL•FO
RenderX

leaders. Written informed consent and demographic data were collected using Qualtrics. One participant did not consent to the use of direct quotes.

Data Collection

Across phases, data collection included semistructured interviews or online surveys. The research team (IH, JM, MK, LH, and KM) developed and pilot-tested the instruments to ensure clarity, appropriateness, and alignment with the study objectives.

Phase 1

A semistructured interview guide was created and modified iteratively (Multimedia Appendix 1). In interpretive description, insights from early interviews inform the design of subsequent interviews, including the interview guide, which begins with broad, open-ended questions and evolves as themes, categories, and theoretical constructs are developed through ongoing analysis [15]. Two experienced qualitative researchers (LH and KM) conducted interviews with women who were diagnosed with breast cancer and were former or current patients, in person at the BDC. Interviews lasted 60 to 90 minutes. During the interviews, participants were asked to map out their breast cancer journey, starting from the time before their diagnosis to the present. For example, they were asked about their activities, the health professionals they interacted with, their feelings, the information they received, and their communication needs at each stage of their journey. Participants were encouraged to take the conversation in new directions and explore concepts not initiated by the interviewers. Our goal was to learn from participants' breast cancer journeys at the BDC to identify gaps in care and their information and support needs, in anticipation of developing a patient-facing app that would address these needs as a complement to the Clinical Navigation Tool aimed at providers.

Phase 2

A semistructured interview guide was developed (for each of the pharmacist and nurse participants) by 3 researchers (IH, JM, and MK; Multimedia Appendix 2). Interviews were conducted in person or virtually (via Zoom [Zoom Video Communications] or Microsoft Teams) in 30-minute time blocks, at the participant's preference. Interview guides followed a semistructured format informed by the RE-AIM framework [18] (ie, to help formulate questions that would elicit responses grounded in implementation theory). To maximize participation, we offered flexible participation options (a single 1-hour session or two 30-minute sessions). Interview questions focused on (1) participants' roles and responsibilities, (2) experience using MyJourney (including duration and frequency of use; how it supports tasks, workflows, and patient management; and whether it meets their needs), (3) MyJourney feature-specific feedback, (4) benefits and challenges of using MyJourney and related recommendations, (5) experience with implementation (what worked well, what did not, any disruptions to workflow, and suggestions for future implementation), and (6) challenges experienced in each of these areas along with associated recommendations, including feedback on how MyJourney could be tailored to support their work better.

Phase 3

The research team developed an online survey for each of the BDC and CC and pilot-tested it for understandability (Multimedia Appendix 3). Surveys were tailored to different audience groups (eg, nurses and clinic staff) and administered via Qualtrics to study participants to collect data on demographics and usability. Usability and user experience were assessed using the System Usability Scale (SUS) [23,24] and the technology acceptance model, which measures ease of use and usefulness of a technology-based innovation [25,26]. Sampling was consistent with established usability guidance that 3‐5 users are typically sufficient to identify most high-severity usability problems in a given design iteration [27]. As such, we aimed to reach at least 3 participants per site who had access to and experience using the platform but were constrained by the number of eligible users during early implementation (2 patients with breast cancer per site). As a result, we treated this as an initial, depth-oriented usability assessment focused on uncovering major usability issues and gathering rich feedback to guide subsequent refinements, rather than on producing generalizable estimates of usability metrics. We also assessed perceptions of workflow alignment, ratings of specific MyJourney features, and overall satisfaction. Demographic variables included age, sex, gender, ethnocultural background, and years in role. In addition, survey responses (stage 1) were used to tailor interview probes for stage 2. For example, if a participant rated a MyJourney feature poorly, the interview explored the reasons and discussed possible solutions to address them.

Data Analysis and Synthesis

Qualitative Data

All interviews were digitally recorded, transcribed verbatim, and reviewed for accuracy. For qualitative analyses across all phases, we used an iterative, inductive, and interpretive process to construct our understanding of MyJourney. Within the interpretive description methodology of Sandelowski [16], qualitative data analysis unfolds as an iterative, inductive, and reflexive process. It begins with immersion in the data to gain contextual understanding, followed by broad initial coding to capture patterns without imposing rigid categories. In phase 1, reviewers (LH and KM) used constant comparison analysis, repeatedly contrasting data segments and cases to develop, refine, and validate emerging insights and themes [16]. This process allowed us to move beyond description, producing a practice-oriented interpretation that informed actionable insights about how MyJourney should be constructed. The codebook was developed iteratively with each subsequent transcript, and its emergent structure became more focused as categories and themes developed. In phase 2, we used directed content analysis, combining deductive coding (guided by RE-AIM and study objectives) and inductive coding to capture themes emerging from participant narratives [28]. Reviewers in this phase (IH, JM, and MK) independently coded initial transcripts and collaboratively developed an initial codebook that evolved as more data were analyzed and was structured around barriers and recommendations. The remaining transcripts were coded in duplicate by applying the codebook. Discrepancies were

XSL•FO
RenderX

resolved by consensus or adjudication with a third reviewer (MK). Thematic categories were refined through team discussions to ensure clarity, distinctiveness, and alignment with the study objectives. Data were then organized by clinical site and mapped onto a figure depicting the workflow, allowing the clinical context in which MyJourney would be used to be identified. Open-ended survey responses were independently coded by 2 reviewers using the same coding structure.

Data Synthesis

Across phases, coded qualitative data were summarized in tables organized by challenges, recommendations, and, where relevant, enabling factors. Representative quotations were selected to illustrate key findings.

Quantitative Data

Survey data were analyzed descriptively using Microsoft Excel. Measures included frequencies, percentages, means, medians, SDs, and IQRs. SUS scores were interpreted using the following benchmarks: >68 (above average), >70 (acceptable), >74 (excellent), and >85 (best imaginable) [29]. Mean scores and SDs were calculated for technology acceptance model dimensions [25]. Due to the small sample size, no inferential statistics were conducted.

Researcher Characteristics and Reflexivity

The evaluation team comprised researchers who identified as female and women from health services and implementation science, who were not involved in the design or delivery of MyJourney. To mitigate potential bias arising from the dual role of a coprincipal investigator (FO) as a platform developer, data collection and analysis were conducted by team members independent of platform design and operations. The interviewer had no prior relationship with the participants. A note-taker attended phase 3 interviews and offered follow-up questions where appropriate. Team members kept reflexive journals and audit trails and met regularly to review assumptions and ground analytic decisions in participant accounts. Member checking was not feasible. Credibility was supported through rich description and direct quotation.

Results

Phase 1: Predesign Engagement With Patients at the Breast Diagnostic Clinic

Participant Characteristics

In total, 13 participants consented and were interviewed between 2018 and 2019 (Table 1). The mean age was 56 years (range 42 - 67 years), all identified as female and a woman, and most identified as White (5/13, 38.5%), South Asian (2/13, 15.4%), and Chinese (2/13, 15.4%), and 1 (7.7%) participant each identified as Jewish, Latin American, West Asian, or White-Aboriginal. The majority of participants (9/13, 69.2%) reported an annual household income of ≥US $659,200. All participants spoke English, and 3 (23.1%) individuals spoke a second language. Of the 12 who specified education, most (8/12, 66.7%) reported at least a bachelor’s degree.

XSL•FO
RenderX

Table . Phase 1 participant characteristics[a].

Characteristic	Values (n=13), n (%)
Age range (n=10)	
41 - 50 years	2 (20)
51 - 60 years	4 (40)
>60 years	4 (40)
Prefer not to say	3 (23.1)
Biological sex	
Female	13 (100)
Gender identity	
Woman	13 (100)
Ethnic or racial background	
Chinese	2 (15.4)
Jewish	1 (7.7)
Latin American	1 (7.7)
South Asian	2 (15.4)
West Asian	1 (7.7)
White	5 (38.5)
White-Aboriginal	1 (7.7)
Education (n=12)	
University—bachelor or more	8 (66.7)
College diploma	3 (25)
University—no degree	1 (8.3)
Prefer not to answer	1 (7.7)
Languages spoken[b]	
English	13 (100)
Arabic	1 (7.7)
Russian	1 (7.7)
Chinese	1 (7.7)
Household income (n=10)	
US $110,000	4 (40)
US $66,000-$88,000	3 (30)
US $88,000-$110,000	2 (20)
US $0-$22,000	1 (10)
US $22,000-$66,000	0 (0)
Prefer not to answer	3 (23.1)
Stage of breast cancer journey (n=13)	
Prediagnosis	2 (15.4)
Diagnosis	3 (23.1)
Treatment	4 (30.8)
Survivorship	4 (30.8)

[a]Percentages may not add to 100% due to rounding.

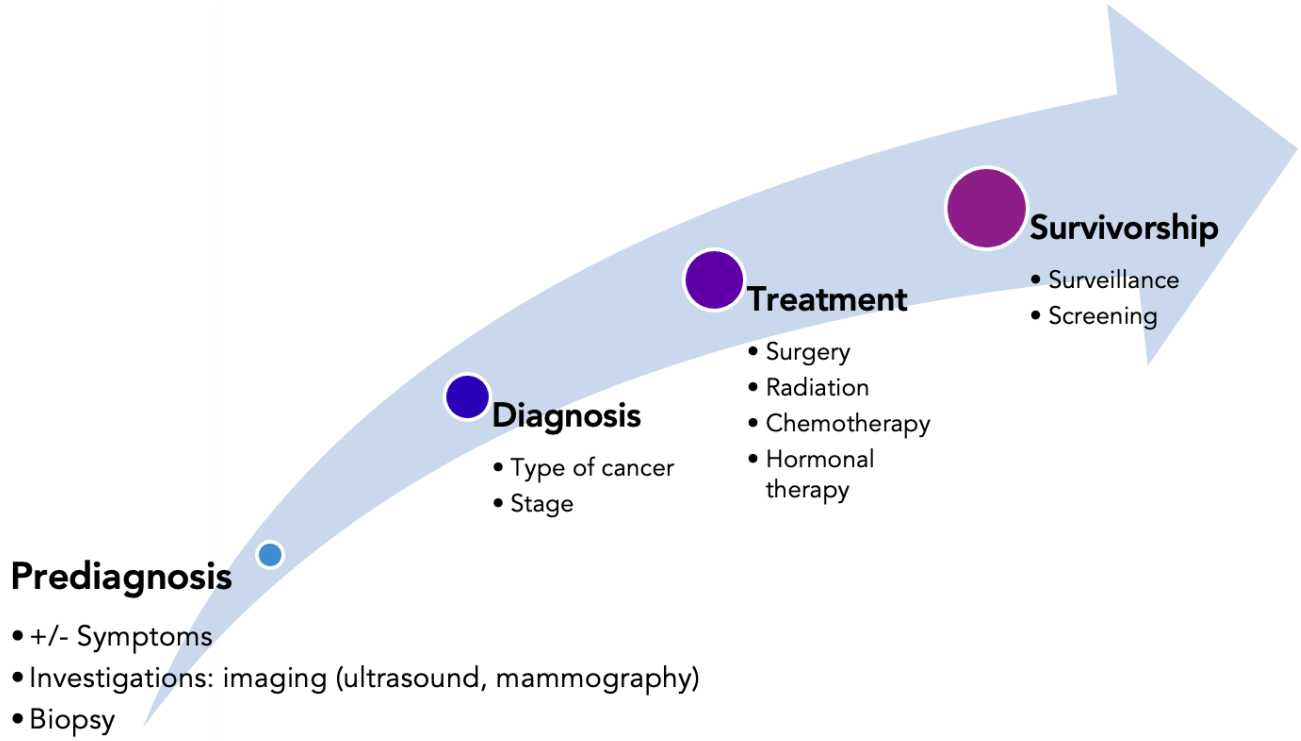
[b]The totals in this section of the table are greater than the sample size, and the percentages add to more than 100%.

XSL•FO
RenderX

Stage of Breast Cancer Journey

Participants described different "phases" of their journey, each with its own associated challenges and communication needs (Figure 2). The phases identified were "prediagnosis," "diagnosis," "treatment" (including active management strategies such as surgery, chemotherapy, and radiation), and "survivorship." These labels mark the key clinical activities during the respective time frame and reflect the clinical perspective that patients with breast cancer are expected to adopt. The majority of participants were at the treatment (4/13, 30.8%) or survivorship (4/13, 30.8%) stage, followed by the diagnosis (3/13, 23.1%) and prediagnosis (2/13, 15.4%) stages. However, participants acknowledged that these phases were not always linear, and not all patients go through all phases. Participants reported that each phase had distinct informational and emotional needs, which were incorporated into the tool's design features. Multimedia Appendix 4 summarizes the themes and relevant quotes describing participants' experiences according to their breast cancer journey stages.

Figure 2. Stages of the typical breast cancer journey as reported by interview participants.



Summary of Qualitative Results

Overview

We developed 8 overarching themes that describe participants' experiences across the breast cancer care trajectory at the BDC and CC at NYGH (Table 2).

XSL•FO
RenderX

Table . Overarching themes that describe participants' experiences across the breast cancer care trajectory at the North York General Hospital Breast Diagnostic Centre (BDC) and Chemotherapy Clinics.

Theme and subtheme	Description	Illustrative quotes
1. Care coordination and workflows		
BDC workflows and processes	Perceptions of the BDC as coordinated, efficient, and well-organized.	• "From the time from mammogram to surgery was only 1 month—that was an amazing experience" [Participant 7]. • "They have a very good process ... After your appointment [they] give you another one ... If you need a referral, they have already sent the referral ... It's a very good process and I don't want it to change" [Participant 1].
Transitions and follow-up	Breakdowns when moving between BDC, chemo, radiation, home care, and primary care, with patients worrying that things will be missed.	• "At the BDC it was a real team approach ... Then, when I went to chemo there was no team approach and it all went downhill from there" [Participant 4]. • "The imaging follow up just doesn't happen. I have to orchestrate it" [Participant 4].
2. Communication and information flow		
Receiving diagnosis	How, where, and by whom the diagnosis is communicated, and its emotional impact.	• "When I got the phone call from my doctor, I was alone at home and told about the diagnosis. That was hard. It would have been better to have someone with me" [Participant 7]. • "I was 'told' that I had breast cancer by the bone scan technician ... before I was given the official diagnosis and that was difficult to hear and very brusque" [Participant 8].
Team communication and contradictions	Inconsistent or conflicting information between clinicians and services, eroding trust.	• "There was lack of communication between groups in this hospital ... I spent the whole week panicking ... and they said 'oh it's fine, we have ice for you'" [Participant 4]. • "They are telling you it's nothing ... They are not doing their research to understand the actual side effects of radiation ... Every time you say it, they dismiss it instead of cataloguing people's experience" [Participant 18].
Role clarity (who to call)	Uncertainty about which clinician is responsible for which problems.	• "There were a few things that I wasn't sure if I should go to the surgeon or the medical oncologist for ... I often wasn't sure who to go to" [Participant 9].
Tone, empathy, and respect	Value placed on clear, respectful, and nondismissive communication.	• "We had an amazing surgeon and oncologist ... They don't try to make you feel bad ... They never show signs of exasperation ... It's very personalized" [Participant 1].
3. Information and resource needs		
Volume and timing of information	Tension between information overload and unmet information needs; importance of timing.	• "They gave me lots of information in the little bag. It was overwhelming ... When I started reading it, it was okay" [Participant 5]. • "The 'welcome to breast cancer' package ... had too much of a lumpectomy focus versus mastectomy" [Participant 4].

XSL•FO
RenderX

Theme and subtheme	Description	Illustrative quotes
Practical “how-to” guidance	Desire for concrete, practical advice about surgery recovery, daily logistics, and self-care.	• “I thought I’d get more info ... what to expect and prepare for after the surgery ... mastectomy bras, etc. I did research online and got a wedgie pillow ... grabbers ... special tables ...” [Participant 2].
Understanding diagnosis and treatment	Need to understand pathology, staging, oncotype, systemic therapies, and their implications.	• “I was trying to understand the HER2. I found this confusing ... It would be good to ... tell us which sites to go to, and not just Google” [Participant 9].
Survivorship “what happens after”	Lack of clear survivorship plans, follow-up schedules, and guidance for life after treatment.	• “After the treatments were over, I was left alone. I want to know, what happens after cancer? ... Now that I have had cancer, it has changed me. How do I deal with this?” [Participant 9].
Safe, reputable online information	Desire for trusted, curated web resources instead of unguided internet searches.	• “I was googling a lot—wish I’d had something reputable to refer to” [Participant 15].
4. Digital tools and patient portal		
Access to records and results	Strong interest in a portal centralizing reports, laboratory results, and imaging.	• “I would like a patient portal with all of my reports and bloodwork results ... upcoming appointments, information on drugs I’m taking, and resources, like Wellspring” [Participant 6]. • “She loves Sunnybrook mychart ... she can bring the results to other medical appointments, so they have the information” [Participant 4].
Appointments and chronology	Desire for a clear visual timeline of appointments and treatments.	• “Would like to see the chronology of her appointments and treatments in a” [Participant 7].
Laboratory result “lookup”	Need for plain-language explanations of test results.	• “What about a lookup feature for what the lab results all mean—this would be great, so you don’t have to go Google” [Participant 8].
Messaging and chat	Preference for asynchronous written communication for nonurgent questions.	• “A good option would be to have a chat function ... I don’t necessarily need an answer right away and I don’t want to talk on the phone” [Participant 3].
Technology comfort and training	Wide variation in digital literacy and format preference (portal vs paper).	• “I would be very interested in using the portal, but someone would have to show me how to use it because I’m not good at using computers” [Participant 5]. • “I am a paper person. Would like to have all the educational materials paper-based” [Participant 8].
5. Experience of quality of care		

XSL•FO
RenderX

Theme and subtheme	Description	Illustrative quotes
BDC environment and team	Perceptions of the clinic environment, staff attitudes, and teamwork.	• "The whole team—everybody. Every time I am here I feel that I am coming home" [Participant 1].
Chemo and radiation experiences	Physical and emotional experiences of systemic and radiation treatments.	• "Chemo is the worst thing ever ... The chemotherapy treatments involved long waits" [Participant 6]. • "The nurses there were amazing at the chemo clinic ... They really tried to make a really bad situation bearable" [Participant 4].
Home care or CCAC[a]	Challenges with community nursing, reliability, and competence.	• "CCAC came and took care of my wound. It was a comedy of errors ... They had no idea whether they were coming or going, and whether they had supplies" [Participant 4].
6. Psychosocial and emotional support		
Distress and uncertainty	Intense distress at diagnosis and during periods of waiting or uncertainty.	• "The hardest part was from when I thought I probably had breast cancer, but didn't know for sure ... The hardest part was not knowing whether I would be able to see my son go to school" [Participant 16].
Ongoing anxiety and fear of recurrence	Persistent psychological impact after treatment ends.	• "I still struggle with the psychological aftercare ... You are always on guard for another lump" [Participant 9].
Support services and gaps	Value of programs like Wellspring and perceived gaps in psychosocial support.	• "The hospital just focuses on the body. They don't focus on the mental and the cognitive. I was happy to find Wellspring ... I go to a lot of classes" [Participant 6].
Humor and coping styles	Use of humor and individualized coping strategies.	• "Am I stressed, absolutely. To deal with it, I sit in the car and do stand-up humour for myself" [Participant 5].
7. Patient agency and self-management		
Self-advocacy and orchestrating care	Patients pushing for referrals, chasing results, and coordinating their own care.	• "I have to orchestrate it ... If you aren't willing to drive the car, let me. I want my test results. I want them to be accessible" [Participant 4].
Personal tracking and journaling	Using notebooks, binders, and journals to track care and questions.	• "After each appointment, I wrote in my book what happened. So, I have this record of what care I received" [Participant 12].
Doing own research and lifestyle change	Independent research on treatments and lifestyle changes to support recovery.	• "I wanted to be part of my recovery, so I did some digging and reading ... I removed sugar from my diet ... removed meat ... took 2 minute cold showers ... and drank a lot of water before chemo" [Participant 3].
8. Survivorship care and transitions		

XSL•FO
RenderX

Theme and subtheme	Description	Illustrative quotes
Who leads survivorship	Confusion and preferences around oncologist- versus GP[b]-led follow-up.	• "What happens now? Who follows me? Who do I speak with? ... There is no coordinated care. No one is in charge. It's you" [Participant 18]. • "I would prefer having my GP follow me once my treatment is complete ... She phones me and tells me my results" [Participant 11].
Requirements for GP-led care	Perceived capabilities and training needed in primary care.	• "The family doctor would have to have a lot of training regarding ... post-treatment side effects ... and the softer skills side of it" [Participant 7].
Long-term anxiety and follow-up	Ongoing fear of recurrence and desire for continued monitoring as years pass.	• "As the years go by, you think, nothing has happened. And you start to worry ... That's when the anxiety starts to come back with the waiting" [Participant 13].
Survivorship identity and peer support	Seeking peer support that matches age, life stage, and experience.	• "I would be really interested in a survivorship focus group ... After the treatments were over, I was left alone" [Participant 9]. • "I still haven't found someone like me ... a mother of two young kids ... To have that kind of peer support is really powerful" [Participant 16].

[a]CCAC: community care access center.

[b]GP: general practitioner.

Care Coordination and Workflows

Participants described the BDC as a coordinated, efficient entry point into care, characterized by rapid movement from abnormal imaging to biopsy and surgery and by clear internal communication. Women emphasized short wait times and proactive scheduling of subsequent appointments and referrals, which reduced uncertainty and signaled that "the system" was functioning well. In contrast, this perceived coordination often deteriorated once patients transitioned into chemotherapy, radiation, community nursing, or primary care. Several participants reported that the "team approach" appeared to end after surgery, with no single clinician or service clearly overseeing the remainder of care. Delays in imaging follow-up, confusion around orders, and inconsistent tracking of test results were frequently described. As a result, many participants felt responsible for navigating and organizing their own care, including obtaining their results, confirming appointments, and ensuring referrals had been made, a role they found exhausting in the context of illness and emotional distress.

Communication and Information Flow

How and where the diagnosis was communicated strongly shaped participants' early experiences. Some received the diagnosis alone at home by telephone and described this as particularly distressing; others were informed indirectly by nonphysician staff (eg, imaging technicians) before any formal consultation, which they found abrupt and insensitive. Beyond diagnosis, participants highlighted inconsistent messages between clinicians and services and poor intrateam communication, including contradictory advice between oncology and treatment units, eroding trust and increasing anxiety. Many were uncertain whether to contact their surgeon, medical oncologist, radiation oncologist, family physician, or nurse navigator when new questions or symptoms arose. At the same time, participants expressed high appreciation for clinicians who communicated in clear, plain language, invited questions, and repeated information without frustration. Participants also reported that initial information packages were sometimes overwhelming and poorly tailored (eg, lumpectomy-focused materials given to women undergoing mastectomy) and found gaps in practical "how-to" guidance (eg, what to buy and prepare before surgery, managing drains, and concrete strategies for side effects) and in survivorship-oriented information about "what happens after" treatment, follow-up schedules, and recurrence risk.

Information and Resource Needs

Information needs were complex and shifted across the cancer trajectory. Many participants recalled receiving large information packages at diagnosis that felt overwhelming or poorly tailored (eg, materials focused on lumpectomy when they were having mastectomy), even as they still lacked answers to very practical questions about what to expect after surgery and how to prepare their home. They sought clear, accessible explanations of pathology results, risk scores, and treatments, as well as guidance on what is "normal" versus concerning during and after treatment. Critically, once active treatment ended, people described a stark information void about survivorship—wanting to know what happens next, how long follow-up would last, what symptoms to watch for, and how to live with the ongoing impact of cancer. Many expressed a desire

XSL•FO
RenderX

for curated, trustworthy online resources to reduce reliance on unsupervised internet searches.

Digital Tools and Patient Portal

Participants suggested that a cancer-focused portal to address multiple unmet needs in one place would be helpful. Those with experience of existing portals valued having access to laboratory results, imaging reports, and appointment lists, and being able to bring printed results to other appointments. They imagined a more tailored tool that would provide a clear chronology of appointments and treatments, explanations of laboratory values in plain language, and secure messaging or chat for nonurgent questions. However, our data also highlighted wide variation in participants' comfort level with digital solutions: some described themselves as "information junkies" who would "be on it all the time," while others said they would need to be walked through how to use such a system or preferred paper-based materials. Any digital solution would therefore need to support both high- and low-technology users, offer training, and coexist with paper formats rather than fully replace them. Multimedia Appendix 5 summarizes the features participants reported they would be interested in seeing in a future patient portal.

Experience of Quality of Care

The overall quality of care was experienced as uneven. Many participants spoke positively about the BDC environment, highlighting friendly staff, clean, welcoming spaces, and a sense of being "at home" and well-supported. Chemotherapy units were described as both physically and emotionally challenging, yet often made more bearable by attentive, kind nurses. In contrast, community-based home care services were frequently characterized as disorganized, inconsistent, or poorly coordinated, with missed visits, questionable technique, or lack of supplies. Long waits for some specialist appointments, crowded spaces, and privacy concerns in certain treatment areas also shaped perceptions of quality. These contrasting experiences underscored how much the attitudes, organization, and resources of specific teams influence patients' sense of safety and trust.

Psychosocial and Emotional Support

Across the cancer journey, participants reported substantial psychosocial burden. The period between suspecting cancer and receiving a definitive diagnosis was described as especially distressing, particularly for those with young children who feared not living to see key milestones. They described intense fear and uncertainty at the time of diagnosis and during periods of waiting for test results or treatment decisions. Even after treatment, many reported ongoing psychological impacts, including heightened worry before follow-up imaging and a sense of being "on guard" for recurrence. Some participants found that strong support through community-based survivorship centers (such as psychosocial programs and survivorship classes) was essential to addressing emotional and cognitive impacts that hospital-based services, which "just focus on the body," often neglect and do not adequately address mental health, cognitive changes, or other concerns. Participants emphasized the value of humor and individualized coping strategies, and having space for levity, storytelling, and emotional validation alongside more formal counseling or group support. Transitions into survivorship were characterized by anxiety and uncertainty about who was responsible for ongoing care. Some participants were comfortable transitioning to a trusted family physician, while others questioned whether primary care had sufficient oncology expertise and time to manage late effects and psychological needs. Many described feeling "left alone" once active treatment concluded.

Patient Agency and Self-Management

Across interviews, patient participants portrayed themselves as active agents who often had to fill gaps in the system. Many described pushing for referrals, insisting on further tests, or repeatedly calling to obtain results, reflecting both a strong desire for control and a lack of trust that the system would automatically "catch" everything. Self-management practices were common. For example, participants kept notebooks or binders to track appointments, results, and questions, and some engaged in extensive independent research on treatments, nutrition, exercise, and complementary strategies to support recovery. These behaviors were framed both as empowering and burdensome, as patients felt they "shouldn't have to" organize their own care to this degree; yet, they would have valued tools and information that would have allowed them to do these activities more effectively.

Survivorship Care and Transitions

The transition from active treatment to survivorship was a major source of uncertainty and distress. Participants frequently reported not knowing who was now "in charge" of their care, how often they would be seen, or what tests would be done in follow-up. Some were comfortable with their family physician assuming a lead role, particularly when they had longstanding, trusting relationships and good access. However, others doubted that primary care providers had sufficient oncology-specific knowledge or time to manage their physical and psychological needs without additional training. Survivorship was also described as a distinct identity shift, whereby they no longer felt like active patients but did not feel "back to normal" either. Many sought survivorship programs, focus groups, or peer connections that recognized and supported this ongoing phase, rather than treating the end of treatment as the end of the cancer experience.

How Findings Informed the Design of the Patient-Facing App of MyJourney

Overall, participants emphasized the importance of structured follow-up plans, clear communication about recurrence risk, and accessible survivorship resources. Table 3 highlights the features identified by participants as to be considered for the patient-facing app. Multimedia Appendix 6 presents the minimum viable feature set used to build the patient-facing app.

XSL•FO
RenderX

Table . Features that were considered in the development of the patient-facing app.

App feature area	Directly addresses participant needs or quotes
Timeline and roadmap	Need for chronology, pathway or flowchart, and "knowing what happens next" including aftercare and survivorship.
Results and lookup	Desire for laboratory and test access, explanations (oncotype, seed lock, and radiation effects), and faster and clearer results.
Messaging and routing	Confusion about who to contact, wish for chat, frustration with missed callbacks, and vacation gaps.
Education library	Information overload plus gaps; wish for curated sites, practical tips, emotional support, and survivorship information.
Psychosocial or peer support	Requests for psychologists, survivorship groups, younger-women groups, peer stories, and contact.
Self-management tools	Patients documenting care, tracking steps, changing diet, and orchestrating follow-up themselves.

Phase 2: Preimplementation Interviews at the CC

Participant Characteristics

In total, 3 nurses and 5 pharmacists were interviewed between November 23 and 29, 2023, representing all relevant users of the MyJourney platform at the CC. Most participants (7/8, 87.5%) identified as female and a woman, aged 41 - 50 years (5/8, 62.5%), primarily with a Chinese (2/8, 25%), East African (1/8, 12.5%), South Asian (3/8, 37.5%), or White (2/8, 25%) racial background. Most participants had been in their current role for under 5 years (5/8, 62.5%; Table 4). At the CC, nurses were primarily responsible for coordinating care, managing symptoms, and answering patient inquiries. Pharmacists provided patient education, addressed treatment-related issues, secured funding, and offered tailored resources. Two pharmacist roles existed: Clinical One, who conducted scheduled education sessions before the first chemotherapy cycle, and Clinical Two, who provided chairside counseling to patients in clinical trials or to those who had not received prior education on the day of their first treatment.

XSL•FO
RenderX

Table . Phase 2 participant characteristics[a].

Characteristic	Values (n=8), n (%)
Age range (years)	
30 - 40	1 (12.5)
41 - 50	5 (62.5)
51 - 60	2 (25)
>60	0 (0)
Biological sex	
Female	7 (87.5)
Male	1 (12.5)
Other	0 (0)
Gender identity	
Woman	7 (87.5)
Man	1 (12.5)
Other	0 (0)
Ethnic or racial background[b]	
White	2 (25)
Chinese	2 (25)
Black	0 (0)
Chinese	2 (25)
East African	1 (12.5)
Latin American	0 (0)
South Asian	3 (37.5)
White	2 (25)
East African	1 (12.5)
Prefer not to say	2 (25)
Years in current role	
<3	1 (12.5)
3 - 5	4 (50)
6 - 10	1 (12.5)
>10	2 (25)

[a]Percentages may not add to 100% due to rounding.

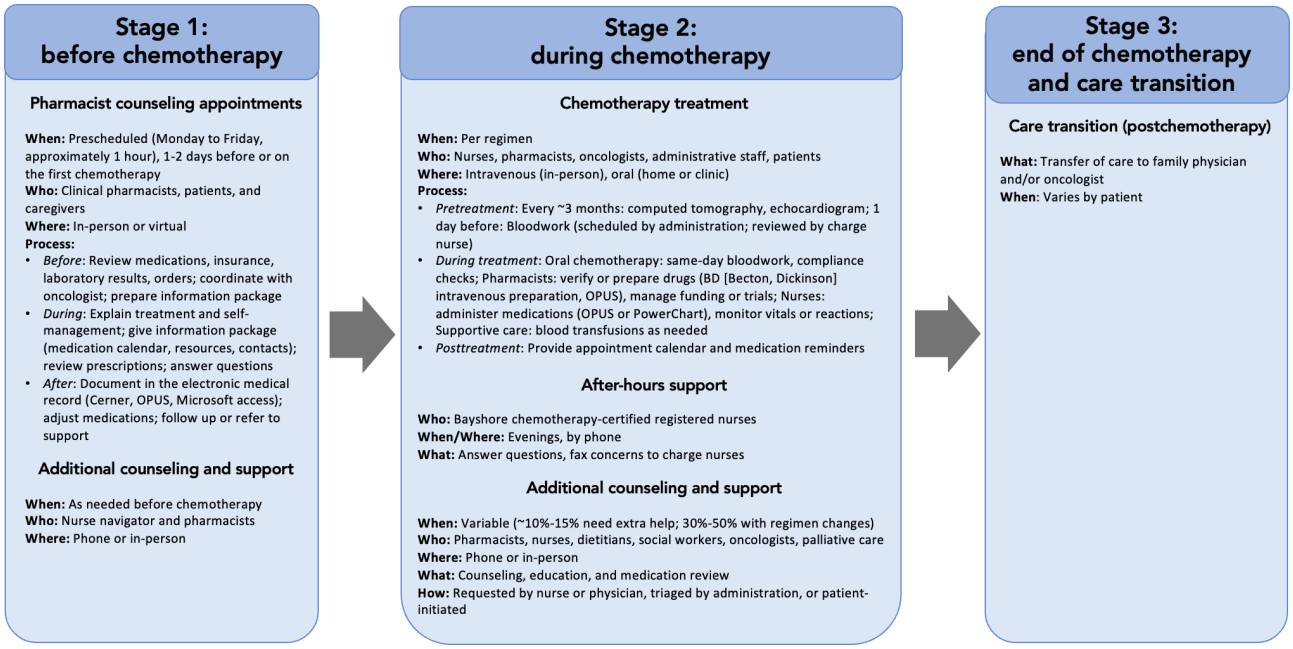
[b]The totals in this section of the table are greater than the sample size and the percentages add to greater than 100%.

Workflow Mapping: Phases of Chemotherapy Care

Workflows described by participants spanned three stages of the patient journey during their chemotherapy treatment: (1) before chemotherapy begins (including pretreatment education and preparation), (2) during chemotherapy (including treatment delivery, monitoring, and follow-up), and (3) end-of-treatment and transition to the next phase of care (Figure 3).

XSL•FO
RenderX

Figure 3. Workflow stages during chemotherapy treatment. OPUS: Ontario Public Health Standards.



Challenges and Recommendations

Preimplementation interviews with pharmacists and nurses uncovered multilevel challenges to care coordination across the chemotherapy care continuum, which we organized into five domains: (1) patient-level, (2) administrative and workflow inefficiencies, (3) health information technology systems limitations (eg, poor electronic health record usability and lack of interoperability between Ontario Public Health Standards, Cerner, and Connecting Ontario), (4) health resource constraints (eg, nurse staffing shortages and inadequate physical space), and (5) quality of care gaps (eg, lack of pharmacist continuity and no secure patient communication channels). Table 5 summarizes these challenges and MyJourney design recommendations to address them across the 3 stages of chemotherapy care: stage 1 (before), stage 2 (during), and stage 3 (end).

XSL•FO
RenderX

Table . Implementation challenges across the chemotherapy care continuum and MyJourney design recommendations to address them.

Challenge domain	Stage 1: before chemotherapy (Multimedia Appendix 7)	Stage 2: during chemotherapy (Multimedia Appendix 8)	Stage 3: end of chemotherapy and transition (Multimedia Appendix 9)	MyJourney design recommendations to address challenges
1. Patient-level	Lack of access to appointment information; privacy concerns about digital resources; limited digital literacy and technology access.	Information overload when large volumes of information were delivered on treatment days; walk-in patients seeking urgent support; limited after-hours access to clinicians.	Regional differences in community care coordination contributing to incomplete medication information and gaps at transition.	Patient-facing app centralizing appointment details and visit summaries; layered, on-demand education to reduce overload; mobile-friendly, plain-language content to support varying digital literacy; future secure messaging to support questions between visits and during transitions.
2. Administrative and workflow	Incomplete or unsigned prescriptions; late counseling notifications; delayed funding reviews; manual preparation of regimen-specific materials; heavy follow-up demands.	Reliance on verbal updates or email for after-hours issues; incomplete information leading to duplicated counseling; nurses unaware of appointment changes; high call volumes; manual record checks and after-hours reports.	Regional community care access differences hindering coordination and resulting in incomplete medication lists at transition.	Clinical Navigation Tool supporting task lists, standardized documentation templates, and structured prompts; centralized record of counseling and follow-up to reduce duplication; integration of key appointment and treatment information to support proactive planning and smoother transitions.
3. Health IT systems	Time-intensive Cerner documentation; restricted access and incomplete patient information; limited interoperability between OPUS[a], Connecting Ontario, and PowerChart; pharmacists unable to view oncologist appointments.	Manual order entry and frequent program switching; lost interprofessional notes on admission; unclear note headings; ongoing limited interoperability between OPUS and Cerner.	Fragmented information across hospital and community systems contributing to incomplete medication histories.	MyJourney designed as a single navigation workspace that aggregates key information from multiple systems where possible, reduces duplicate data entry through structured fields, and surfaces oncology appointments and treatment plans in a consolidated view to support navigation across settings.
4. Health resources	Complex funding confirmation processes; strained patient-to-nurse ratios; staffing constraints; shared counseling space with a vaccine clinic.	Nurse staffing shortages; physical space limitations restricting in-clinic transfusions; high-risk treatments requiring extended monitoring.	Ongoing resource constraints in community and hospital settings affecting follow-up capacity.	Workflow tools to streamline preparation for counseling, reduce repetitive administrative tasks, and support team-based navigation so workload can be shared across pharmacists and nurses, with clearer task allocation over time.
5. Quality of care	Medication charts lacking visuals or digital access; strict time limits that truncate counseling; privacy concerns limiting use of educational videos; inefficient coordination of bloodwork and education appointments.	Delayed drug preparation prolonging chair time; no 24/7 clinician availability; no dedicated pharmacist continuity; no secure electronic communication with patients.	Coordination gaps at transition leading to fragmented follow-up and incomplete information for community providers.	Visual, phase-specific patient education content; structured symptom and concern tracking to focus limited visit time; clear assignment of navigation tasks within the care team to enhance continuity and follow-up; digital information sharing to support more consistent transitions.

[a]OPUS: Ontario Public Health Standards.

Briefly, across stages, staff described patients' difficulty accessing technology, tracking appointments, and experiencing information overload when large volumes of education were delivered on treatment days and had limited options for secure communication and follow-up between visits. They also emphasized administrative inefficiencies (eg, incomplete

XSL•FO
RenderX

prescriptions, delayed notifications, and duplicated counseling); manual, fragmented, and time-intensive documentation across Cerner, Ontario Public Health Standards, and Connecting Ontario; staffing and space constraints; and coordination gaps at transitions to community care. These findings contributed to the design of MyJourney's patient-facing app (centralized appointment and treatment information, phase-specific educational content, and tools to support symptom tracking) and the Clinical Navigation Tool (structured prompts, standardized documentation templates, and a shared workspace to reduce duplication and improve visibility across the team), with the goal of addressing the specific workflow, communication, and information-access barriers identified by frontline clinicians and staff. See the complete set of challenges, recommendations, and related participant quotes for each stage in Multimedia Appendices 7-9.

Phase 3: Evaluation of MyJourney's Implementation and Usability

Participant Characteristics

In total, 2 charge nurses (CC), 1 charge nurse, and 1 administrative staff member (BDC clinic) participated in the usability evaluation. This sample comprised female participants aged 30 - 60 years who identified as women, with Black, East African, South Asian, and White racial backgrounds. Participants had been in their professional roles for 5 - 11 years (Multimedia Appendix 10). Phase 3 findings primarily reflect the RE-AIM domains of implementation and maintenance. Implementation-related results (stages 1 and 2) describe how patients navigated the app, the extent to which MyJourney fit within their routines, and specific usability issues that affected completion of key tasks, whereas maintenance-related results (stage 3) focus on participants' intentions to continue using MyJourney and their views on its ongoing role in their cancer journey

Stage 1 Evaluation (Usability Survey Results)

BDC Clinic

Usability was rated as excellent (mean SUS score 81.3, SD 8.8). Participants agreed that MyJourney was useful (mean score 3.9 of 5) and easy to use (mean score 4.0) and were very satisfied with the platform overall (mean score 4.5). Staff described MyJourney as helpful for workflow and daily tasks, with appointment visibility and patient organization consistently rated as easy to use, well-organized, functional, and satisfying. Notifications and diagnostic information were also valued, although perceptions of helpfulness varied slightly by role: "For me, it's about having fewer places to look ... it helps me do my job better and save time" (Participant 09, BDC).

Chemotherapy Clinic

Usability was rated as excellent with a mean SUS score of 76.3 (SD 8.8)6. Participants consistently found the platform useful (mean score 3.9 of 5), very easy to use (mean score 4.3), and very satisfied overall (mean score 4.5). Features were considered understandable, well-organized, functional, and helpful, with satisfaction ranging from agree for notifications and task reminders to strongly agree for treatment summaries, appointment views, and notetaking: "It's very easy to use and organized ... instead of clicking or going to a million places in Cerner, it's right there" (Participant 01, CC).

Stage 2 Evaluation (Interview Results)

BDC Clinic

Interviews supported survey results. Participants described MyJourney as efficient, with clear benefits for referrals, patient tracking, and information management. Shared visibility across the care team was considered a strength, and staff emphasized the value of appointment visibility, triage functions, and diagnostic notifications to support workflow. Participants also identified challenges, including glitches such as missing or outdated appointments, inconsistent notifications, and the absence of document upload capability. Recommendations included adding accountability features, digitizing paper charts, allowing customizable lists, and eventually integrating patients into multidisciplinary cancer conferences. Despite these challenges, staff expressed strong support for the platform. They described MyJourney as a communication tool that saves time, prevents patients from being missed, and offers a clear improvement over older workflows. Long-term use was considered both feasible and beneficial, with participants anticipating continued improvements and the forthcoming patient app (see Multimedia Appendix 4 for representative quotes).

Chemotherapy Clinic

Interviews reinforced survey findings, with participants consistently describing MyJourney as easy to use, clearly organized, and helpful for daily workflow. Staff highlighted that it took only minutes to learn to navigate the system, which replaced paper-based tracking and consolidated treatment information into a single platform. Features such as notes, tasks, treatment summaries, and appointment views were identified as particularly valuable. At the same time, participants reported challenges with reminders and notifications that were only visible during active sessions. They recommended adding email notifications and requested further integration of pathology, bloodwork, and imaging results. Limited customization and the timing of installation during work hours were also reported as concerns. Participants recommended greater flexibility in tailoring views, broader integration with existing systems, and extending MyJourney beyond breast cancer to other types of cancer.

Stage 3 Evaluation (Repeat Usability Survey at the CC)

At the 6-week follow-up, perceived usability increased from a mean SUS score of 76.3 (SD 8.8) to 86.3 (SD 1.8). This score is interpreted as "best imaginable usability" [24]. Participants found the platform useful (mean score 4.0 of 5) and very easy to use (mean score 4.8 of 5), and they were overall very satisfied (mean score 4.5). Overall, participants expressed strong support for MyJourney, reporting that it streamlined workflows by addressing challenges identified in phase 2, showed perceived improvement in communication within the care team, and was feasible for long-term use. Perceived benefits were described as: "... everything is right there and ... you're able to see a

XSL•FO
RenderX

snapshot" (Participant 01, CC); "Before I used to have to check Cerner and SharePoint ... Now a lot of times I can just look in one place in MyJourney" (Participant 09, BDC). Based on our collective findings, five MyJourney features were developed to meet the needs of staff at the BDC: (1) diagnostic imaging and pathology notifications, (2) upcoming appointment information, (3) triaging patients accordingly, (4) diagnosis information for patients with cancer, and (5) organizing BDC patients according to their appointment types. At the CC, six features of MyJourney were identified to address the identified challenges: (1) treatment summary; (2) upcoming or past appointments view; (3) customized notes, tasks, and reminders; (4) task reminders at the patient level; (5) educational content via mobile app; and (6) overview of bloodwork appointments or results. Multimedia Appendix 5 summarizes recommendations to address challenges for specific MyJourney features.

Technical Integration of MyJourney

Collectively, findings from each phase were translated into phase-specific prompts, content scheduling, and task workflows for both the Clinical Navigation Tool and the patient app. Regarding technical integration, the only data element entered directly by patients was the provincial health card number. All other core patient demographics (name, gender, date of birth, and age) were retrieved automatically from the NYGH electronic medical record system (Cerner) via system integration. For each patient, we linked the hospital medical record number and health card number to an existing referral in our system. If no referral record was present, we created one and manually entered the patient's demographic information. In practice, referrals were available for the vast majority of patients, as clinic policy requires a referral from a family physician prior to consultation. Cancer diagnoses were recorded manually within the innovation platform using a structured drop-down field. Following completion of the diagnostic work-up, the confirmed diagnosis was entered to transition the patient into the "cancer diagnosis" pathway, which, in turn, enabled automated capture of subsequent hospital visits and oncology appointments. For patients with stage IV disease, clinicians manually advanced the record to the stage IV journey by selecting a dedicated "stage 4" option after the diagnosis was entered; this step was based on clinical staging, including cross-sectional imaging demonstrating distant metastases. All remaining data were captured automatically through integration with the electronic medical record. The system continuously identified relevant laboratory results and scheduled appointments, moved patients into the "cancer journey" based on appointment data, updated their status across treatment states, and generated treatment summaries from encounters with chemotherapy, surgery, and radiation without additional manual data entry.

Discussion

Principal Findings

Findings across the 3 phases of this research demonstrate how patient and provider input shaped MyJourney, how it functioned in practice to support the breast cancer journey for providers and patients, and its potential to reduce administrative burden while improving care coordination. Our multiphase studies demonstrated that MyJourney was feasible to implement in 2 oncology settings at an academic community hospital in Ontario, Canada, and that end users found it highly usable and helpful. The phased, iterative approach to our methods enabled insights from patients and providers to inform the design, local adaptation, and early evaluation of MyJourney, which likely contributed to its rapid acceptance.

Phase 1 provided a patient-centered map of the breast cancer journey, along with key features to be incorporated into the design of the MyJourney platform components. This phase provided foundational input for both design and implementation. Patients described the BDC as a coordinated entry point into breast cancer care but highlighted substantial gaps when transitioning into chemotherapy and survivorship, including information overload, unclear points of contact, and a need to self-manage coordination across services. For example, these insights identified the timing and content of the navigation touchpoints for the Clinical Navigation Tool and the patient app. Although this paper primarily focused on the provider-facing tool (ie, clinical navigation), the phase 1 findings are foundational, demonstrating how patient priorities can be systematically translated into platform design.

Phase 2 revealed workflow pain points in information access, communication, task tracking, and system interoperability at the CC. Nurses and pharmacists described heavy reliance on manual processes, incomplete or inconsistent appointment information, and fragmented systems. These findings positioned MyJourney as a practical solution by consolidating information, reducing manual entry, and providing notes, tasks, and reminders that supported cognitive load. The preimplementation assessment generated site-specific requirements that informed the configuration prior to rollout. These insights were used to define navigation touchpoints, content, and features in MyJourney, particularly around transitions from diagnosis to treatment and beyond, and were subsequently leveraged, alongside CC workflow data from phase 2, to guide the localization of MyJourney for the CC. By systematically mapping observed challenges to specific MyJourney design recommendations (Table 5), we demonstrated how frontline implementation barriers informed the platform's core functions, rather than designing in isolation from workflow realities. This alignment between identified needs and design decisions is likely to have contributed to early acceptability and perceived implementation fit.

Phase 3 findings indicated strong acceptance among tool users, with improvements in usability over time. Early acceptance after only 1.5 weeks likely reflects the value of preimplementation work and the platform's maturation since its initial BDC deployment in 2022. Staff described mechanisms through which the platform delivers benefits: replacing paper-based notes and task tracking, centralizing essential patient information, reducing duplicate system checks, and externalizing memory work through notifications and reminders. Treatment summaries and appointment views created a shared mental model across team members, which supported coordination and handoffs. Together, the layout, content, and features of MyJourney coalesced to save time and improve efficiency for daily workflows and tasks. RE-AIM also helped

XSL•FO
RenderX

us to interpret phase 3 findings, revealing that, despite the small usability sample, MyJourney demonstrated promising implementation, with participants able to complete core tasks and articulating clear benefits for organizing information and tracking care. At the same time, their reflections on when and how they would continue to use the app, and on conditions needed for ongoing integration into clinical workflows, directly informed our understanding of maintenance and our plans for scale-up and refinement.

Opportunities for improvement were clear and actionable. Reliability needed to be strengthened by resolving errors to maintain user confidence, particularly during busy periods. For example, appointment views need to display entries consistently. Integrations should be expanded to include bloodwork, diagnostic imaging, and pathology, with straightforward navigation to source systems where embedding is not yet possible. Additional patient details, the ability to upload key letters or documents, and role-based list filtering would better align the tool with daily work. Out-of-app notifications, such as email triggers for reminders, could support time-sensitive follow-up when users are not actively in the system.

Implications for Practice and Policy

The results confirm and extend existing evidence regarding the benefits and challenges of digital health platforms in oncology care. Oncology nurse navigation programs are well established and have been shown to reduce barriers, improve timely care, and enhance patient experience, particularly for patients with complex cancer trajectories [30-32]. However, navigation activities are often supported by fragmented tools (eg, electronic health record notes, spreadsheets, and paper trackers) with limited interoperability or standardized digital support for scheduling, task management, and program-level data use [31,33]. Our findings suggest that MyJourney extends these traditional models by embedding navigation functions within an integrated, provider-facing platform that consolidates appointments, treatment summaries, key laboratory information, and navigation tasks for charge nurses, pharmacists, and administrative staff in the CC. In contrast to models that rely on a single nurse navigator to manually orchestrate care, MyJourney distributes navigation activities across roles through shared dashboards and reminders, while automatically drawing core patient and appointment data from the electronic medical record to reduce duplicate documentation. By localizing the Clinical Navigation Tool to the CC, this study demonstrates how navigation principles can be operationalized within the day-to-day delivery of systemic therapy, complementing but not replacing existing nurse navigation. It also addresses calls for interoperable, scalable digital infrastructure to support navigation across the cancer continuum, which is largely lacking [30,33].

The use of integrated digital platforms aligns with global health system priorities, such as the Quintuple Aim, by simultaneously addressing patient and provider experience, enhancing workflow efficiency, and supporting continuity of care [14]. Evidence has shown the fragmented nature of cancer care and the high burden placed on clinicians by noninteroperable health information systems [34-37]. MyJourney's design addresses these documented challenges by consolidating provider-facing functionality and facilitating the timely exchange of information. Platforms that coordinate tasks across teams represent a promising strategy for enhancing multidisciplinary collaboration and reducing provider workload, findings consistent with recent studies on digital navigation tools and care pathway supports [34-37]. The iterative, user-centered development process undertaken in this study further highlights the importance of interest-holder engagement in implementation research to ensure the practical relevance, adaptation, and uptake of new health technologies.

Provider-facing digital tools can reduce administrative burden when configured to local workflows and introduced through a preimplementation assessment [38-40]. Health systems should invest in interoperability and lightweight notification pathways, allowing navigators and nurses to act on results without repeatedly checking the system. Workforce implications are also notable: features such as task reminders, appointment views, and shared dashboards directly address the cognitive and administrative load of nurses and pharmacists, suggesting that digital platforms can be part of a broader strategy to sustain oncology workforce capacity [37-43]. The adaptability of MyJourney across both mature (BDC) and new (CC) implementations suggests that it has scalability potential across cancer programs and possibly to other diseases requiring similar care-delivery support. While this study focused on the provider-facing tool, the linked patient app may also reduce call volumes and enhance self-management, creating system-wide benefits.

Strengths and Limitations

Our study had several strengths. First, our evaluation used a multimethod, iterative design that incorporated qualitative interviews and quantitative surveys at multiple stages, including the usability evaluation. Second, we engaged diverse knowledge user groups within MyJourney, including patients, nurses, pharmacists, and administrative staff, to gain a comprehensive understanding of the platform's potential impact and areas for improvement. Third, the use of the interpretive description methodology provided detailed insights, closely grounded in participants' experiences and clinical contexts. The integration of patient insights and provider workflow mapping into the design and rollout of the MyJourney platform at the CC helped us refine it to better align with real-world needs. Finally, interpreting our findings through RE-AIM (phase 3) offers preliminary evidence of promising implementation and potential maintenance of MyJourney, with participants able to complete core tasks and describing the app as helpful for organizing information and coordinating aspects of their care. Patients and clinicians reported perceived reductions in documentation burden and smoother workflow coordination when using MyJourney.

Limitations include a small sample confined to 2 clinics within 1 hospital, so we acknowledge that the generalizability of our findings to smaller hospitals, rural settings, or low-resource environments may be limited. However, this approach allowed us to observe both a mature implementation (ie, the BDC clinic) and an early-stage adoption (CC) of MyJourney within the same

XSL•FO
RenderX

institution. Nevertheless, we acknowledge that the implementation of innovations such as MyJourney is dependent on institutional readiness, leadership buy-in, and interoperability capacity. There is also a possibility that our purposive sampling strategy to identify eligible patients, providers, and staff introduced selection bias, as our recruitment methods incorporated convenience and self-selection (eg, posters, prior consent lists, and voluntary participation by staff). This may limit the transferability of our findings. However, we included nearly all MyJourney users in our sample. Phase 2 interviews did not include oncologists or surgeons, reflecting their limited routine use of the CC Clinical Navigation Tool at this stage; future work should more directly incorporate physician perspectives, as the platform scales and physician-facing components are further integrated. Another related limitation was that the patient participants in phase 1 were predominantly English-speaking, highly educated, and reported higher income, which does not reflect the North York community. In future evaluations, we will ensure the recruitment of patients with diverse intersecting identities, including gender identity, race, language, and accessibility needs. Outcomes are primarily perceptual and descriptive at this stage of technology evaluation (particularly for describing workflow efficiency), and integration with external systems has not been tested. We did not collect objective use or long-term sustainability metrics (eg, log-in frequency and time-on-task) for the BDC implementation; therefore, our assessment of implementation "success" is based on user-reported usability, perceived utility, and integration into workflows. As such, our results should be interpreted as evidence of early implementation success rather than definitive proof of sustained adoption (ie, they should be understood as descriptive and perceptual indicators of implementation success and anticipated sustainability). Future work should incorporate objective workflow and clinical metrics, as well as measures (eg, time-motion or productivity metrics), and longer follow-up to more rigorously assess efficiency and maintenance over time.

Recommendations for Future Research

Future work should evaluate operational outcomes such as time to key treatment steps, missed appointment rates, task completion latency, and call volumes. Studies should assess workload and burden using validated measures, paired with time-and-motion studies or log data to capture changes in work patterns. Longitudinal research is needed to track sustained use and maintenance and to understand the sustainability of MyJourney and what is required to ensure it becomes part of routine care across different clinical settings. Equity impacts warrant close attention, including experiences of patients with lower health or digital literacy and the broader accessibility of the patient app. Resource limitations prevented us from evaluating the patient app, so assessing this feature of MyJourney will be an essential step. Comparative research across different clinical settings, including rural or resource-limited sites, could also be undertaken to test MyJourney's scalability potential. Finally, integrating use analytics from the clinical tool and the patient app with qualitative accounts will clarify mechanisms of effect and guide ongoing optimization.

Conclusions

Results from our research demonstrated that MyJourney's Clinical Navigation Tool may contribute to workflow efficiency, care coordination, and information management for breast cancer teams at the NYGH's BDC and CC. The high usability scores and positive user experiences at these clinics support the MyJourney platform's potential for scaling and integration into other settings to optimize cancer care pathways. Continued refinement of digital platforms based on user feedback and expanding integration with electronic medical records will be essential to maximize their benefit and advance high-quality, multidisciplinary oncology care.

Acknowledgments

The authors acknowledge Abbas Ali for helping with technology development, Sana Butt for designing the features of MyJourney based on user feedback, and Rathini Sivasubramaniam and Téjia Bain for their help in the quantitative and qualitative data collection. Artificial intelligence tools were used at any stage in the preparation of this manuscript.

Funding

This work was supported by an Ontario Centre of Excellence award and an Innovation Digital Health Solutions Fund through the Ontario Centre of Innovation.

Data Availability

All data generated or analyzed during this study are included in this published paper and its multimedia appendices.

Authors' Contributions

FO conceived the study, and IH, JM, LH, KM, and MK conducted the research and analyzed the data. All authors contributed to the interpretation. MK and IH wrote the initial draft of the manuscript, and all authors reviewed, revised, and approved the final version.

Conflicts of Interest

FO is the founder of MyJourney Health. All other authors have no conflicts of interest to declare.

XSL•FO
RenderX

Multimedia Appendix 1
Phase 1 interview guide.
[DOCX File, 18 KB - cancer_v12i1e87973_app1.docx]

Multimedia Appendix 2
Phase 2 interview guides.
[DOCX File, 1242 KB - cancer_v12i1e87973_app2.docx]

Multimedia Appendix 3
Phase 3: evaluation surveys for the Breast Diagnostic and Chemotherapy Clinics.
[DOCX File, 2245 KB - cancer_v12i1e87973_app3.docx]

Multimedia Appendix 4
Phase 1: themes and relevant quotes describing participants' experience throughout their breast cancer journey stages.
[DOCX File, 19 KB - cancer_v12i1e87973_app4.docx]

Multimedia Appendix 5
Phase 1: features suggested by participants to include in a future patient tool.
[DOCX File, 14 KB - cancer_v12i1e87973_app5.docx]

Multimedia Appendix 6
Minimum viable feature set that was used to build the patient-facing app.
[DOCX File, 19 KB - cancer_v12i1e87973_app6.docx]

Multimedia Appendix 7
Phase 2: challenges and recommendations as reported by participants—before starting chemotherapy.
[DOCX File, 33 KB - cancer_v12i1e87973_app7.docx]

Multimedia Appendix 8
Phase 2: challenges and recommendations as reported by participants—during chemotherapy.
[DOCX File, 25 KB - cancer_v12i1e87973_app8.docx]

Multimedia Appendix 9
Phase 2: challenges and recommendations as reported by participants—end of chemotherapy and care transition.
[DOCX File, 14 KB - cancer_v12i1e87973_app9.docx]

Multimedia Appendix 10
Phase 3: usability evaluation results with representative quotes.
[DOCX File, 19 KB - cancer_v12i1e87973_app10.docx]

References

1. Public Health Agency of Canada (PHAC). Breast cancer. Government of Canada. 2023. URL: https://www.canada.ca/en/public-health/services/chronic-diseases/cancer/breast-cancer.html [accessed 2026-04-29]
2. Breast cancer statistics. Canadian Cancer Society. 2023. URL: https://cancer.ca/en/cancer-information/cancer-types/breast/statistics [accessed 2026-05-15]
3. Wagle NS, Nogueira L, Devasia TP, et al. Cancer treatment and survivorship statistics, 2025. CA Cancer J Clin 2025;75(4):308-340. [doi: 10.3322/caac.70011] [Medline: 40445120]
4. Rogers CC, Pope S, Whitfield F, Cohn WF, Valdez RS. The lived experience during the peri-diagnostic period of breast cancer: a scoping review. Patient Educ Couns 2022 Mar;105(3):547-585. [doi: 10.1016/j.pec.2021.06.017] [Medline: 34210570]
5. Williams F, Jeanetta SC. Lived experiences of breast cancer survivors after diagnosis, treatment and beyond: qualitative study. Health Expect 2016 Jun;19(3):631-642. [doi: 10.1111/hex.12372] [Medline: 25953316]
6. Campbell-Enns HJ, Woodgate RL, Chochinov HM. Barriers to information provision regarding breast cancer and its treatment. Support Care Cancer 2017 Oct;25(10):3209-3216. [doi: 10.1007/s00520-017-3730-8] [Medline: 28439727]

XSL•FO
RenderX

7. Loo S, Mullikin K, Robbins C, et al. Patient navigator team perceptions on the implementation of a citywide breast cancer patient navigation protocol: a qualitative study. BMC Health Serv Res 2022 May 21;22(1):683. [doi: 10.1186/s12913-022-08090-3] [Medline: 35597947]
8. Silva V, Cardoso J, Fonseca J, Cruz-Correia R. Study of clinical workflow and information flow of a breast care unit. Presented at: Proceedeings of the Third International Conference on Health Informatics; Jan 20-23, 2010. [doi: 10.5220/0002744502940301]
9. Charalambous A. Utilizing the advances in digital health solutions to manage care in cancer patients. Asia Pac J Oncol Nurs 2019;6(3):234-237. [doi: 10.4103/apjon.apjon_72_18] [Medline: 31259218]
10. Kruse CS, Stein A, Thomas H, Kaur H. The use of electronic health records to support population health: a systematic review of the literature. J Med Syst 2018 Sep 29;42(11):214. [doi: 10.1007/s10916-018-1075-6] [Medline: 30269237]
11. Casanova NL, LeClair AM, Xiao V, et al. Development of a workflow process mapping protocol to inform the implementation of regional patient navigation programs in breast oncology. Cancer 2022 Jul 1;128(Suppl 13):2649-2658. [doi: 10.1002/cncr.33944] [Medline: 35699611]
12. Mertz BG, Dunn-Henriksen AK, Kroman N, et al. The effects of individually tailored nurse navigation for patients with newly diagnosed breast cancer: a randomized pilot study. Acta Oncol 2017 Dec 2;56(12):1682-1689. [doi: 10.1080/0284186X.2017.1358462]
13. Auschra C. Barriers to the integration of care in inter-organisational settings: a literature review. Int J Integr Care 2018 Jan 16;18(1):5. [doi: 10.5334/ijic.3068] [Medline: 29632455]
14. Nundy S, Cooper LA, Mate KS. The Quintuple Aim for health care improvement: a new imperative to advance health equity. JAMA 2022 Feb 8;327(6):521-522. [doi: 10.1001/jama.2021.25181] [Medline: 35061006]
15. Thorne S. Interpretive Description: Qualitative Research for Applied Practice, 2nd edition: Routledge; 2016. [doi: 10.4324/9781315545196]
16. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
17. Sullivan-Bolyai S, Bova C, Harper D. Developing and refining interventions in persons with health disparities: the use of qualitative description. Nurs Outlook 2005;53(3):127-133. [doi: 10.1016/j.outlook.2005.03.005] [Medline: 15988449]
18. Holtrop JS, Estabrooks PA, Gaglio B, et al. Understanding and applying the RE-AIM framework: clarifications and resources. J Clin Trans Sci 2021;5(1):e126. [doi: 10.1017/cts.2021.789]
19. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
20. Cancer care. North York General Hospital. URL: https://www.nygh.on.ca/care-service/cancer-care/ [accessed 2025-09-21]
21. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544. [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
22. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018 Jul;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8]
23. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum-Comput Interact 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
24. Brooke J. SUS—a quick and dirty usability scale. In: Usability Evaluation in Industry: Taylor & Francis; 1996:4-7. [doi: 10.1201/9781498710411-35]
25. Gagnon MP, Orruño E, Asua J, Abdeljelil AB, Emparanza J. Using a modified technology acceptance model to evaluate healthcare professionals' adoption of a new telemonitoring system. Telemed J E Health 2012;18(1):54-59. [doi: 10.1089/tmj.2011.0066] [Medline: 22082108]
26. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
27. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? Hum Factors 1992 Aug;34(4):457-468. [doi: 10.1177/001872089203400407]
28. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
29. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009 May 1;4(3):114-123. [doi: 10.5555/2835587.2835589]
30. Kirsch EP, Kunte SA, Wu KA, et al. Digital health platforms for breast cancer care: a scoping review. J Clin Med 2024 Mar 27;13(7):1937. [doi: 10.3390/jcm13071937] [Medline: 38610702]
31. Phillips S, Raskin S, Zhang Y, Pratt-Chapman M. Perspectives from oncology patient navigation programs on information management practices and needs: a descriptive study. Support Care Cancer 2020 Feb;28(2):515-524. [doi: 10.1007/s00520-019-04837-7] [Medline: 31073852]
32. McMullen L. Oncology nurse navigators and the continuum of cancer care. Semin Oncol Nurs 2013 May;29(2):105-117. [doi: 10.1016/j.soncn.2013.02.005] [Medline: 23651680]

XSL•FO
RenderX

33. Rohan EA, Slotman B, Tristani EG, et al. Evaluating the feasibility of using a mobile app to track oncology patient navigation activities and outcomes. J Oncol Navig Surviv 2019 Mar;10(3):98-104. [Medline: 37200970]
34. Chen M, Wu VS, Falk D, Cheatham C, Cullen J, Hoehn R. Patient navigation in cancer treatment: a systematic review. Curr Oncol Rep 2024 May;26(5):504-537. [doi: 10.1007/s11912-024-01514-9] [Medline: 38581470]
35. Aapro M, Bossi P, Dasari A, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct;28(10):4589-4612. [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
36. Jacobs F, D'Amico S, Zazzetti E, et al. Digital innovations in breast cancer care: exploring the potential and challenges of digital therapeutics and clinical decision support systems. Digit Health 2024;10:20552076241288821. [doi: 10.1177/20552076241288821] [Medline: 39502478]
37. Institute of Medicine. Delivering High-Quality Cancer Care: Charting a New Course for a System in Crisis: National Academies Press; 2013. [doi: 10.17226/18359]
38. Thomas Craig KJ, Willis VC, Gruen D, Rhee K, Jackson GP. The burden of the digital environment: a systematic review on organization-directed workplace interventions to mitigate physician burnout. J Am Med Inform Assoc 2021 Apr 23;28(5):985-997. [doi: 10.1093/jamia/ocaa301]
39. Shah SJ, Devon-Sand A, Ma SP, et al. Ambient artificial intelligence scribes: physician burnout and perspectives on usability and documentation burden. J Am Med Inform Assoc 2025 Feb 1;32(2):375-380. [doi: 10.1093/jamia/ocae295]
40. Mohindra NA, Coughlin A, Kircher S, et al. Implementing a patient-reported outcome dashboard in oncology telemedicine encounters: clinician and patient adoption and acceptability. JCO Oncol Pract 2024 Mar;20(3):409-418. [doi: 10.1200/OP.23.00493] [Medline: 38207229]
41. Arora S, Ryals C, Rodriguez JA, Byers E, Clewett E. Leveraging digital technology to reduce cancer care inequities. Am Soc Clin Oncol Educ Book 2022 Apr;42:559-566. [doi: 10.1200/EDBK_350151] [Medline: 35503982]
42. Katsaros D, Hawthorne J, Patel J, Pothier K, Aungst T, Franzese C. Optimizing social support in oncology with digital platforms. JMIR Cancer 2022 Jun 24;8(2):e36258. [doi: 10.2196/36258] [Medline: 35749161]
43. Arditi C, Rège-Walther M, Durieux P, Burnand B. Computer-generated reminders delivered on paper to healthcare professionals: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2017 Jul 6;7(7):CD001175. [doi: 10.1002/14651858.CD001175.pub4] [Medline: 28681432]

Abbreviations

BDC: Breast Diagnostic Centre
CC: Chemotherapy Clinic
NYGH: North York General Hospital
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance
SRQR: Standards for Reporting Qualitative Research
SUS: System Usability Scale

Edited by M Balcarras; submitted 17.Nov.2025; peer-reviewed by J Zhao, O Afolayan; revised version received 30.Mar.2026; accepted 13.Apr.2026; published 10.Jun.2026.

Please cite as:
Herrington I, Makarski J, Amog K, Hayden L, Mossman K, Abbo N, Jedrzejko N, Osman F, Kastner M
Supporting Patients With Breast Cancer and Providers Through Treatment and Survivorship: Multimethod Implementation Study of the MyJourney Platform
JMIR Cancer 2026;12:e87973
URL: https://cancer.jmir.org/2026/1/e87973
doi:10.2196/87973

© Isabella Herrington, Julie Makarski, Krystle Amog, Leigh Hayden, Kate Mossman, Norna Abbo, Nicole Jedrzejko, Fahima Osman, Monika Kastner. Originally published in JMIR Cancer (https://cancer.jmir.org), 10.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

ChatGPT Versus DeepSeek for Breast Cancer Information Retrieval: Quantitative Comparative Study

Rima Hajjo[1], PhD; Dima A Sabbah[1], PhD; Sanaa K Bardaweel[2], PhD

[1]Department of Pharmacy, Faculty of Pharmacy, Al-Zaytoonah University of Jordan, Airport Street, P.O. Box 130, Amman, Jordan

[2]Department of Pharmaceutical Sciences, School of Pharmacy, University of Jordan, Amman, Jordan

Corresponding Author:
Rima Hajjo, PhD
Department of Pharmacy, Faculty of Pharmacy, Al-Zaytoonah University of Jordan, Airport Street, P.O. Box 130, Amman, Jordan

Abstract

Background: Artificial intelligence (AI) is increasingly used to generate medical content, yet its performance in delivering clinically relevant and reliable information remains underexplored, especially in complex areas such as breast cancer.

Objective: This study aimed to compare ChatGPT-4.0 and DeepSeek-V3 in generating breast cancer information, focusing on readability, content quality, and citation reliability.

Methods: On the basis of publicly available patient education materials, 10 frequently asked questions were selected. Each model generated 60 responses. Three expert reviewers rated each response using a 7-point Likert scale across 5 dimensions (ie, accuracy, completeness, clarity, depth and insight, and alignment with expert answers). Readability was assessed using Flesch-Kincaid Grade Level scores. Information reliability was evaluated through interrater agreement metrics, including Cohen κ and Fleiss κ. Paired *t* tests were used for statistical comparisons.

Results: AI models produced significantly more readable content than expert references (mean Flesch-Kincaid Grade Level difference −2.60; $P<.001$). ChatGPT-4.0 responses were more stylistically consistent with a median Flesch-Kincaid Grade Level score of 10.66 (IQR 0.98), whereas DeepSeek-V3 showed greater variability with a median Flesch-Kincaid Grade Level score of 10.17 (IQR 1.41). Content quality scores were DeepSeek-V3 achieving a higher mean score than ChatGPT-4.0 (6.22 [SD 0.43] vs 6.01 [SD 0.49]). In the multiresponse analysis, DeepSeek-V3 demonstrated a statistically significant advantage in accuracy ($P=.041$), while differences across other criteria were not statistically significant ($P>.05$). Human raters showed almost perfect agreement when judging source reliability (Fleiss κ=0.842 for ChatGPT's citations and 0.935 for DeepSeek's citations). Agreement between each model's citation reliability scores and the expert majority was substantial for ChatGPT (Cohen κ=0.665) and higher for DeepSeek (Cohen κ=0.782).

Conclusions: Both models generated readable and clinically relevant content with comparable overall performance. ChatGPT provided more consistent readability, while DeepSeek offered more diverse references with stronger alignment to expert ratings. Continued evaluation and quality assurance are essential for the responsible clinical use of AI-generated content.

(JMIR Cancer 2026;12:e72839) doi:10.2196/72839

KEYWORDS

artificial intelligence; breast cancer; ChatGPT; DeepSeek; large language models; AI; LLM

Introduction

Breast cancer is the most commonly diagnosed cancer among women and remains a major global health burden. Despite advancements in detection and treatment, late-stage presentation continues to hinder outcomes [1-4]. The World Health Organization highlights the need for improved screening and care access [5], while socioeconomic factors such as education and income shape care-seeking behavior. Many women turn to the internet for information on early signs and symptoms [6,7], although online health content is often inaccurate or misleading [8,9].

In response, large language models (LLMs), such as ChatGPT and DeepSeek, have emerged as potential tools for generating medical information and answering patient queries [10-17]. Trained on extensive biomedical literature and general data sources, these generative artificial intelligence (AI) models can produce natural language responses to diverse health-related questions [18,19], potentially supporting patient education and access to general medical knowledge [20]. Promising results have been demonstrated across multiple areas, including medical imaging, disease detection, drug discovery, and personalized oncology, with some AI models surpassing traditional diagnostic accuracy [21-24]. In oncology, AI aids in drug delivery, genomics integration, and real-time monitoring to improve outcomes [2,5-8,25,26]. Nonetheless, concerns persist regarding

XSL•FO
RenderX

the accuracy, readability, and reliability of AI-generated content, especially in complex domains such as breast cancer, necessitating further validation studies [27-30].

A recent work by Ye et al [31] systematically evaluated ChatGPT-3.5's performance in responding to frequently asked questions (FAQs) on cervical and breast cancer. Their study used 10 validated FAQs developed through expert consensus and compared ChatGPT's responses to those of experienced gynecologists and mammography specialists. Using a 7-point Likert scale, they assessed multiple dimensions of response quality, including accuracy, readability, consistency, reliability, efficiency, and relevance. While ChatGPT 3.5 demonstrated high accuracy, often comparable to or exceeding physician responses, it also exhibited limitations in readability, consistency, and the reliability of cited sources. Although foundational, this study did not evaluate more advanced generative AI models, such as ChatGPT-4.0 or DeepSeek-V3, whose capabilities remain largely unexplored [31-33]. Thus, further research is needed to assess how newer models compare in generating high-quality, clinically relevant patient education content.

Therefore, the aim of this study was to systematically evaluate and compare ChatGPT-4.0 and DeepSeek-V3 in their ability to generate high-quality, patient-facing breast cancer information. Using an established, expert-validated set of breast cancer FAQs, we assessed each model across multiple dimensions of response quality, including readability, accuracy, completeness, clarity, depth, insight, alignment with expert reference answers, and citation reliability, with the goal of determining whether these systems can provide trustworthy, clinically relevant educational content for patients.

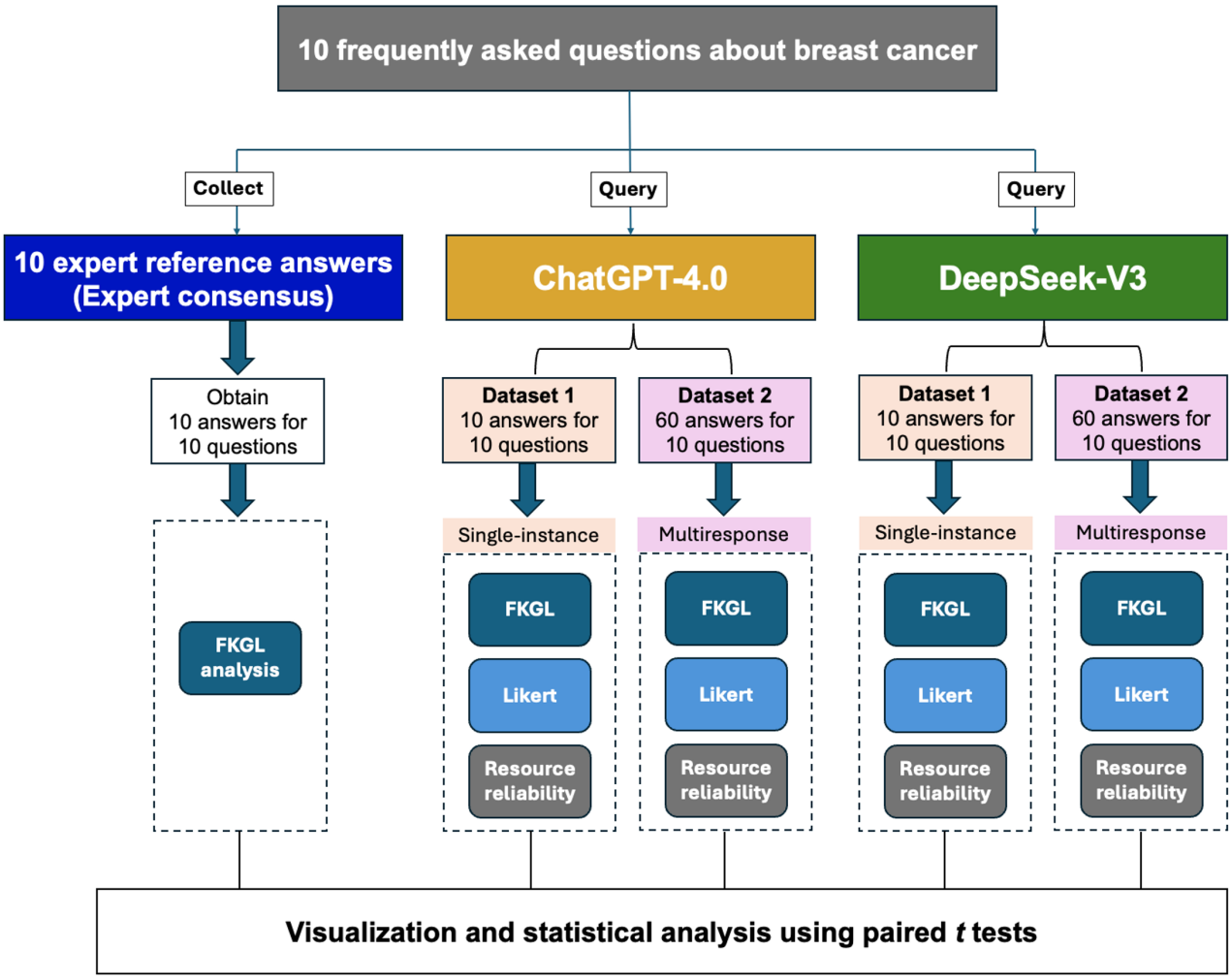
Methods

Study Design and Workflow

This study evaluated the performance of generative AI models in answering 10 FAQs related to breast cancer as described in the workflow shown in Figure 1. The 10 validated questions and their corresponding expert reference answers were adopted from the study by Ye et al [31], which served as the benchmark dataset for this evaluation. In their foundational study, Ye et al [31] developed a list of FAQs regarding breast cancer and cervical cancer based on popular science materials and current hot topics related to tertiary prevention. A set of 10 questions was designed for each cancer type. Subsequently, experts were invited to review the proposed questions and to formulate expert consensus answers. The final set of questions and expert consensus answers were determined through several rounds of consultation and assessment.

XSL•FO
RenderX

Figure 1. Workflow for evaluating the performance of ChatGPT and DeepSeek in retrieving and presenting medical information. FKGL: Flesch-Kincaid Grade Level.

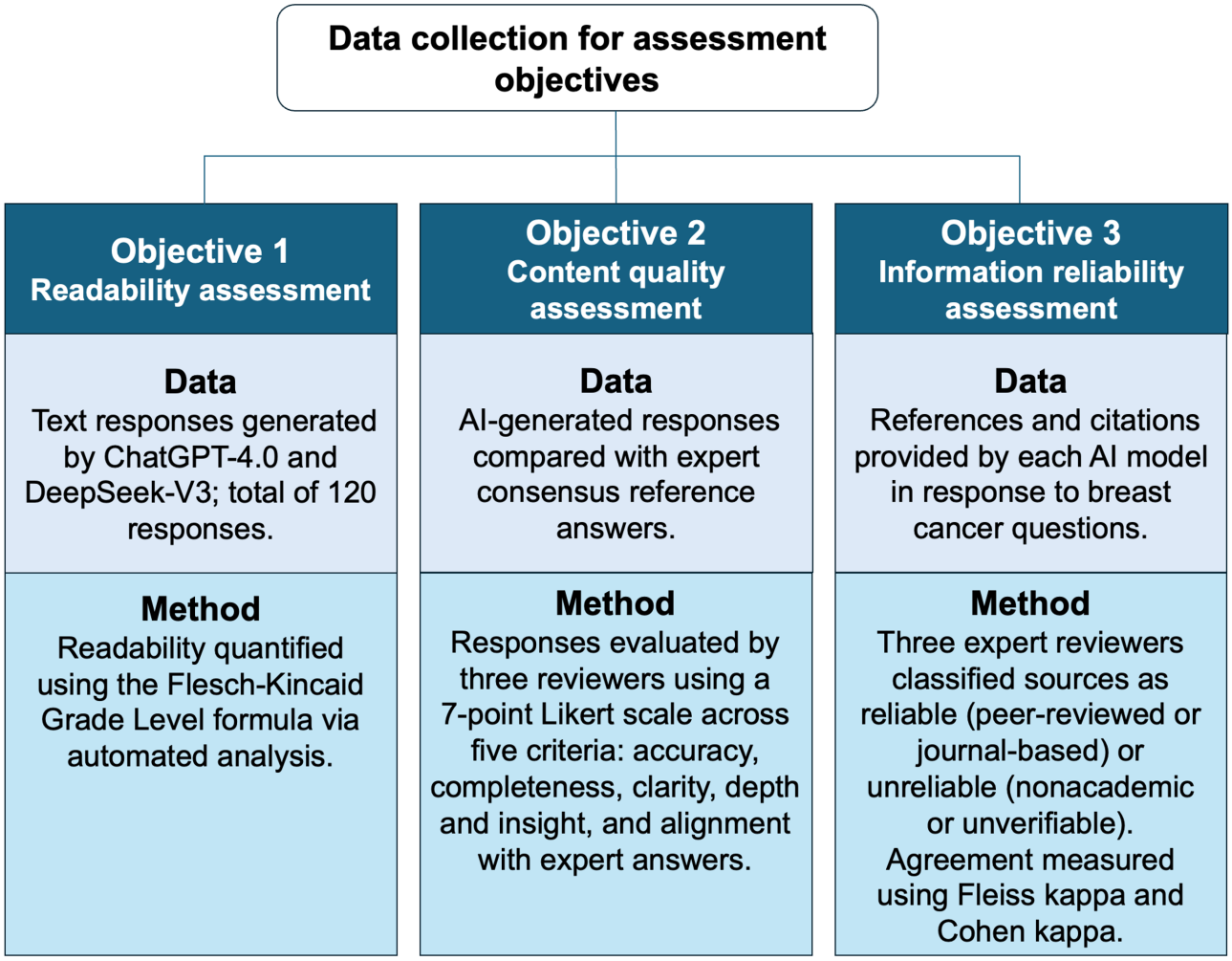


Data Collection and Processing

Figure 2 summarizes the data collection and assessment procedures used to evaluate the performance of ChatGPT-4.0 and DeepSeek-V3 across three core dimensions: readability, content quality, and information reliability. Each objective shown in Figure 2 was addressed using a tailored methodology aligned with specific types of data and evaluation criteria.

XSL•FO
RenderX

Figure 2. Workflow of data collection for assessment objectives. AI: artificial intelligence.



AI Response Generation

To evaluate AI-generated content, we submitted 10 breast cancer FAQs to two generative AI models: ChatGPT-4.0 (OpenAI) and DeepSeek-V3 (DeepSeek). To ensure reproducibility and minimize bias, 3 independent researchers queried each model twice per question at different times, locations, and internet connections, yielding 6 responses per question per model (total: 60 responses per model and 120 responses overall).

Expert Reference Answers

Expert reference answers for each FAQ were obtained from the consensus of 5 board-certified physicians specializing in mammography, as reported by Ye et al [31]. These physicians independently reviewed each FAQ and provided responses, which were then harmonized into a single consensus answer per question. This consensus served as the gold standard for evaluating AI performance.

The 10 breast cancer FAQs and the corresponding AI-generated answers from each independent researcher are provided in Tables S1-S3 in Multimedia Appendices 1-3, and the expert consensus reference answers are provided in Table S4 in Multimedia Appendix 4.

Data Analysis

Readability Assessment

The readability evaluation was performed using the Flesch-Kincaid Grade Level (FKGL) analysis [34,35]. The textstat library (version 0.7.7; developed by Shivam Bansal, open-source project) [36] in Python was used to compute the FKGL scores for expert reference answers and responses generated by ChatGPT-4.0 and DeepSeek-V3. FKGL estimates the US school grade level required to understand a text, with higher scores indicating greater complexity. The FKGL score was calculated using the following formula:

$$FKGL = 0.39\left(\frac{\text{Total words}}{\text{Total sentences}}\right) + 11.8\left(\frac{\text{Total syllables}}{\text{Total words}}\right) - 15.59$$

This analysis used two datasets: dataset 1, where expert reference answers were compared to the initial responses from ChatGPT-4.0 or DeepSeek-V3 for all questions (Q1-Q10); and dataset 2, where 3 researchers generated 2 responses each from ChatGPT-4.0 and DeepSeek-V3 for Q1-Q10, resulting in 60 responses per model. To evaluate statistical significance, paired *t* tests were conducted using the SciPy library (scipy.stats library; SciPy Developers, open-source project) within Python [37], as FKGL scores were normally distributed. These tests assessed mean differences between ChatGPT-4.0 and DeepSeek-V3 while ensuring paired comparisons across identical question sets.

XSL•FO
RenderX

Content Quality Assessment

Content quality was evaluated across five predefined dimensions: accuracy, completeness, clarity, depth and insight, and alignment with expert reference answers. These dimensions were adapted from established guidelines for assessing the quality of health information and AI-generated content [38,39]. Each AI-generated answer was independently scored by 3 expert reviewers using a 7-point Likert-type scale, where 1 represented poor performance and 7 represented excellent performance.

To compare the performance of ChatGPT-4.0 and DeepSeek-V3 across these dimensions, statistical analysis was performed using Python's SciPy library (scipy.stats module, version 1.15.1; SciPy Developers, open-source project) [37]. The Shapiro-Wilk test was first applied to assess the normality of score distributions [40,41]. As the data met the normality assumptions, paired *t* tests [42] were conducted to compare mean scores across the 5 evaluation criteria. This approach enabled a robust and consistent comparison of content quality between the 2 models using identical question sets and standardized scoring methods.

Information Reliability Assessment

The information sources cited by ChatGPT-4.0 and DeepSeek-V3 were independently evaluated by 3 domain experts (RH, DAS, and SKB) with expertise in medicinal chemistry (all), molecular oncology (SKB), and computational drug discovery (RH and DAS). Each unique cited reference was assigned a binary reliability score: 1 for reliable sources (peer-reviewed journal articles, clinical guidelines, or official or authoritative bodies, such as national health agencies or major academic medical centers) and 0 for unreliable sources (commercial health media, foundation blogs, unsourced claims, or otherwise nonverifiable material). Complete source lists and scores are provided in Table S5 in Multimedia Appendix 5 (ChatGPT-4.0) and Table S6 in Multimedia Appendix 6 (DeepSeek-V3), and the full question-level reference data for both ChatGPT-4.0 and DeepSeek-V3 are provided in Table S7 in Multimedia Appendix 7.

For each AI model, we then performed 2 levels of agreement analysis. First, we assessed interrater reliability among the three human experts using Fleiss κ [43], and we also quantified the proportion of references with perfect agreement (all 3 raters assigned the same score), majority agreement (at least two raters agreed on the same score), and complete disagreement. Second, we evaluated how well each model aligned with human judgment by comparing the model's score for each reference to the majority human score (the score agreed upon by ≥2 raters). Agreement between each AI model and this consensus was quantified using Cohen κ [44].

To further characterize disagreement, we identified references where ChatGPT or DeepSeek diverged from the expert majority score and summarized the frequency and nature of those disagreements. All scoring data (individual rater scores, AI-assigned scores, and majority scores) and all agreement calculations (Fleiss κ and Cohen κ) were generated in Python (pandas, NumPy, and scikit-learn) [45-47] and compiled into Microsoft Excel for audit trails, including disagreement tables and summary statistics for each model.

This evaluation framework allowed us to simultaneously measure (1) consistency among human raters when judging citation reliability and (2) how closely each AI system's choices agreed with what experts considered acceptable evidence.

Ethical Considerations

This study did not involve human participants, patient data, or intervention and therefore did not require review or approval by an institutional review board or research ethics committee. The analyses were based exclusively on AI-generated text responses and publicly available information sources, without collection of any personal, identifiable, or sensitive data. As no human subjects were recruited or interacted with, informed consent was not required, and no consent waiver was necessary. No compensation was provided. All data were handled in accordance with principles of data minimization and confidentiality, and no private or proprietary information was accessed or stored during the study.

Results

Study Design and Analytical Framework

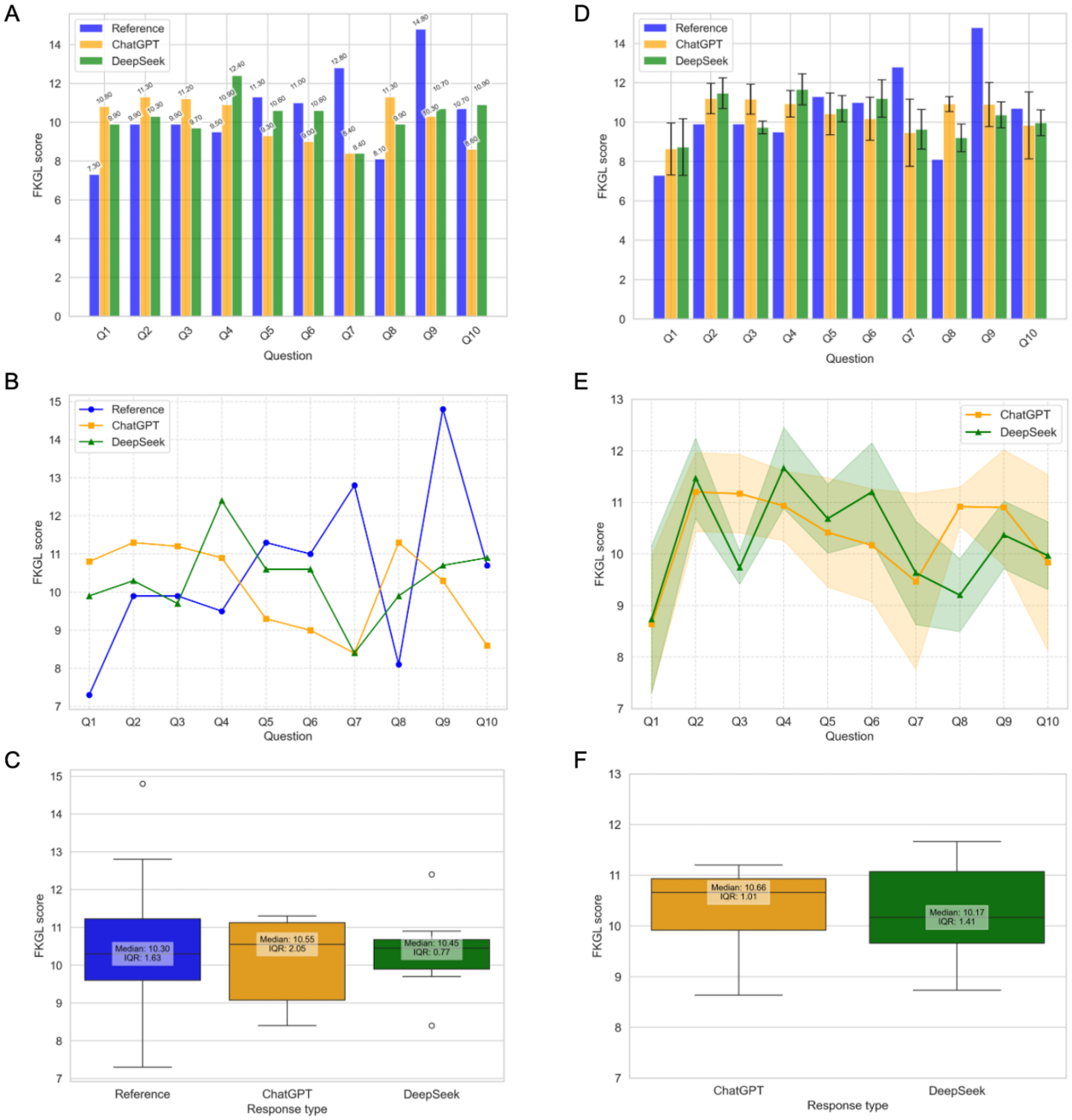
To evaluate the performance of ChatGPT and DeepSeek, a structured analytical framework was implemented, as illustrated in Figure 1. A comparison of ChatGPT and DeepSeek was conducted using 10 FAQs about breast cancer, with responses evaluated against consensus expert reference answers as outlined in the Methods section. The workflow, illustrated in Figure 1, consisted of three key steps: data collection, model querying, and visual and statistical analysis. In the data collection phase, 10 common breast cancer questions were identified, and expert consensus answers were established. Three independent researchers then queried ChatGPT and DeepSeek, generating responses. Two datasets were created for each model: dataset 1, consisting of single-instance responses (1 response per question); and dataset 2, containing multiresponse outputs (multiple responses per question), resulting in 10 responses from each model for dataset 1 and 60 responses from each model for dataset 2Figure 1.

Readability Assessment

The FKGL readability scores for responses generated by ChatGPT and DeepSeek, compared against expert reference answers, are presented in Tables 1 and 2 and Figure 3. Two types of analyses were performed: single-instance analysis, where each AI model provided a single response per question; and multiresponse analysis, where multiple responses per question were collected by 3 researchers independently from both models.

XSL•FO
RenderX

Figure 3. Readability assessment using Flesch-Kincaid Grade Level (FKGL) scores for expert reference answers, ChatGPT, and DeepSeek responses. (A) Bar plot comparing FKGL scores for 10 questions using 1 response per question from each platform. The bars represent the FKGL scores for reference, ChatGPT, and DeepSeek. (B) Line plot showing the overall trends of FKGL scores across all questions using 1 response per question from each platform. The lines represent the FKGL scores for reference, ChatGPT, and DeepSeek. (C) Box plot showing the distribution of FKGL scores using 1 response per question from each platform. The boxes represent the IQR, with the median annotated inside each box. Reference, ChatGPT, and DeepSeek are represented in blue, orange, and green, respectively. (D) Bar plot comparing FKGL scores for 10 questions using 6 responses per question per platform. The bars represent the average FKGL scores, with error bars indicating the SD. ChatGPT responses are shown in orange, and DeepSeek responses are shown in green. (E) Line plot showing the overall trends of FKGL scores for ChatGPT and DeepSeek responses across 10 questions using 6 responses per question per platform. The solid lines represent the mean FKGL scores, and the shaded regions indicate the SDs. (F) Box plot showing the distribution of FKGL scores for ChatGPT and DeepSeek responses across 10 questions using 6 responses per question per platform. The boxes represent the IQR, with the median annotated inside each box. The *y*-axis for all line and box plots ranges from 7 to 15 to ensure a clear comparison of FKGL variations.



XSL•FO
RenderX

Table . Comparison of readability scores of Flesch-Kincaid Grade Level (FKGL) of single-instance analysis for expert reference answers, ChatGPT-4.0, and DeepSeek-V3.

Question	Expert reference answer[a]	ChatGPT-4.0[b]		DeepSeek-V3[c]	
	FKGL score	FKGL score	Difference[d]	FKGL score	Difference[d]
Q1	7.30	10.80	3.50	9.90	2.60
Q2	9.90	11.30	1.40	10.30	0.40
Q3	9.90	11.20	1.30	9.70	−0.20
Q4	9.50	10.90	1.40	12.40	2.90
Q5	11.30	9.30	−2.00	10.60	−0.70
Q6	11.00	9.00	−2.00	10.60	−0.40
Q7	12.80	8.40	−4.40	8.40	−4.40
Q8	8.10	11.30	3.20	9.90	1.80
Q9	14.80	10.30	−4.50	10.70	−4.10
Q10	10.70	8.60	−2.10	10.90	0.20

[a]Statistics across all the responses for 10 questions are as follows: mean 10.53 (SD 2.17) and median 10.30 (IQR 1.63).

[b]Statistics across all the responses for 10 questions are as follows: mean 10.11 (SD 1.17) and median 10.55 (IQR 2.05).

[c]Statistics across all the responses for 10 questions are as follows: mean 10.53 (SD 1.02) and median 10.45 (IQR 0.77).

[d]Difference in comparison to reference. Positive difference indicates that the FKGL score was higher than that of reference, whereas negative difference means that the KFGL scores were lower than that of reference.

Table . Comparison of readability scores of Flesch-Kincaid Grade Level (FKGL) of multiresponse analysis for expert reference answers, ChatGPT-4.0, and DeepSeek-V3.

Question	Expert reference answer[a]	ChatGPT-4.0[b]			DeepSeek-V3[c]		
	FKGL score	FKGL score	Difference[d]	SD	FKGL score	Difference[d]	SD
Q1	7.30	8.63	1.33	1.32	8.73	1.43	1.43
Q2	9.90	11.20	1.30	0.77	11.47	1.57	0.78
Q3	9.90	11.17	1.27	0.76	9.73	−0.17	0.32
Q4	9.50	10.93	1.43	0.67	11.67	2.17	0.79
Q5	11.30	10.42	−0.88	1.06	10.68	−0.62	0.67
Q6	11.00	10.17	−0.83	1.09	11.2	0.20	0.96
Q7	12.80	9.47	−3.33	1.71	9.63	−3.17	1.00
Q8	8.10	10.92	2.82	0.38	9.20	1.10	0.71
Q9	14.80	10.90	−3.90	1.12	10.37	−4.43	0.66
Q10	10.70	9.83	−0.87	1.70	9.97	−0.73	0.66

[a]Statistics across all the responses for 10 questions are as follows: mean 10.53 (SD 2.17) and median 10.30 (IQR 1.63).

[b]Statistics across all the responses for 10 questions are as follows: mean 10.36 (SD 0.84) and median 10.66 (IQR 1.01).

[c]Statistics across all the responses for 10 questions are as follows: mean 10.27 (SD 0.98) and median 10.17 (IQR: 1.41).

[d]Difference in comparison to reference. Positive difference indicates that the FKGL score was higher than that of reference, whereas negative difference means that the FKGL scores were lower than that of reference.

In the single-instance analysis, ChatGPT's mean FKGL score (10.11, SD 1.17) was slightly lower than the expert reference (10.53, SD 2.17), whereas DeepSeek's mean score (10.53, SD 1.02) matched the expert reference answer exactly. In the multiresponse analysis, ChatGPT's mean FKGL increased to 10.36 (SD 0.84), whereas DeepSeek's mean FKGL slightly decreased to 10.27 (SD 0.98). ChatGPT demonstrated reduced variability in the multiresponse case (SD 0.84) compared to the single-instance case (SD 1.17), indicating greater consistency when averaging multiple responses. In contrast, DeepSeek's variability remained relatively stable across both analyses (SD 1.02 in single instance and 0.98 in multiresponse).

In both analyses, DeepSeek's median FKGL scores were slightly lower than ChatGPT's, suggesting that DeepSeek's responses

XSL•FO
RenderX

may be more concise or readable. DeepSeek also exhibited less fluctuation in readability scores, as indicated by its smaller IQR 0.77 (median 10.45) in single instance and 1.41 (median 10.17) in multiresponse. In contrast, ChatGPT's variability decreased in the multiresponse case, with its IQR narrowing from 2.05 (median 10.55) to 1.01 (median 10.66).

A detailed evaluation of FKGL readability scores reveals variations in how ChatGPT and DeepSeek generate responses compared to expert references. For some questions (eg, Q1, Q2, Q3, and Q4), both models produced more complex responses (higher FKGL scores), potentially reducing accessibility for general readers but could be more suitable for professionals and specialists seeking in-depth information.

In contrast, for other questions (eg, Q5, Q6, Q7, and Q9), DeepSeek generally remained closer to the reference readability (reference) but showed greater variability, whereas ChatGPT exhibited larger deviations, sometimes producing more complex or simpler responses (Figure 3A, B, D and E). The multiresponse analysis improved consistency (Figure 3F), smoothing out extreme deviations (eg, outliers evident in Figure 3C) and leading to more balanced readability. Full per-question FKGL values and paired comparisons for single-instance and multiresponse data are provided in Tables S8-S11 in Multimedia Appendices 8-11.

A paired *t* test was conducted to assess the statistical significance of mean differences in FKGL scores among the expert reference answer, ChatGPT-4.0, and DeepSeek-V3 groups (Table 3). The Shapiro-Wilk test confirmed normality ($P>.05$ for all groups), validating the use of the *t* test. Results indicated no significant differences in readability scores between the expert reference answers and those generated by ChatGPT-4.0 or DeepSeek-V3, nor between ChatGPT-4.0 and DeepSeek-V3 ($P=.61$ for single instance and $P=.73$ for multiresponse). However, high variability in scores across questions was observed. Multiresponse analyses showed smaller differences between AI models and the reference, suggesting that averaging multiple responses yields more balanced readability. Normality was further confirmed by Shapiro-Wilk *P* values of .16 (ChatGPT-4.0) and .82 (DeepSeek-V3). The corresponding statistical summaries are provided in Tables S8 and S10 in Multimedia Appendices 8 and 10.

Table . Statistical comparison of readability (FKGL scores) among expert references, ChatGPT-4.0, and DeepSeek-V3 using paired *t* tests.

Comparison	Single-instance			Multiresponse		
	t test *(df)*[a]	*P* value[b]	95% CI[c]	*t* test *(df)*[a]	*P* value[b]	95% CI[c]
ChatGPT-4.0 versus reference	−0.45 (9)	.66	−1.69 to 2.53	−0.24 (9)	.82	−1.40 to 1.74
DeepSeek-V3 versus reference	−0.24 (9)	.81	−1.58 to 1.96	−0.40 (9)	.70	−1.25 to 1.78
ChatGPT-4.0 versus DeepSeek-V3	0.52 (9)	.61	−1.23 to 0.77	0.35 (9)	.73	−0.53 to 0.72

[a]*t* test *(df)*: measures the magnitude and direction of the difference between paired samples. A negative *t* value indicates that ChatGPT had lower scores than DeepSeek, while a positive *t* value indicates the opposite. *t* statistics are reported, all tests were 2-tailed paired *t* tests. Degrees of freedom shown in parentheses, and were calculated as n−1 for paired comparisons (*df*=9 for all single analyses).

[b]*P* value represents the probability of observing the test results under the null hypothesis. A *P* value ≤.05 is considered statistically significant.

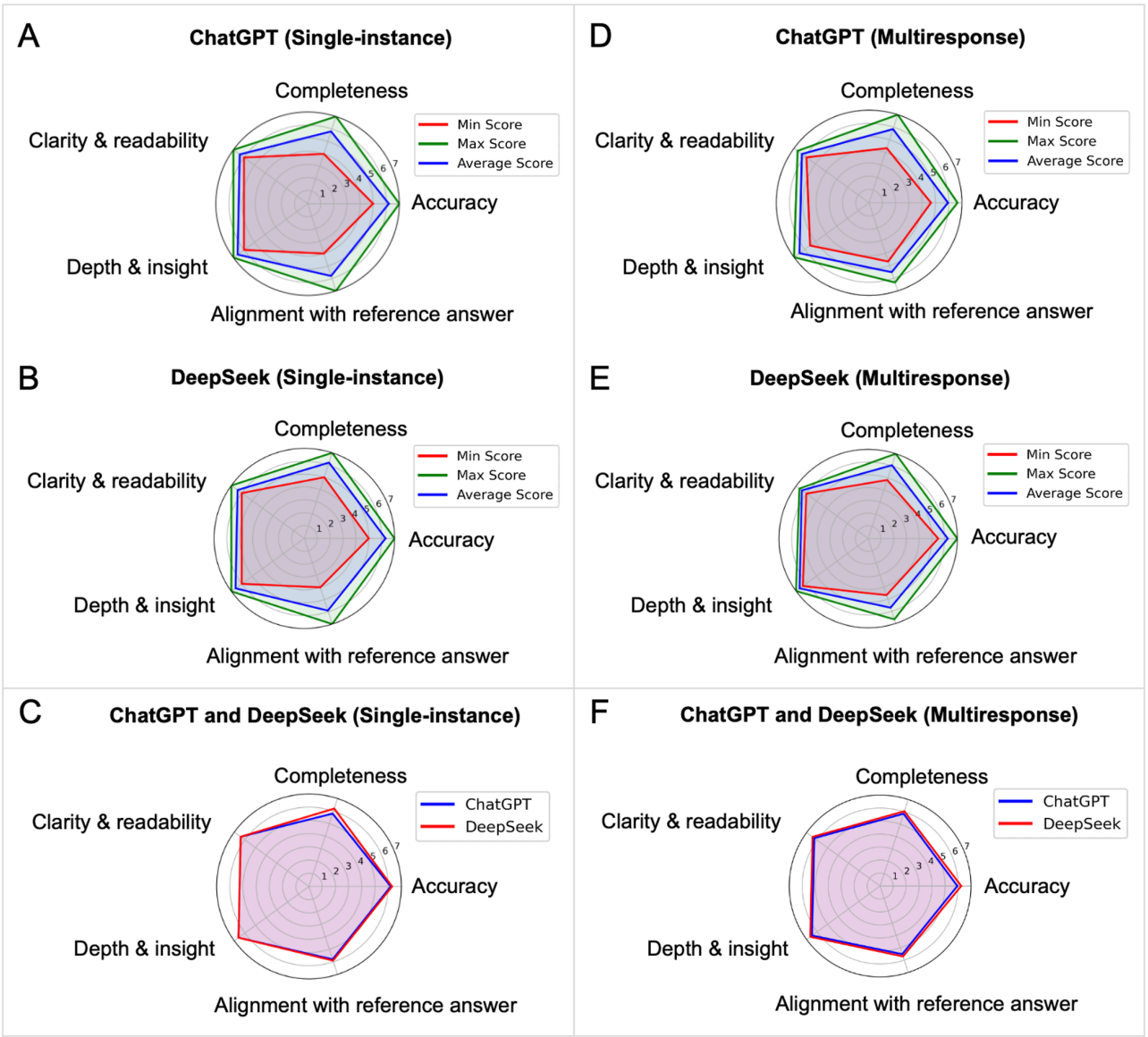
[c]95% CI for the mean difference.

Content Quality Assessment

Response quality was evaluated using a 7-point Likert scale across five criteria: accuracy, clarity and readability, completeness, depth and insight, and alignment with reference answers. In the single-response analysis, ChatGPT achieved an overall mean Likert score of 5.36 (SD 0.53), whereas DeepSeek had a slightly higher mean score of 5.60 (SD 0.45) across all criteria and questions. At the criterion level, ChatGPT scored higher in alignment and accuracy, whereas DeepSeek performed slightly better in the remaining criteria (Figure 4A–C).

XSL•FO
RenderX

Figure 4. Content quality ratings for ChatGPT-4.0 and DeepSeek-V3 across 5 criteria. (A) ChatGPT response evaluation based on the average scores of 10 answers. (B) DeepSeek response evaluation based on average score of 10 answers. (C) ChatGPT versus DeepSeek comparison based on average scores of 10 answers. (D) ChatGPT response evaluation based on 60 answers. (E) DeepSeek response evaluation based on 60 answers. (F) ChatGPT versus DeepSeek comparison based on average scores of 10 answers.



In the multiresponse analysis (dataset 2), ChatGPT achieved a mean Likert score of 6.01 (SD 0.49), while DeepSeek achieved a comparable mean score of 6.22 (SD 0.43) across all criteria and responses (Figure 4D–F). Differences between the two models were minimal and not statistically significant (P>.05), with overlapping 95% CIs for completeness, clarity & readability, depth & insight, and alignment with the reference answer. However, DeepSeek outperformed ChatGPT based on accuracy (ChatGPT: 5.95 vs DeepSeek: 6.25). This difference was statistically significant (t = −2.377, P = .041; 95% CI −0.585 to −0.015), indicating consistently higher accuracy scores for DeepSeek across the 10 questions.Overall, these findings indicate that both models provide comparably high-quality responses across the assessed dimensions. Detailed summary statistics are provided in Table 4, and full per-question and per-criterion Likert ratings for both the single-instance and multiresponse datasets are provided in Table S12 in Multimedia Appendix 12 and Table S13 in Multimedia Appendix 13.

XSL•FO
RenderX

Table . Comparison of ChatGPT-4.0 and DeepSeek-V3 content quality across 5 evaluation criteria using paired *t* tests for single- and multiresponse analyses.

Evaluation criterion	ChatGPT-4.0 score (mean)	DeepSeek-V3 score (mean)	*t* test (*df*)[a]	*P* value[b]	95% CI[c]
Single-instance results					
Accuracy	6.20	6.30	−0.43 (9)	.68	−0.63 to 0.43
Completeness	5.80	6.20	−1.50 (9)	.17	−1.00 to 0.20
Clarity and readability	6.40	6.40	0.00 (9)	1.00	−0.34 to 0.34
Depth and insight	6.60	6.60	0.00 (9)	1.00	−0.34 to 0.34
Alignment with reference answer	5.80	5.90	−0.43 (9)	.68	−0.63 to 0.43
Multiresponse results					
Accuracy	5.95	6.25	−2.38 (9)	.04	−0.59 to −0.01
Completeness	5.85	6.05	−1.91 (9)	.09	−0.44 to 0.04
Clarity and readability	6.27	6.43	−1.73 (9)	.12	−0.38 to 0.05
Depth and insight	6.48	6.65	−1.79 (9)	.11	−0.38 to 0.04
Alignment with reference answer	5.52	5.70	−1.49 (9)	.17	−0.46 to 0.09

[a]*t* test: measures the magnitude and direction of the difference between paired samples. A negative *t* value indicates that ChatGPT had lower scores than DeepSeek, whereas a positive *t* value indicates the opposite, all tests were 2-tailed paired *t* tests Degrees of freedom shown in parentheses, and were calculated as n−1 for paired comparisons (*df*=9 for all analyses).

[b]*P* value represents the probability of observing the test results under the null hypothesis. A *P* value ≤.05 is considered statistically significant.

[c]95% CI of the mean difference.

Information Reliability Assessment

We analyzed the reliability of citations provided by ChatGPT-4.0 and DeepSeek-V3 in response to 10 clinically oriented breast cancer questions. For ChatGPT, 45 distinct references were extracted and scored by all 3 experts. Human raters showed very high internal consistency: Fleiss κ was 0.842, indicating almost perfect agreement. Of these 45 references, 40 (88.9%) received perfect agreement from all 3 raters, and 44 (97.8%) reached majority agreement (ie, at least two raters agreed on the same label). Only 1/45 (2.2%) reference showed complete disagreement among the raters. Detailed results are provided in Table S14 in Multimedia Appendix 14.

For DeepSeek, we evaluated 268 unique references. Interrater reliability among the three experts was even higher: Fleiss κ was 0.935, again in the *almost perfect* range. Across these 268 references, 260 (97.0%) had perfect agreement, and 267 (99.6%) had majority agreement. Only 1/268 references (0.4%) showed complete disagreement. These results confirm that, despite differences in style and depth of citation between the models, human experts were able to judge source reliability in a highly consistent way across both datasets. All details are provided in Table S15 in Multimedia Appendix 15.

We then assessed how well each AI model's citation-level reliability scores aligned with human judgment (ie, each reference set was evaluated by both ChatGPT and DeepSeek as reported in Table S11 and Table S12 in Multimedia Appendices 11 and 12, respectively). For the ChatGPT reference set, agreement between ChatGPT's scores and the human majority score was substantial (Cohen κ=0.665). For the DeepSeek reference set, agreement between DeepSeek's scores and the human majority score was higher (κ=0.782). When disagreement occurred, ChatGPT differed from the human majority on 7 of its 45 references, whereas DeepSeek differed on 32 of its 268 references. These agreement metrics and disagreement profiles are summarized in Table S14 and Table 15 in Multimedia Appendices 7 and 14, respectively.

To enable direct comparison, we also evaluated each model on the other model's reference set. In the ChatGPT set, DeepSeek's scoring of those same sources achieved κ=0.800, which exceeded ChatGPT's own κ=0.665. In the DeepSeek set, ChatGPT's scoring of DeepSeek's citations achieved κ=0.600, which was lower than DeepSeek's κ=0.782. In other words, when both models were evaluated against the same expert consensus, DeepSeek was consistently closer to the human majority than ChatGPT.

Qualitatively, ChatGPT tended to cite large US medical organizations and public-facing health information portals (eg, national cancer societies, specialty associations, and patient education sites). These sources were commonly accepted as reliable when they represented established clinical authorities, but they occasionally triggered disagreement when they

XSL•FO
RenderX

originated from advocacy foundations or commercial health media. DeepSeek more often cited recent peer-reviewed literature, systematic reviews, and oncology-specific clinical studies, which generally scored as reliable, although some entries contained incomplete metadata (eg, partial author or year strings or nonresolving links) that required manual adjudication.

Taken together, these results show 3 things. First, human expert scoring of reference quality was internally very stable for both models (Fleiss κ=0.842 for ChatGPT's reference set and 0.935 for DeepSeek's reference set) with majority agreement observed for 44/45 (97.8%) and 267/268 (99.6%) references, respectively, indicating that *reliable versus unreliable* source is a reproducible judgment in this context. Second, both AI systems demonstrated substantial alignment with expert consensus, but DeepSeek exhibited stronger agreement with the human majority than ChatGPT (κ up to 0.800 on the ChatGPT set and 0.782 on its own set, versus ChatGPT's κ of 0.665 and 0.600, respectively). Third, the main source of residual disagreement was concentrated in borderline cases, such as advocacy organization content, institutional news posts, and partially specified citations, rather than in conventional peer-reviewed oncology literature (Table 5).

Table . Summary of citation reliability and agreement metrics for ChatGPT-4.0 and DeepSeek-V3.

Metric	ChatGPT-4.0	DeepSeek-V3
References evaluated, N	45	268
Reliability score (majority human; mean)	0.64	0.82
Fleiss κ (inter-rater agreement)	0.842	0.935
Perfect agreement (all 3 raters assigned the same label), n/N (%)	40/45 (88.9)	260/268 (97.0)
Majority agreement (≥2 of 3 raters assigned the same label), n/N (%)	44/45 (97.8)	267/268 (99.6)
Complete disagreement (all 3 raters assigned different labels), n/N (%)	1/45 (2.2)	1/268 (0.4)
Cohen κ (AI vs human majority)	0.665	0.782
Cross-evaluation κ (AI scoring the other model's references)	ChatGPT on DeepSeek=0.600	DeepSeek on ChatGPT=0.800

The mean reliability score (majority human) is the mean binary reliability score (1=reliable and 0=unreliable) assigned by the human majority (≥2 of 3 expert raters) across all unique references cited by each model. "Fleiss κ" quantifies agreement among the 3 human experts for each model's reference set. "Perfect agreement" is the proportion of references where all 3 experts independently gave the same reliability score. "Majority agreement" is the proportion of references where at least two of the 3 experts agreed. "Complete disagreement" refers to cases in which all 3 experts gave different scores. "Cohen κ (AI vs human majority score)" measures agreement between each model's own reliability classification of its citations and the human majority classification for the same reference set. "Cross-evaluation κ" measures how well each model's scoring of the other model's references agreed with the human majority for that reference set (eg, "DeepSeek on ChatGPT"=DeepSeek's κ when scoring the ChatGPT reference set). All underlying data are provided in Table S14 in Multimedia Appendix 7 (ChatGPT-4.0) and Table S15 in Multimedia Appendices 14 and 15 (DeepSeek-V3).

Discussion

Principal Findings

This study systematically evaluated the performance of ChatGPT-4.0 and DeepSeek-V3 in generating breast cancer information, focusing on readability, content quality, and reference quality. This multidimensional assessment approach (Figure 1), which is rarely addressed collectively in existing literature, offers valuable insights into the capabilities and limitations of generative AI models in health care. Both models generated clinically relevant and readable responses, with ChatGPT-4.0 showing greater consistency and overall readability, whereas DeepSeek-V3 offered more diverse citations but with greater variability and occasional technical issues.

Readability Analysis

The FKGL readability scores revealed that AI-generated responses generally used simpler language structures than expert reference answers, which may enhance accessibility for general audiences. Expert responses had the highest FKGL scores (mean 10.53, SD 2.17), reflecting more complex vocabulary and sentence construction. In contrast, both ChatGPT and DeepSeek produced responses with lower FKGL scores, indicating simpler, more readable content.

This trend aligns with prior studies showing that LLMs often generate linguistically simplified outputs, which may benefit general users [38,39]. Notably, ChatGPT-4.0 exhibited more consistent readability across multiple outputs, with a median FKGL of 10.66 (IQR 0.98), compared to DeepSeek-V3, which showed greater variability with a median FKGL of 10.17 (IQR 1.41), suggesting greater stability with repeated querying.

For instance, ChatGPT-4.0 demonstrated improved readability consistency when averaging multiple responses, suggesting stability with repeated querying. In contrast, DeepSeek-V3's responses were closer to expert reference readability levels in

XSL•FO
RenderX

single-instance analysis but displayed greater variability in multiresponse cases (IQR 1.41).

However, statistical comparisons did not reveal significant differences in readability between models or relative to expert references. For instance, in the multiresponse setting, the paired *t* test for ChatGPT-4.0 versus reference yielded t_9=−0.24; *P*=.82 (95% CI −1.40 to 1.74), and for DeepSeek-V3 versus reference, t_9=−0.40; *P*=.70 (95% CI −1.25 to 1.78). The difference between ChatGPT and DeepSeek was also nonsignificant (t_9=0.35; *P*=.73). Similar nonsignificant findings were observed in the single-instance analysis.

It is essential to emphasize that FKGL measures linguistic complexity, not clinical accuracy or information richness [35]. Lower FKGL scores do not necessarily indicate that critical medical details were omitted. Therefore, while improved readability may facilitate comprehension, further analysis is necessary to confirm that the content remains accurate, complete, and clinically appropriate.

Content Quality Analysis

Beyond readability, content quality was assessed across 5 predefined evaluation criteria, including accuracy, completeness, clarity, depth and insight, and alignment with expert reference answers. This structured evaluation, based on established frameworks [38,39], allowed a more nuanced assessment of response quality. Across analyses, both models achieved high scores across all criteria. In the multiresponse dataset, DeepSeek achieved a higher overall mean score (6.22, SD 0.43) than ChatGPT (6.01, SD 0.49), with DeepSeek scoring numerically higher across all five criteria. Notably, DeepSeek demonstrated a statistically significant advantage in accuracy in the multiresponse analysis (t = −2.377, *P* = .041; 95% CI −0.585 to −0.015), while differences for completeness, clarity and readability, depth and insight, and alignment with the reference answer did not reach statistical significance (P > .05). Interestingly, averaging multiple responses per question reduced response-level variability and revealed a more consistent performance advantage for DeepSeek, particularly for accuracy.

These findings expand on earlier studies that primarily focused on ChatGPT-3.5 [33] by demonstrating the evolving strengths and remaining limitations of newer generative AI models when applied to clinically relevant educational content. Unlike prior studies limited to single-response analysis, our multiresponse approach highlights how response variability remains an important factor when evaluating AI for health care education.

Information Reliability Assessment

The reference reliability analysis revealed clear differences in both reference selection patterns and interrater scoring agreement for ChatGPT-4.0 and DeepSeek-V3 citations. ChatGPT tended to cite US-based medical institutions, national health organizations, and patient-facing educational resources, whereas DeepSeek more often cited recent peer-reviewed oncology studies and systematic reviews, although some DeepSeek citations were partially specified or contained inaccessible links. Human experts showed very consistent judgments of citation reliability for both models' outputs: interrater reliability was almost perfect for both sets (Fleiss κ=0.842 for ChatGPT's 45 references and 0.935 for DeepSeek's 268 references), with ≥97% majority agreement and only one fully disputed reference in each set. When we compared each model's scoring of source reliability to the majority expert score, agreement was substantial for ChatGPT and higher for DeepSeek. DeepSeek reached Cohen κ values up to 0.800 (for ChatGPT reference set) and 0.782 (for DeepSeek reference set), indicating strong alignment with expert consensus, whereas ChatGPT's ability to assess references reached 0.665 for its own reference set and 0.600 when evaluated DeepSeek's set.

Collectively, these results demonstrate that both models operate within the bounds of medically acceptable sourcing but use distinct strategies. DeepSeek's output, favoring primary literature, more accurately reflects the evidence-based curation of oncology experts. Conversely, ChatGPT's propensity for synthesized institutional guidance results in a higher frequency of citations to borderline sources that require expert judgment.

Implications for Practice and Future Research

Collectively, these findings suggest that while generative AI models offer promising capabilities for generating patient education materials, important limitations remain. Oversimplification of complex medical information, variability in response consistency, and the reliability of cited sources remain persistent concerns. Given the growing interest in deploying AI tools for patient counseling, clinical decision support, and health care communication, these results underscore the need for careful validation of AI-generated content before clinical use [48]. Furthermore, while models such as DeepSeek demonstrate the potential for integrating primary literature more effectively, consistent quality control of both content and source reliability will be essential for safe application in clinical settings.

Finally, the structured multidimensional evaluation framework used in this study provides a replicable approach for future evaluations of AI models in health care. Future studies should continue to assess performance across broader clinical domains, expand evaluation across different languages and health literacy levels, and integrate clinical expert validation to further ensure the safety, accuracy, and equity of AI-generated medical information.

Limitations and Future Directions

This study has several limitations. The FKGL test assessed structural readability but not contextual or clinical accuracy, and it is validated only for English, limiting generalizability across languages and literacy levels [20,34]. Additionally, the study was conducted in a controlled setting without patient or provider involvement, so real-world usability, comprehension, and clinical outcomes remain unexplored. Future work should include usability testing and assess impacts on decision-making and adherence [49]. Moreover, AI responses were not compared directly to clinical guidelines, and future studies should include validation against established standards before clinical integration. Reference accuracy remains a concern, as AI models may generate outdated or fabricated citations, requiring expert oversight [50].

XSL•FO
RenderX

Furthermore, the risk of bias in training data may lead to inaccuracies for underrepresented populations, highlighting the importance of ongoing monitoring and alignment with validated medical sources [38,48,49]. Evidence from clinical research, including variability in therapeutic outcomes across patient subgroups and molecular subtypes, further emphasizes the necessity for such validation [51,52]. Finally, ethical and regulatory considerations are essential for safe deployment of AI in health care; compliance with frameworks from regulatory bodies such as the Food and Drug Administration and the European Medicines Agency will be critical to ensure safety, transparency, and public trust [53]. As these technologies continue to evolve, continuous evaluation, comparative benchmarking, and expert oversight will be necessary to guide their responsible implementation.

Conclusions

This comparative analysis of ChatGPT-4.0 and DeepSeek-V3 demonstrated that both AI platforms effectively retrieve and present medical information on breast cancer, each excelling in different areas. ChatGPT produced more polished, detailed, and readable responses, with strong performance in clarity and completeness, whereas DeepSeek outperformed ChatGPT on accuracy, and provided more comprehensive, globally diverse references with in-text citations resembling academic articles. However, DeepSeek faced challenges such as untagged links, occasional downtime, and corrupted references, which impacted the user experience. Despite these limitations, DeepSeek demonstrated superior citation efficiency and closely aligned with expert consensus answers. Moreover, interrater reliability analyses confirmed high levels of agreement among human evaluators across both models, with Fleiss κ scores exceeding 0.84 and Cohen κ values indicating substantial to almost perfect agreement. These findings support the overall quality of reference sources provided by the models and lend credibility to the evaluation framework used in this study. Furthermore, statistical analysis revealed no significant differences between the models across most evaluation criteria; however, in the multiresponse dataset (i.e., larger datasets), DeepSeek demonstrated a statistically significant advantage in accuracy, underscoring the importance of rigorous evaluation of AI-generated medical information to ensure accuracy, accessibility, and reliability. To optimize their effectiveness in health care applications, future improvements should focus on enhancing platform stability, response consistency, and overall user accessibility.

Acknowledgments

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Funding

RH and DAS acknowledge funding provided by the Deanship of Scientific Research at Al-Zaytoonah University of Jordan (grant 2023-2022/17/50 and 2025-2024/06/29).

Data Availability

All data generated or analyzed during this study, including AI-generated responses, reference metadata, and interrater reliability analyses, are publicly available in a GitHub repository [54].

Authors' Contributions

Conceptualization: RH, SKB
Data Curation: RH, DAS, SKB
Formal analysis: RH
Methodology: RH, DAS, SKB
Writing – original draft: RH, DAS, SKB
Writing – review & editing: RH, DAS, SKB

Conflicts of Interest

None declared.

Multimedia Appendix 1
Frequently asked questions about breast cancer and answers generated by artificial intelligence models (ChatGPT and DeepSeek) by Researcher 1.
[DOCX File, 139 KB - cancer_v12i1e72839_app1.docx]

Multimedia Appendix 2
Frequently asked questions about breast cancer and answers generated by artificial intelligence models (ChatGPT and DeepSeek) by Researcher 2.

XSL•FO
RenderX

[DOCX File, 81 KB - cancer_v12i1e72839_app2.docx]

Multimedia Appendix 3
Frequently asked questions about breast cancer and answers generated by artificial intelligence models (ChatGPT and DeepSeek) by Researcher 3.
[DOCX File, 119 KB - cancer_v12i1e72839_app3.docx]

Multimedia Appendix 4
Frequently asked questions about breast cancer and expert consensus answers obtained from Ye et al.
[XLSX File, 15 KB - cancer_v12i1e72839_app4.xlsx]

Multimedia Appendix 5
Human experts and artificial intelligence reliability scores for references cited by ChatGPT-4.0.
[XLSX File, 11 KB - cancer_v12i1e72839_app5.xlsx]

Multimedia Appendix 6
Human experts and artificial intelligence reliability scores for references cited by DeepSeek-V3.
[XLSX File, 22 KB - cancer_v12i1e72839_app6.xlsx]

Multimedia Appendix 7
Information sources for ChatGPT-4.0 and DeepSeek-V3.
[DOCX File, 109 KB - cancer_v12i1e72839_app7.docx]

Multimedia Appendix 8
Average Flesch-Kincaid Grade Level scores for the single-response analysis and results of the paired *t* test.
[XLSX File, 10 KB - cancer_v12i1e72839_app8.xlsx]

Multimedia Appendix 9
Detailed Flesch-Kincaid Grade Level scores for the single-response analysis.
[XLSX File, 10 KB - cancer_v12i1e72839_app9.xlsx]

Multimedia Appendix 10
Average Flesch-Kincaid Grade Level scores for the multiresponse analysis and results of the paired *t* test.
[XLSX File, 11 KB - cancer_v12i1e72839_app10.xlsx]

Multimedia Appendix 11
Detailed Flesch-Kincaid Grade Level scores for the multiresponse analysis.
[XLSX File, 11 KB - cancer_v12i1e72839_app11.xlsx]

Multimedia Appendix 12
Likert scores for the single-instance analysis.
[XLSX File, 12 KB - cancer_v12i1e72839_app12.xlsx]

Multimedia Appendix 13
Likert scores for the multiresponse analysis.
[XLSX File, 15 KB - cancer_v12i1e72839_app13.xlsx]

Multimedia Appendix 14
Interrater evaluation of ChatGPT-4.0 references for breast cancer information based on ratings by 3 human experts and alignment with artificial intelligence models.
[XLSX File, 15 KB - cancer_v12i1e72839_app14.xlsx]

Multimedia Appendix 15
Interrater evaluation of DeepSeek-V3 references for breast cancer information based on ratings by three human experts and alignment with artificial intelligence models.

XSL•FO
RenderX

[XLSX File, 30 KB - cancer_v12i1e72839_app15.xlsx]

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Sinha A, Naskar M, Pandey M, Rautaray SS. Challenges to the early diagnosis of breast cancer: current scenario and the challenges ahead. SN COMPUT SCI 2024;5(1):170. [doi: 10.1007/s42979-023-02534-1]
3. Sunoqrot S, Abusulieh S, Abusara OH. Identifying synergistic combinations of doxorubicin-loaded polyquercetin nanoparticles and natural products: implications for breast cancer therapy. Int J Pharm 2023 Oct 15;645:123392. [doi: 10.1016/j.ijpharm.2023.123392] [Medline: 37683979]
4. Sweidan K, Elfadel H, Sabbah DA, et al. Novel derivatives of 4,6 - dihydroxy - 2 - quinolone - 3 - carboxamides as potential PI3Kα inhibitors. ChemistrySelect 2022 Aug 26;7(32) [FREE Full text] [doi: 10.1002/slct.202202263]
5. Global Breast Cancer Initiative Implementation Framework: Assessing, Strengthening and Scaling-up of Services for the Early Detection and Management of Breast Cancer: World Health Organization; 2023. URL: https://www.who.int/publications/i/item/9789240067134 [accessed 2026-01-14]
6. Kasper G, Momen M, Sorice KA, et al. Effect of neighborhood and individual-level socioeconomic factors on breast cancer screening adherence in a multiethnic study. BMC Public Health 2024 Jan 2;24(1):63. [doi: 10.1186/s12889-023-17252-9] [Medline: 38166942]
7. Chen J, Duan Y, Xia H, Xiao R, Cai T, Yuan C. Online health information seeking behavior among breast cancer patients and survivors: a scoping review. BMC Womens Health 2025 Jan 3;25(1):1. [doi: 10.1186/s12905-024-03509-x] [Medline: 39754199]
8. Loeb S, Langford AT, Bragg MA, Sherman R, Chan JM. Cancer misinformation on social media. CA Cancer J Clin 2024;74(5):453-464. [doi: 10.3322/caac.21857] [Medline: 38896503]
9. Johnson SB, Parsons M, Dorff T, et al. Cancer misinformation and harmful information on Facebook and other social media: a brief report. J Natl Cancer Inst 2022 Jul 11;114(7):1036-1039. [doi: 10.1093/jnci/djab141] [Medline: 34291289]
10. Walker HL, Ghani S, Kuemmerli C, et al. Reliability of medical information provided by ChatGPT: assessment against clinical guidelines and patient information quality instrument. J Med Internet Res 2023 Jun 30;25:e47479. [doi: 10.2196/47479] [Medline: 37389908]
11. Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. Front Artif Intell 2023;6:1169595. [doi: 10.3389/frai.2023.1169595] [Medline: 37215063]
12. Choi J, Kim JW, Lee YS, et al. Availability of ChatGPT to provide medical information for patients with kidney cancer. Sci Rep 2024;14(1):1542. [doi: 10.1038/s41598-024-51531-8]
13. Peng Y, Malin BA, Rousseau JF, et al. From GPT to DeepSeek: significant gaps remain in realizing AI in health care. J Biomed Inform 2025 Mar;163:104791. [doi: 10.1016/j.jbi.2025.104791] [Medline: 39938624]
14. Kincaid JP, Braby R, Mears JE. Electronic authoring and delivery of technical information. J Instr Dev 1988 Jun;11(2):8-13. [doi: 10.1007/BF02904998]
15. Sullivan GM, Artino AR Jr. Analyzing and interpreting data from likert-type scales. J Grad Med Educ 2013 Dec;5(4):541-542. [doi: 10.4300/JGME-5-4-18] [Medline: 24454995]
16. AlZu'bi S, Zreiqat A, Radi W, Mughaid A, Abualigah L. An intelligent health care monitoring system-based novel deep learning approach for detecting covid-19 from x-rays images. Multimed Tools Appl 2024;83(23):63479-63496. [doi: 10.1007/s11042-023-18056-0]
17. Muhairat M, Alzyadat W, Shaheen A, Alhroob A, Asfour AN. Leveraging machine learning for predictive pathways in higher education: a case study at Al-Zaytoonah University of Jordan. SSRG-IJECE 2024;11(11):28-44. [doi: 10.14445/23488549/IJECE-V11I11P104]
18. Jarab AS, Al-Qerem W, Al-Hajjeh DM, et al. Artificial intelligence utilization in the health care setting: perceptions of the public in the UAE. Int J Environ Health Res 2025 Mar;35(3):585-593. [doi: 10.1080/09603123.2024.2363472] [Medline: 38832887]
19. Buch VH, Ahmed I, Maruthappu M. Artificial intelligence in medicine: current trends and future possibilities. Br J Gen Pract 2018 Mar;68(668):143-144. [doi: 10.3399/bjgp18X695213] [Medline: 29472224]
20. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med 2022 Jan;28(1):31-38. [doi: 10.1038/s41591-021-01614-0] [Medline: 35058619]
21. Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. Curr Cardiol Rep 2014 Jan;16(1):441. [doi: 10.1007/s11886-013-0441-8] [Medline: 24338557]
22. Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. J Family Med Prim Care 2019 Jul;8(7):2328-2331. [doi: 10.4103/jfmpc.jfmpc_440_19] [Medline: 31463251]

XSL•FO
RenderX

23. Sahu M, Gupta R, Ambasta RK, Kumar P. Artificial intelligence and machine learning in precision medicine: a paradigm shift in big data analysis. Prog Mol Biol Transl Sci 2022;190(1):57-100. [doi: 10.1016/bs.pmbts.2022.03.002] [Medline: 36008002]
24. Hajjo R, Sabbah DA, Al Bawab AQ. Unlocking the potential of the human microbiome for identifying disease diagnostic biomarkers. Diagnostics (Basel) 2022 Jul 19;12(7):1742. [doi: 10.3390/diagnostics12071742] [Medline: 35885645]
25. Boscardin CK, Gin B, Golde PB, Hauer KE. ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity. Acad Med 2024 Jan 1;99(1):22-27. [doi: 10.1097/ACM.0000000000005439] [Medline: 37651677]
26. Hajjo R, Sabbah DA, Bardaweel SK, Tropsha A. Identification of tumor-specific MRI biomarkers using machine learning (ML). Diagnostics (Basel) 2021 Apr 21;11(5):742. [doi: 10.3390/diagnostics11050742] [Medline: 33919342]
27. Wenderott K, Krups J, Weigl M, Wooldridge AR. Facilitators and barriers to implementing AI in routine medical imaging: systematic review and qualitative analysis. J Med Internet Res 2025 Jul 21;27:e63649. [doi: 10.2196/63649] [Medline: 40690758]
28. Yang YM, Wang T, Chan HY, Huang YM. Key elements and theoretical foundations for the design and delivery of text messages to boost medication adherence in patients with diabetes, hypertension, and hyperlipidemia: scoping review. J Med Internet Res 2025 Jul 21;27:e71982. [doi: 10.2196/71982] [Medline: 40690759]
29. Lai CC, Chen CY, Chang TH. Predicting pathological complete response following neoadjuvant therapy in patients with breast cancer: development of machine learning-based prediction models in a retrospective study. JMIR Cancer 2025 Jul 18;11:e64685. [doi: 10.2196/64685] [Medline: 40680158]
30. Xu F, Yang Y, Luo Y, et al. Efficacy of Fuzheng Quxie formula against postoperative metastasis of lung cancer in stage IIA-IIIA with negative driver genes: protocol for a multicenter, double-blind, randomized controlled trial. JMIR Res Protoc 2025 Oct 24;14:e66342. [doi: 10.2196/66342] [Medline: 41135941]
31. Ye Z, Zhang B, Zhang K, et al. An assessment of ChatGPT's responses to frequently asked questions about cervical and breast cancer. BMC Womens Health 2024 Sep 2;24(1):482. [doi: 10.1186/s12905-024-03320-8] [Medline: 39223612]
32. A shocking chinese AI advancement called deepseek is sending US stocks plunging. CNN Business. URL: https://edition.cnn.com/2025/01/27/tech/deepseek-stocks-ai-china/index.html [accessed 2025-12-13]
33. Korea biomedical review. The DeepSeek dilemma: navigating innovation and security concerns in Korea's health care industry. URL: https://www.koreabiomed.com/news/articleView.html?idxno=26581 [accessed 2025-02-13]
34. FLESCH R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
35. Badarudeen S, Sabharwal S. Readability of patient education materials from the American Academy of Orthopaedic Surgeons and Pediatric Orthopaedic Society of North America web sites. J Bone Joint Surg Am 2008 Jan;90(1):199-204. [doi: 10.2106/JBJS.G.00347] [Medline: 18171975]
36. Textstat. Python Package Index. URL: https://pypi.org/project/textstat/ [accessed 2025-12-13]
37. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020 Mar;17(3):261-272. [doi: 10.1038/s41592-019-0686-2] [Medline: 32015543]
38. Johnson SB, King AJ, Warner EL, Aneja S, Kann BH, Bylund CL. Using ChatGPT to evaluate cancer myths and misconceptions: artificial intelligence and cancer information. JNCI Cancer Spectr 2023 Mar 1;7(2):pkad015. [doi: 10.1093/jncics/pkad015] [Medline: 36929393]
39. Haver HL, Ambinder EB, Bahl M, Oluyemi ET, Jeudy J, Yi PH. Appropriateness of breast cancer prevention and screening recommendations provided by ChatGPT. Radiology 2023 May;307(4):e230424. [doi: 10.1148/radiol.230424] [Medline: 37014239]
40. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965 Dec;52(3/4):591. [doi: 10.2307/2333709]
41. Arnastauskaitė J, Ruzgas T, Bražėnas M. An exhaustive power comparison of normality tests. Mathematics 2021;9(7):788. [doi: 10.3390/math9070788]
42. Ruxton GD. The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. Behav Ecol 2006 Jul 1;17(4):688-690. [doi: 10.1093/beheco/ark016]
43. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull 1971;76(5):378-382. [doi: 10.1037/h0031619]
44. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960 Apr;20(1):37-46. [doi: 10.1177/001316446002000104]
45. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. Nature New Biol 2020 Sep;585(7825):357-362. [doi: 10.1038/s41586-020-2649-2] [Medline: 32939066]
46. Scikit-learn: machine learning in python — scikit-learn 170 documentation. Scikit-learn. URL: https://scikit-learn.org/stable/ [accessed 2025-12-13]
47. Python data analysis library. pandas. URL: https://pandas.pydata.org/ [accessed 2025-12-13]
48. Vellido A. Societal issues concerning the application of artificial intelligence in medicine. Kidney Dis (Basel) 2019 Feb;5(1):11-17. [doi: 10.1159/000492428] [Medline: 30815459]
49. Kitsios F, Kamariotou M, Syngelakis AI, Talias MA. Recent advances of artificial intelligence in health care: a systematic literature review. Appl Sci (Basel) 2023;13(13):7479. [doi: 10.3390/app13137479]

XSL•FO
RenderX

50. Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. N Engl J Med 2023 Mar 30;388(13):1233-1239. [doi: 10.1056/NEJMsr2214184] [Medline: 36988602]
51. Hajjo R, Sabbah DA, Bardaweel SK, Zhong HA. Targeting the EGFR/RAS/RAF signaling pathway in anticancer research: a recent update on inhibitor design and clinical trials (2020-2023). Expert Opin Ther Pat 2024;34(1-2):51-69. [doi: 10.1080/13543776.2024.2327307] [Medline: 38450537]
52. Sabbah DA, Hajjo R, Bardaweel SK, Zhong HA. Targeting the PI3K/AKT signaling pathway in anticancer research: a recent update on inhibitor design and clinical trials (2020-2023). Expert Opin Ther Pat 2024 Mar;34(3):141-158. [doi: 10.1080/13543776.2024.2338100] [Medline: 38557273]
53. Pasas-Farmer S, Jain R. From discovery to delivery: governance of AI in the pharmaceutical industry. Green Analytical Chemistry 2025 Jun;13:100268. [doi: 10.1016/j.greeac.2025.100268]
54. Hajjo R. AI-models [github repository]. GitHub. URL: https://github.com/rhajjo/AI-Models [accessed 2026-01-04]

Abbreviations:

AI: artificial intelligence
FAQ: frequently asked question
FKGL: Flesch-Kincaid Grade Level
LLM: large language model

Edited by J Bender; submitted 19.Feb.2025; peer-reviewed by B Mesko, D Xu, J Ogunsakin; revised version received 27.Oct.2025; accepted 19.Nov.2025; published 27.Feb.2026.

Please cite as:
Hajjo R, Sabbah DA, Bardaweel SK
ChatGPT Versus DeepSeek for Breast Cancer Information Retrieval: Quantitative Comparative Study
JMIR Cancer 2026;12:e72839
URL: https://cancer.jmir.org/2026/1/e72839
doi:10.2196/72839

© Rima Hajjo, Dima A Sabbah, Sanaa K Bardaweel. Originally published in JMIR Cancer (https://cancer.jmir.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Planning and Developing a Symptom Diary Intervention for Breast Cancer Survivors With Concerns About Medication Brands (ENABLE Study): User-Centered Design Approach

Yolanda Eraso[1], PhD; Duncan Stewart[1], PhD; Ruth Edwards[2], PhD; Bal Nanray[3]; Jacqueline Cochrane[4], MBPsS

[1]School of Social Sciences and Professions, London Metropolitan University, London, England, United Kingdom

[2]Department of Pharmacy, University of Wolverhampton, Wolverhampton, United Kingdom

[3]Patient representative, Essex, United Kingdom

[4]The Open University, Milton Keynes, England, United Kingdom

Corresponding Author:
Yolanda Eraso, PhD
School of Social Sciences and Professions
London Metropolitan University
166-220 Holloway Road
London, England, N7 8DB
United Kingdom
Phone: 44 20 7133 2623
Email: y.eraso@londonmet.ac.uk

Abstract

Background: Approximately 80% of breast cancers are estrogen receptor positive, and following initial tumor treatment, patients are prescribed hormone therapy (HT) drugs (tamoxifen, letrozole, anastrozole, and exemestane) for 5-10 years. These drugs are known to cause several side effects. Additionally, a small number of studies have identified that changing medication brands (generics) can negatively affect patients' side effects, attitudes, and acceptance of HT. However, no effective intervention currently exists to address patients' concerns about generic switching.

Objective: This study explores whether a symptom diary appears to be an acceptable approach for patients and pharmacists to address brand concerns, and provides a detailed report of the planning and development stages of a coproduced diary intervention.

Methods: This paper presents the studies conducted during the planning and development stages of the symptom diary, following the person-based approach, which combines theory-, evidence-, and person-based approaches to ensure interventions are acceptable and easy to implement. The study included the following stages: (1) the planning stage, which involved analysis of an online patient forum (n=277), interviews with patients (n=9), and interviews with pharmacists (n=7). Data triangulation was used to present an integrated set of findings. (2) The development stage, which involved a rapid scoping review of the literature on the use of diaries among patients with cancer, coproduction workshops with patients and pharmacists (n=17), pharmacist interviews (n=11), and consultation with our patient advisory group (n=5). Analysis was informed by evidence synthesis, behavior change analyses, and iterative stakeholder consultations.

Results: In the intervention planning stage, patients and pharmacists perceived community pharmacies as the ideal setting for HT medication consultations. Both groups proposed using a patient diary to support awareness and attribution of drug-related symptoms. Pharmacists described patients' requests for specific brands as difficult to manage, whereas patients reported that switching brands undermined confidence in medication-taking. In the development stage, the scoping review (n=29 papers) identified barriers and facilitators related to diary use, and workshops explored diary format and content. These findings informed the intervention planning table and the first diary prototype. Pharmacist interviews and patient feedback informed the development of 2 further prototypes. The symptom diary is a self-monitoring tool designed to enhance self-efficacy in HT medication-taking behaviors. It includes recording side effects from brands alongside physical, psychological, and environmental factors; actions taken to manage symptoms; and problem-solving and planning before a medication consultation to facilitate personalized feedback from pharmacists.

Conclusions: We preliminarily developed an intervention that appears engaging, relevant, and acceptable for patients to self-monitor their symptoms. Further revisions informed by user feedback will be incorporated after testing ENABLE (Medication

XSL•FO
RenderX

Brand Changes in Hormone Therapy for Breast Cancer: A Community Pharmacy Intervention Development to Improve Patients' Adherence and Quality of Life) in community pharmacies.

Trial Registration: ISRCTN Registry ISRCTN15089342; http://www.isrctn.com/ISRCTN15089342

(JMIR Cancer 2026;12:e91234) doi:10.2196/91234

KEYWORDS

intervention development; behavioral sciences; breast cancer survivors; symptom diary; qualitative methods; community pharmacy

Introduction

Breast cancer (BC) is the most common cancer in the United Kingdom, with an estimated 56,400 new cases diagnosed in women each year [1]. Around 80% of BCs are estrogen receptor positive, meaning that, for the great majority of BC survivors, treatment does not end with surgery, radiation, or chemotherapy. Hormone therapy (HT) drugs (tamoxifen, letrozole, anastrozole, and exemestane) are prescribed for 5-10 years to reduce the risk of recurrence and mortality. These drugs are known to cause several side effects and have been consistently identified in the literature as a key factor in nonadherence [2] and poor quality of life (QoL) [3]. Adherence to HT declines by an average of 25.5% each year, and by the fifth year of treatment, adherence is estimated to range from 33.3% to 88.6%, while treatment continuation ranges from 45.2% to 87.4% [4].

In the United Kingdom, it is common practice for community pharmacies to dispense different generic brands of the same drug when patients collect their prescriptions, depending on availability and wholesalers. Currently, the British National Formulary lists 10 generic brands for tamoxifen (20 mg), 13 for letrozole, 17 for anastrozole, and 13 for exemestane, in addition to their proprietary names, when available [5]. Clinical studies on women's experiences with HT medications have revealed that the use of generic drugs, either to replace the proprietary name (generic substitution) or to switch between generic brands (generic switching), can negatively affect patients' side effects, attitudes, and acceptance of HT [6,7]. Among women switching tamoxifen generic brands, a small but significant proportion of patients in the United Kingdom (11%-13%) attributed severe menopausal symptoms to the brand switch, with symptoms resolving after taking a different brand. The authors suggested that patients' perceived changes in medication side effects were due to manufacturers' use of different excipient profiles, namely, the nonactive ingredients used to enhance manufacturing, patient acceptance, and bioavailability [6,7]. A US study on inactive ingredients concluded that this is a "poorly appreciated" area, considering the adverse reactions (allergies and irritants) that can occur in population groups taking more than 1 medication daily [8]. To date, only 1 study [9] on HT drugs has identified drug allergies as a strong predictor of adherence.

Additionally, experimental studies have shown that patients often misattribute symptoms arising from everyday activities, disease, or environmental factors as medication side effects [10], highlighting the need to assist patients in recognizing drug-related symptoms. Experiences of new side effects or worsening of existing ones after generic switching or substitution (whether excipient-related or belief-induced) can negatively affect the QoL and treatment continuation of patients with BC [6,7,11].

A growing body of evidence points to the need for interventions that provide enhanced support for patients with BC to better manage their medication [12-14]. The Royal Pharmaceutical Society acknowledges "considerable scope" for pharmacist-led interventions in cancer care, including medicines advice [15]. A systematic review on effective medicines use found that interventions that appear to improve medication adherence include medication self-monitoring and self-management programs, while promising interventions include medicines reviews, consultation with patients to resolve medication-related problems, development of a care plan, and provision of follow-up [16]. Similarly, other reviews have reported that patient medication adherence can be improved through specific community pharmacy approaches, including patient identification and enrollment, medication review, and patient assessment [17,18]. To date, no community pharmacy service or effective intervention exists to address patients' concerns regarding HT generic switching.

This paper reports on the planning and development of a self-monitoring tool, a symptom diary, as part of the ENABLE (Medication Brand Changes in Hormone Therapy for Breast Cancer: A Community Pharmacy Intervention Development to Improve Patients' Adherence and Quality of Life) study (NIHR206589 award; ISRCTN [International Standard Randomised Controlled Trial Number] registry ISRCTN15089342). ENABLE aims to coproduce, with patients with BC and pharmacists, an intervention to improve medication brand change (MBC) consultations in community pharmacy settings through the delivery of a patient-centered approach that could improve adherence and QoL.

The symptom diary is the patient-facing intervention component of the ENABLE study. The aims of this paper are 2-fold: (1) to explore whether a self-monitoring tool (symptom diary) appears to be an acceptable approach for patients and pharmacists to address MBC concerns; and (2) to provide a detailed account of the planning and development stages of a diary coproduced with patients, pharmacists, and stakeholders, as well as the empirical evidence and behavioral analysis that shaped its content. The latter will help explain whether a diary is an effective approach to improving consultations about HT generic brand switching, adherence, and QoL in BC survivors.

XSL•FO
RenderX

Methods

Study Design and Theoretical Framework

This was a multimethod qualitative intervention development study. It aligned with the person-based approach (PBA), which focuses on user-centered design to ensure that interventions are acceptable, engaging, feasible, and effective in their implementation [19]. Its theoretical approach was grounded in behavioral science and qualitative research methods to collect data on users' perspectives at different stages of the development process. Through its combination of evidence (users' needs and barriers and facilitators to intervention success), theory (behavioral science), and person-centered methods (qualitative research used to iteratively gather user feedback), the PBA is recognized by the Medical Research Council [20] as one of the approaches for developing complex interventions.

The intervention development methods for the symptom diary are described below in 2 main stages: planning and development. The Consolidated Criteria for Reporting Qualitative Research (COREQ) were used as the reporting guidelines for the study (Multimedia Appendix 1).

Intervention Planning

Planning Stage: Identification of User Needs and Behavioral Determinants

The planning stage used triangulation of qualitative data from 3 preliminary studies to identify users' needs, barriers and facilitators, and behavioral determinants: analysis of posts from a UK online BC forum, and interviews with patients and community pharmacists.

Online BC Patient Forum

Online patient forums are frequently accessed as valuable sources of health-related information, particularly for managing side effects, where patients can find supportive, experience-focused comments and advice [21]. The aim was to identify how patients with concerns about MBCs sought support for their medication, what actions they took, and what responses they received from health care professionals (HCPs). Searches were conducted in the "Hormone Therapy" section of the Breast Cancer Now forum (data accessed May 15, 2019-December 31, 2020) using the following keywords: "brand*," "tamoxifen," "letrozole," "anastrozole," "exemestane," and specific drug brand names in use at the time. Searches covered the period from January 1, 2013, to December 31, 2020. A total of 290 women were identified, of whom 277 (95.5%) experienced new or worsening side effects after brand switches. A research assistant extracted data into a Microsoft Excel spreadsheet, including brands, year of posts, patients' age, years of treatment, drug names, side effects, contacts with pharmacists, and interactions with other HCPs. Multiple posts by the same individual were combined into a single entry so that each individual had only 1 record.

Interviews With Patients With BC

A second study interviewed women with lived experience of MBCs recruited through the Breast Cancer Now charity. Participants were recruited through an invitation advertised on the charity's research web page. Interested patients completed an online form and were contacted by email. Eligible participants were aged 18 years or older, had previously been prescribed HT, and had experience with MBCs. Nine women were included in the study and interviewed by a research assistant between July 9 and September 19, 2021. Interviews were conducted via Microsoft Teams and Zoom (Zoom Communications, Inc; n=7 interviews) or by telephone (n=2 interviews) and lasted 24-63 minutes. Participants lived in different parts of England and had a mean age of 57 years. A semistructured interview explored patients' attitudes, needs, challenges, and preferences regarding HT brand changes and was informed by interpretative phenomenological analysis (IPA) [22].

Interviews With Community Pharmacists

Informed by the previous 2 studies, semistructured interviews with 7 community pharmacists located in London and the South East region were conducted by YE between March 16 and April 11, 2022. This exploratory analysis aimed to provide a more comprehensive understanding of the MBC phenomenon by triangulating our existing data with pharmacists' perspectives on the topic. An interview guide consisting of open-ended questions explored pharmacists' experiences with MBCs, management of consultations, the potential role of pharmacists, and the feasibility of a contract to deliver a pharmacy service for patients with BC. Saturation was sought when patterns across all 4 topics were identified. Participants were purposively selected. Eligible participants were registered pharmacists or pharmacy technicians with experience dispensing HT medication. Participants were recruited through visits to pharmacies in London and Oxford, where participant information sheets (PISs) containing a link to register interest were distributed, and through snowball sampling, whereby participants shared the PISs with contacts in the region. One pharmacist agreed to participate but did not respond to attempts to arrange the interview. No participants were known to the researchers before the interviews. Individual interviews were conducted face to face (n=3), by telephone (n=2), and via Microsoft Teams with audiovisual recording (n=2), and lasted 15-25 minutes. Participant characteristics are presented in Table S1 in Multimedia Appendix 2.

Analysis

Content analysis [23] was used to analyze women's posts in the online patient forum. First, the research assistant inductively coded the posts, and the coding was audited by YE. Second, the codes were categorized according to behavioral factors using the Theoretical Domains Framework [24].

All interviews were audio recorded, transcribed verbatim, and analyzed. A patient interview guide was developed using an IPA approach [22], primarily because MBCs related to HT had not previously been qualitatively explored, and an experiential approach was sought to generate preliminary in-depth data to inform future interventions based on a PBA [19]. Further details of the study are published elsewhere [11]. Pharmacists' interviews were analyzed using thematic analysis [25], in which codes were derived inductively by an experienced research assistant and audited by YE (25% of transcripts). The

XSL•FO
RenderX

identification of themes was mapped to the study aims and discussed and agreed on by the researchers.

Finally, data triangulation across the 3 data sources was undertaken to support cross-validation and enhance the credibility of the MBC phenomenon as experienced by pharmacists and patients [26]. Findings were organized into barriers and facilitators and mapped onto the Theoretical Domains Framework [24] to analyze behavioral determinants. The synthesis is presented with an assessment of agreements and disagreements across the 3 data sets. In addition, a detailed behavioral analysis was conducted to document how the qualitative studies undertaken during the planning stage informed the development of the symptom diary intervention. Although the rich idiographic approach of the IPA study is not fully retained in this paper, data triangulation allowed us to strengthen the validity of the selected empirical evidence used in the behavioral analysis.

Intervention Development

Overview

Following the planning stage, our team embarked on an intervention development process guided by the PBA [19]. A description of the methods used at each step is summarized below, following the Guidance for the Reporting of Intervention Development (GUIDED) [27] checklist for reporting intervention development (Multimedia Appendix 3).

Theory

The intervention followed Bandura's social cognitive theory [28], which addresses the interplay between self-regulatory (personal) and environmental (social) determinants of health behavior. Previous systematic reviews have identified self-efficacy in medication taking as a potential predictor and modifiable factor influencing adherence and persistence with HT drugs [29,30]. Self-efficacy is a core concept in Bandura's social cognitive theory [31] and refers to an individual's belief or confidence in their ability to perform a specific behavior to achieve a desired outcome. Higher self-efficacy for medication taking has been significantly associated with greater HT adherence [32]. Evidence also suggests that self-efficacy is an important determinant of intentional nonadherence related to 2 other distinct behavioral components: self-efficacy in managing side effects [29] and self-efficacy in communicating with HCPs [32-34]. The latter is particularly relevant because concerns regarding MBCs increase patient-HCP communication, mainly with pharmacists and general physicians (GPs), as shown in our analysis of the online forum data below.

Self-efficacy is a precursor to self-management. In Bandura's theory, cognitive and emotional factors (participant modeling and cognitive processing), alongside environmental factors such as verbal persuasion (social support), interact to influence self-management behaviors. Self-management in chronic conditions refers to the ability to monitor the condition (eg, side effects) and to achieve the motivation and emotional responses necessary to maintain an acceptable QoL [35]. While self-monitoring strategies, such as a symptom diary, are instrumental in increasing patients' awareness and improving symptom management, evidence indicates that addressing a single factor to improve adherence is ineffective because of the multifaceted nature of adherence dimensions [36]. A meta-analysis on HT adherence found that only interventions involving bidirectional communication (ie, eliciting information from and providing information to patients) improved adherence compared with controls [37]. Thus, environmental factors such as social support, in the form of feedback on behaviors, represent an important source of self-efficacy for successfully managing medication-taking behaviors.

In addition, self-management programs for patients with BC are often not co-designed with patients and HCPs, which may reduce their effectiveness. Relatedly, in the United Kingdom, nonadherence to HT among ethnically minoritized patient groups is 1.48 times higher than among White British patients [38]. Patients with cancer living in deprived areas are more likely to experience delays in starting treatment (33% higher compared with those living in the least deprived areas) or delays in seeking help because they feel they would not be taken seriously [39]. These findings highlight the need for inclusive co-design approaches, as reflected in our recruitment strategies.

Rapid Scoping Review

Rapid scoping reviews are particularly relevant for intervention development studies because of their exploratory nature and broader scope in determining what evidence is available across heterogeneous sources [40]. For this study, the review question was: "What patient diary interventions are available to support self-monitoring behaviors in cancer survivors after active treatment?" Additional elements of interest included intervention components and data on patients' experiences, usability, acceptability, adherence, and QoL outcomes. Studies were excluded if diaries were used solely to report symptom scales, episode frequencies (eg, pain or hot flashes), or activities performed, without promoting further user engagement (ie, not intended to facilitate patient self-management behaviors or feedback from HCPs).

Searches included systematic reviews and primary studies of any design in the following electronic databases: Cochrane Library, PubMed, CINAHL, and APA PsycINFO. Because of time constraints, searches were limited to August 1, 2004, to August 31, 2024. The search terms used for each database were adapted to derive the most meaningful results, using a combination of free-text terms, MeSH (Medical Subject Headings) terms, and subject headings. Additional key papers were identified through reference lists, targeted author searches, forward citation searching, and the authors' expertise.

The scoping review protocol has been published [41], and the review was conducted by JC and YE. The findings were used to inform workshop discussions on the most appropriate information for patients to record about symptoms, approaches to enhance self-efficacy, and problem-solving strategies to support information exchange during medication consultations.

Coproduction Workshops

Coproduction methods [42] informed the workshops, with the aim of developing a symptom diary in partnership with patients with BC and pharmacists. Recruitment of patients with BC took place in London, and Breast Cancer Voices and South Asian

XSL•FO
RenderX

Health Action circulated the call for contributors. Women with lived experience of MBCs and diverse sociodemographic characteristics (age, ethnicity, language, and socioeconomic status) were selected. Community pharmacy participants were recruited through Local Pharmaceutical Committee networks via RNDD North London (Community Pharmacy Research Champions) and included pharmacists, pharmacy technicians, and locums with experience dispensing HT medication. Interested individuals were asked to register via Microsoft Forms.

We planned for a workshop sample of 20 participants, equally divided between patients and pharmacists and organized into 2 mixed groups. Although there is no established guidance on workshop participant numbers [43], we considered that smaller mixed groups would better facilitate effective communication and engagement across different perspectives (patients and pharmacists). Workshop attendees included 12 patients with BC and 5 community pharmacists. A diverse group of patients with BC was purposively selected during recruitment and included participants of different ages (range 38-73 years; mean 56 years), ethnicities (8 White and 4 from minoritized ethnic groups), residential backgrounds from the most deprived areas (n=8) and least deprived areas (n=4), and 1 participant with a physical impairment. Community pharmacists' roles included locum pharmacist (n=1), superintendent pharmacist (n=1), pharmacy managers (n=2), and pharmacist (n=1), with a range of work experience (1-35 years; mean 20.6 years), serving communities in the most deprived areas (n=3) and least deprived areas (n=2). Attendee characteristics are presented in Table S2 in Multimedia Appendix 2. A summary of findings from the rapid scoping review was circulated to participants in advance.

A face-to-face workshop was held at London Metropolitan University in October 2024 and lasted 4 hours, including lunch and coffee breaks. The meeting began with a presentation of the ENABLE study, an introduction to the National Institute for Health and Care Research key principles of coproduction [44], the role of patients and stakeholders as research contributors, and the planned activities. In all workshops, pharmacists and patients worked together in small groups with 1 facilitator (1 Breast Cancer Now nurse, or 1 patient representative, or members of the research team [DS, RE, or YE]) to encourage understanding of each other's perspectives, experiences, and challenges. This shaped group discussions toward a more feasible and practice-focused approach to symptom diary implementation. A handbook guide was developed for each facilitator, including instructions on the aims, questions, probes, and group management. A voting system was used to complement discussions for some activities, such as a diary format and content. The session concluded with a whole-group discussion and feedback on the workshop. The workshop was audio-recorded. Additionally, 3 online workshops were held on October 25 and 30 and November 5, 2024, via Microsoft Teams with audiovisual recording, following the same structure described above. These workshops included screen sharing of presentations, tasks, and materials, with voting activities completed orally. Workshops lasted 1.5-2 hours.

Intervention Guiding Principles, Logic Model, and Intervention Planning Table

Based on the data triangulation conducted during the planning stage, theory-informed evidence on HT adherence interventions, the rapid scoping review, and workshop discussions, we developed the intervention guiding principles and logic model, as well as the intervention planning table (Table 3). The latter involved a behavioral analysis that first identified the mechanisms of action (MoAs) and subsequently the behavior change techniques (BCTs), based on the best available evidence reported in the Theory and Technique Tool [45], a resource for the development of theory-based interventions.

Patient Advisory Group

The group included 5 patients with BC with experience of MBCs from different ethnic backgrounds (South Asian, n=1; Indian, n=1; and White British, n=3), as well as BN, the patient and public involvement and engagement (PPIE) lead who chaired the group. As in previous recruitment strategies, participant selection aimed to ensure a plurality of lived experiences and demographic characteristics (see Table S3 in Multimedia Appendix 2). The initial patient advisory group (PAG) meetings were scheduled to provide feedback on the intervention components without prior involvement with the research team, in order to maximize members' perspectives on the intervention from a fresh viewpoint. The PAG members are based in England, and meetings take place online. Breast Cancer Now and South Asian Health Action supported recruitment. Members participate in a series of scheduled meetings throughout the study and receive payment for their time in line with National Institute for Health and Care Research recommendations. The PAG reviewed diary prototype 1 and provided suggestions for prototype 2.

Interviews With Pharmacists

A purposive sample of community pharmacists (n=6) and pharmacists representing professional bodies (n=5), including the Royal Pharmaceutical Society, Community Pharmacy England, and National Pharmacy Association, were interviewed (see Table S3 in Multimedia Appendix 2). The PBA recommends a flexible approach to sample size during intervention design (5-10 participants, depending on study scope) [46]. The aim was to gather feedback from target users and key stakeholders on their views and practice-based experiences of the intervention design, offering new insights to complement those of the PAG, whose members had been involved in the study from the outset. Community pharmacists were recruited using the same criteria and through the organizations described above, while representatives from professional bodies were recruited through snowball sampling via a contact provided by RE. No participants were known to the researchers before the interviews, and there were no dropouts. All participants received a PIS explaining the purpose of the study. The semistructured individual interviews used a topic guide to explore views on the intervention and insights into barriers and facilitators to implementation. A synopsis of the intervention was circulated in advance, and pharmacists provided suggestions for diary prototype 2. Interviews were conducted by YE via Microsoft Teams with audiovisual

XSL•FO
RenderX

recording and lasted 17-46 minutes. Interviews took place between November 28 and December 19, 2024.

Research Team and Reflexivity

Development Team

The core development team included those with expertise in community pharmacy and pharmacy education, applied health research, adherence behaviors in BC, patient involvement in research, qualitative methods, and the lived experience of BC. JC (health psychologist) has conducted research on postmastectomy lived experiences using qualitative approaches. RE (pharmacist and academic in education) and BN (patient representative) had experience with MBCs related to HT. DS (applied health researcher) has extensive experience researching pharmacy medication review services. YE (health studies) has expertise in qualitative research, with research interests in women's health, including BC adherence, behavioral science, and intervention development. Stakeholder input included a Breast Cancer Now nurse (collaborative partner), who supports patients through the charity's helpline and facilitated workshops, as well as reviewing diary revisions with the team. Research team meetings took place online.

Reflexivity

The research team, which represented a range of academic backgrounds and differing experiences and interests related to the research topic, agreed on the final analyses and interpretations presented here. Previously, we engaged in a coproduction process that addressed power imbalances through respectful and equitable relationships, while collaboratively challenging assumptions and generating and interpreting data. The iterative feedback process involving embedded groups (PAG) and external participants (pharmacists) engaged the research team in ongoing reflection on previous assumptions underlying the use of a diary.

Development of the Symptom Diary

We first produced a prototype of the diary, drawing on findings from the planning stage, rapid scoping review, and codevelopment workshops, which informed its format and content. The intervention planning table, alongside the intervention guiding principles and logic model, provided a framework for designing and mapping diary sections and items relevant to patient monitoring according to their potential to effect behavior change (ie, self-efficacy). Two further prototypes (2 and 3) incorporated modifications following feedback from the PAG, pharmacists, and the research team.

Analysis

Workshops, interviews, and meetings were recorded and transcribed verbatim by JC. Workshop materials were also collected for data analysis. Interviews with pharmacists followed a thematic analysis approach [25], including independent inductive coding by JC and YE. Data collection initially focused on capturing all participants' preferences and suggestions, followed by consideration of disagreements and consensus based on the number of contributors supporting each view. Final decisions regarding diary design and proposed modifications were based on their likely impact on facilitating behavior change and their alignment with the study protocol (eg, suggestions to use a mobile diary were documented but were not planned within the scope of this study). Modifications were recorded in a table of changes.

Ethical Considerations

Planning Stage

The studies received ethical approval from the London Metropolitan University Research Ethics Committee (approval number SSSP-4050119/SSPR 022). Written informed consent was obtained from all interviewed participants. Pharmacists received a £20 (US $27) voucher as a token of appreciation for their time. For the patient forum posts, individual consent was not obtained because of the practical challenges associated with obtaining authorization from a large online community. Authorization was granted by Breast Cancer Now forum moderators, and, following recommendations from the London Metropolitan University Research Ethics Committee, posts were paraphrased to minimize the risk of traceability through online searches.

Development Stage

Ethical approval was obtained from the Health Research Authority NHS (National Health Service) West Midlands-South Birmingham Research Ethics Committee (approval number 24/WM/0137) and the London Metropolitan University Research Ethics Committee (approval number SSSP-2024-025).

Workshops

All participants signed a consent form and received £75 (US $101) for their involvement, in line with National Institute for Health and Care Research recommendations, as well as a certificate of attendance. Pharmacists were also offered a fee for locum backfill.

Pharmacist Interviews

All participants provided electronic consent before the interview and were offered a £30 (US $40) voucher as a token of appreciation for their time.

Results

Intervention Planning

Barriers and facilitators experienced by patients and pharmacists across the 3 studies (data triangulation) are presented in Table S1 in Multimedia Appendix 4, alongside the behavioral determinants identified using the Theoretical Domains Framework [24].

Synthesis of Three Studies

Beyond Patient Factors: The Role of Health Care Systems

The findings provide insight into "health care team and system factors," including HCP interactions, drug distribution systems, drug availability (shortages), HCP beliefs about generics and side effects, and responsiveness to patients' concerns. Health care team and system factors represent one of the multifaceted domains that the World Health Organization associates with

XSL•FO
RenderX

medication nonadherence in chronic conditions [47]. Other domains include social and economic factors affecting patients, disease-related factors (disease progression and risk), therapy-related factors (treatment burden), and patient-related factors (self-efficacy and beliefs about medication) [47]. Overall, research on patients' behaviors and psychological needs has received more scholarly attention than system factors or environmental context and resources, despite their potential to influence patients' capability and motivation to adhere to treatment.

Convergences

All data sets provided evidence that some women are brand sensitive and that these experiences, alongside the search for a single or alternative brand(s), often led to fortuitous solutions depending on pharmacy stock availability, pharmacists' efforts to procure a specific brand, or GP agreement to prescribe by manufacturer name. Community pharmacists interviewed appeared responsive to patients' brand-seeking behaviors and attempted to secure medications within the constraints described above (see Table S2 in Multimedia Appendix 4). Most pharmacists (with one exception) believed that different brands could cause different side effects, and their explanations reflected both a biomedical model (different excipients and E numbers) and a belief in patients' lived experiences. Although there is no scientific evidence explaining how HT brands may cause different side effects, the explanations proposed by pharmacists, particularly the possibility that excipients may affect bioavailability, were shared by patients, who often scrutinized nonactive ingredients in search of validation for their perceived differences. Patients and pharmacists viewed community pharmacies as the ideal setting for HT medication consultations. Both groups recognized the need for pharmacist training on HT drugs and patients' medication concerns. Pharmacists also anticipated the need for clear guidelines to manage these conversations, similar to the New Medicines Service [48], to ensure consistent advice for patients. Both groups proposed the use of a patient diary to support greater awareness of side effects, drug attribution, and beliefs about medications.

Dissonances

Pharmacists described patients' requests for specific brands as challenging (time-consuming and, at times, costly for the pharmacy) and difficult to manage within the constraints of National Health Service prescription services and the Drug Tariff [49]. Although patients considered cost savings to the NHS as the reason underlying generic switching, they often did not understand the NHS prescribing framework or the financial implications for pharmacies operating as businesses. Some patients in the online forum described the process as "Every prescription pick-up is an anxiety-inducing moment" [P32-PaF], particularly when they sought only 1 specific brand, whereas pharmacists viewed this behavior as frustrating and, at times, unnecessarily self-inflicted:

> *So at least some people, they're willing to try something if they can't get it, but I'm like, why would you want to restrict yourself?* [CP5]

This was compounded by patients who claimed to follow their HCP's advice to stick to 1 brand,

> *So when I mentioned a side effect to the oncologist at my six-monthly review, he said: "Always stick to the same brand"...I would say [to the pharmacist]: "Can I have the AR* brand name please?" and they would say: "Well, we're not allowed to do that". And to be honest, I've been backwards and forwards so much about trying to check [this brand] that I just gave up. [Julie, interview quoted in* [11] p.10]

This highlights, first, that brand sensitivity is not uniquely triggered by patients and, second, a disconnect within the care team regarding how the medication supply framework operates. Patients reported that concerns about MBCs undermined their confidence in taking medication and could affect treatment continuation:

> *I've got one friend in the group..., who just stopped it altogether because she couldn't find a brand that was supportive to her, so she stopped. But again, she was like me, she couldn't get a discussion going about what the differences might be. [Diane, interview quoted in* [11], p.9-10]

We aimed to capture women's lived experiences in our behavioral analysis (see Table S3 in Multimedia Appendix 4), scrutinizing the identified barriers, negative emotions, lack of knowledge, limited confidence in managing symptoms, and feelings of being undermined (not listened to) during conversations with HCPs, who often questioned the validity of their experiences. The latter was carefully considered in our analysis of self-monitoring behaviors (diary use) proposed by women themselves, in a way that balanced their experiences with MBCs and those of pharmacists, for whom the need for specific communication skills and guidance in managing these conversations was acknowledged. Self-monitoring through digital technologies is widely used in the prevention and treatment of chronic conditions, but it has also been associated with "individual responsibility," shifting the burden of care onto patients [50]. Overall, women made it clear that support after discharge from the hospital was needed and that HCPs should listen to their concerns and offer better options for managing medication-related issues [11]. The research team envisaged pharmacy training and guidance as requiring alignment between existing medication consultation skills and the specificities of HT treatment in BC survivors. Further details on the online forum analysis and the patient and pharmacist interviews are available in Tables S1-S3 in Multimedia Appendix 4.

Taken together, these studies suggest that many patient interactions with community pharmacists could be construed as missed opportunities to provide care and advice for symptom management beyond a narrow focus on brand-related discussions. As it stands, system factors (wholesaler stock, medication shortages, and GP prescribing by brand name) and scientific principles (generic bioequivalence) appear to place self-management and side effect control beyond patients' capabilities and limit the support pharmacists can provide. In summary, the 3 intervention components identified during the planning stage were informed by the preferences of patients and

XSL•FO
RenderX

pharmacists and included a self-monitoring diary, a pharmacist training package, and consultation guidelines. The next section reports findings from the diary development process.

Development of a Symptom Diary

Rapid Scoping Review

A total of 29 reports (including 19 studies) published between 2007 and 2024 were identified, of which 8 were from the United Kingdom. Study designs included protocols (n=4); intervention design studies using mixed methods (n=3) and qualitative methods (n=3); feasibility studies (n=4); pilot studies (n=5); randomized controlled trials (n=3); postintervention studies using qualitative (n=1), quantitative (n=1), and mixed methods approaches (n=1); systematic reviews (n=2); an overview of reviews (n=1); and a Delphi study (n=1). Most were intervention studies (n=24) involving different cancer types (eg, uterine/cervical, breast, colon, and prostate and lung cancers, among others). Diaries were used to improve physical activity and healthy diet (n=18), QoL (n=2), symptom management (n=3), and medication adherence (n=2). Diary use required completion periods ranging from 3 days to 12 months, and the format was primarily digital (web- or app-based), paper only (n=3), or both paper and digital (n=1). Most articles used diaries to record activities, including the use of scales, whereas 2 studies used a diary more as a reflective tool [51,52]. Key diary components and outcomes were organized into potential barriers and facilitators (Table 1). Although evidence regarding the effectiveness of diary interventions remains limited (only 3 randomized controlled trials [52-54]), positive outcomes were reported across heterogeneous interventions (eg, increased physical activity, improved diet, or enhanced medication adherence) using behavioral targets relevant to supporting self-efficacy in our intervention (goal setting, problem solving, and social support). Evidence regarding factors that appeared to support or hinder engagement, usability, and acceptability of diary use among patients with cancer was also considered during workshop discussions.

Table 1. Summary of topics, facilitators, and barriers from the rapid scoping review of the literature.

Facilitator	Barriers
Patient enrollment/completion	
• High intervention/study completion [52,53,55-61]	• Team's inability to reimburse lost wages or offer interactions with a health care professional outside working hours [52]. • Intervention is inconvenient to everyday life/transport [58]. • Participants' unwillingness to self-manage lifestyle [52,62,63]. • Anxiety about participation [55]. • Ill health restricting participation [51,53,55,63-65].
Engagement, acceptability, and usability of the diary	
• Diaries seen as useful [54-58,61,65-69] • Easy to learn and use [55,56,58,61,65,66,69,70] • Cancer survivors appreciated knowledge gained about intervention target (eg, medication adherence/diet/physical activity) [51,54,58,66,69-71]	• Burden, difficulty, and time commitment required for diary keeping [55,56,63]. • Memory issues and fatigue [56]. • Technical issues with digital diaries [55,56,61,63,66,70]. • Literacy and language barriers for paper-based and electronic diaries [51,52]. • Diary lacked complexity or flexibility [72].
Interventions with a diary component and behavioral targets	
• Focus on goal setting [51,52,55,57-59,61,63,64,67,69-71,73-75] • Problem solving [51,52,54,57,58,64,67,69,71,73-76] • Social support: discussion with a health care professional and feedback on plans/goals/activities [51-55,57-60,62-64,67,71,73-78]	• No personalized information in automated feedback [69]. • Sustained lifestyle change is hard [61]. • Diary lacked interactivity [70].
Outcomes of interventions using a diary component (self-monitoring)	
• Increased self-management/control [53-55,66,69,70,72] • Intervention aided in symptom control, supporting participants to become aware of (medications and symptoms, food quality, exercise) and plan daily routines accordingly [52-55,66,69] • Improved communication with health care professionals [52,54,55,63,66,76,79] • Improved intervention outcomes [52-54,60,64,65] • Increased confidence [55,60,69]	• A few patients found recording side effects as drawing too much attention to them [66]. • Nurses found it challenging to no longer inform and advise patients in general. • They were expected to facilitate patients' own reflection [76].

Co-Development Workshops

Discussions were guided toward addressing key diary components based on patients' and pharmacists' needs and underlying behaviors identified during the planning stage, theoretical approaches related to self-efficacy in medication-taking behaviors, and findings from the scoping review identifying barriers and facilitators to diary engagement and relevant behavioral targets. The team outlined 2 broad

XSL•FO
RenderX

sections for workshop discussions: (1) diary content and (2) preparation of key information for a medication consultation with a pharmacist. The discussion topics and aims are listed in Table 2.

Table 2. Discussion topics and their aims.

Topic for discussion	Aim
• Usefulness of the diary	To explore what could be useful to record
• Diary format: frequency, medium; free text and tick boxes, scales, emojis; length; and reminders	To identify components that will ensure engagement and usability
• Information (hormone therapy drug common side effects and when to seek help) • Symptom recording (physical and emotional) • Personal circumstances and external factors (environmental and other medications) • Problem solving (actions taken to control symptoms)	To gather participants' preferences about the diary's main content
• Content summary (case scenario) • Problem solving (graph and table with summary of key points)	To organize and identify key information written in the diary to use in the discussion with the pharmacist

Participants' illustrative quotes and votes for the diary format are available in Tables S1 and S2 in Multimedia Appendix 5. At this stage, we drafted the intervention guiding principles and the logic model (see Tables S3 and S4 in Multimedia Appendix 5), which provided a framework to design and map diary sections and items for the intervention content.

Intervention Planning Table

The analysis reflects the barriers identified among patients with BC with concerns about MBCs, factors influencing diary engagement among patients with cancer (informed by existing evidence), and the diary components and content most likely to address these issues. The behavioral analysis identified the MoAs and BCTs most likely to be effective in facilitating behavior change, based on the best available evidence reported in the Theory and Technique Tool [45]. Table 3 presents the intervention planning table.

XSL•FO
RenderX

Table 3. Intervention planning table.

Barriers to target behavior	Intervention components	Mechanisms of action	BCT[a]	Content
Lack of knowledge on HT[b] medication and side effects (PS[c], T[d], WS[e], and SR[f])	• Information about potential side effects of HT drugs—supports women in identifying side effects.	• Knowledge, beliefs about consequences, attitude toward the behavior, and skill (self-efficacy in deciding when to seek help)	• Information about health consequences (BCT 5.1), credible source (BCT 9.1), and instruction on how to perform the behavior (BCT 4.1)	• A summary list of 4 HT drugs (tamoxifen, letrozole, anastrozole, and exemestane) with common, occasional, and rare side effects provided by Cancer Research UK—Advice on when to seek professional help.
Low confidence in managing side effects from the new brand (PS, T, WS, and SR)	a. Support women to self-monitor their HT medication. Keeping a record of side effects from different brands: physical, psychological, and environmental factors. b. Record actions taken to control symptoms c. Raising awareness about side effects from brands (avoidance of drug misattribution)	a. Goals, behavioral regulation, and behavioral cuing b. Behavioral regulation c. Attitudes toward the behavior	a. Goal setting (behavior) (BCT 1.1), action planning (BCT 1.4), and self-monitoring of the behavior (BCT 2.3) b. Problem solving (BCT 1.2) c. Framing/reframing (BCT 13.2)	a. The "My Symptom Diary" section has a set of questions about brand, date of new symptoms, physical/psychological (location, severity, and impact [activities, sleep, emotions]), and environmental factors (other medications, personal circumstances, and external factors). b. A question prompts women to record anything done to control symptoms, and if it worked. c. Repeat completion, revision, and reflection on responses.
Experience disbelief about symptoms from the new brand; disregard about concerns (PS and WS)	• Key summary of the diary and plans ahead of medication consultation.	• Goals, motivation, beliefs about capabilities, and behavioral regulation	• Goal setting (outcome) (BCT 1.3), problem solving (BCT 1.2), and reduce negative emotions (BCT 11.2)	• A dedicated section in preparation for the medication consultation—summary (symptoms, environmental factors, actions to control, and medication expectations)
Lack of meaningful engagement and support from professionals regarding HT side effects (PS, T, WS, and SR)	• Personalized pharmacist feedback informed by key diary summary points during a planned medication consultation (person-centered approach incorporating shared decision-making).	• Goals, motivation, feedback process, environmental context and resources, and social influences	• Review outcome goals (BCT 1.7), feedback on behavior (BCT 2.2), and social support (practical) (BCT 3.2)	• Medication consultation starts with key summary points (diary). Medication guidelines for pharmacists (6 steps). Pharmacists and patients discuss options until a mutually agreed new approach to deal with symptoms seems acceptable. Alternatives/resolution for no decision. Notes taking, follow-up consultation.

XSL•FO
RenderX

Barriers to target behavior	Intervention components	Mechanisms of action	BCT[a]	Content
Lack of acceptability and engagement with a diary (SR and WS): cumbersome tasks, burden (time-consuming/long), health literacy, nonnative speaker, and medium preference (paper and online)	a. Clear introduction on how to complete the diary b. Allow flexibility for diary completion (minimum 1 month to maximum 3 months) c. Lay language and easy to understand d. Discrete space for responses to questions with the option to expand e. Option to request from the pharmacist a paper copy or an electronic version f. Non-English speakers	a. Skills and belief about capabilities b. Behavioral cueing c. Environmental context and resources d. Emotion e. Environmental context and resources f. Environmental context and resources	a. Instruction on how to perform behavior (BCT 4.1) b. Action planning (BCT 1.4) and habit formation (BCT 8.3) c. Social support (practical) (BCT 3.2) d. Reduce negative emotions (BCT 11.2) e. Social support (practical) (BCT 3.2) f. Social support (practical) (BCT 3.2)	a. Simplicity of instructions b. Format is easy to navigate. Recommend writing 3 times a week. c. Key points: changes in brand and new/worsening of side effects; and simple, plain language, provided by the patient advisory group. d. The layout for responses allows for expansion (electronic version) or adding pages (paper version). e. Electronic version available in a free word processor; also offline to ensure accessibility. f. Offer translated versions of consent forms and diary in languages other than English.

[a]BCT: behavior change technique.

[b]HT: hormone therapy.

[c]PS: planning stage.

[d]T: theoretical approaches.

[e]WS: workshop.

[f]SR: scoping review.

Barriers to the target behavior identified through qualitative research informed the behavioral analysis. Some barriers, such as "experience disbelief about symptoms," were specifically derived from data collected during the planning stage (see Table S3 in Multimedia Appendix 4) and further expanded during the development stage through workshop discussions (see Table S1 in Multimedia Appendix 5). During the workshops, participants introduced a new perspective of the diary as a form of "evidence" that could support patients' reports of experiencing different side effects:

> *it might make him [pharmacist] think well, actually, you know, this does matter to her, it is important, it can happen.* [PA10]

There was agreement regarding the types of issues to record, except for "environmental factors/life circumstances," for which concerns about privacy were raised by participants from minoritized ethnic groups (patients and pharmacists). This consideration was incorporated into the diary instructions as an optional component. Regarding the "key summary of the diary" prepared for the consultation (problem solving), 1 patient found documenting actions taken to control symptoms problematic because "it dilutes the fact that it's the brand that is causing this specific problem" [PA8]. These points will be important to examine in future research, as they suggest that the extent of diary completion may differ among patients who feel more certain that side effects are brand-related, which may or may not influence reflective processes before the consultation. Of particular importance were patients' discussions about the need for pharmacist feedback to balance the burden of completing a diary, as expressed by 1 patient: "that means that you can use it [diary] to empower yourself beyond your own self patience" [PA5]. A pharmacist similarly stated, "It'll help them [patient] and the caring team make better decisions, definitely" [CP-1]. Social support was also identified in the scoping review [66] and theoretical approaches [37]. Finally, issues related to the acceptability of and engagement with the diary were informed by findings from the scoping review and workshops, during which participants voted on the main diary features (see Table S2 in Multimedia Appendix 5).

Once the barriers to MBCs had been identified through the planning stage and workshops, they were mapped onto the MoAs, a range of behavioral theoretical constructs, including those within the Theoretical Domains Framework, that represent the processes through which specific BCTs influence behavior. The team agreed to select BCTs for which evidence of a link to a specific MoA was available in the Theory and Techniques Tool [45].

Symptom Diary

Prototype 1 of the symptom diary consisted of 5 sections: (1) an introduction explaining the purpose of the diary and presenting 2 scales (pain and mood) for later reference; (2) "My Symptom Diary," containing 8 questions; (3) "Preparing Your Medication Consultation With the Community Pharmacist," designed to summarize key points; (4) feedback on the experience of using the diary; and (5) an appendix listing side

XSL•FO
RenderX

effects for each of the 4 HT drugs, categorized as common, occasional, or rare.

Details of subsequent changes to this prototype version are discussed below.

PAG and Pharmacist Interviews: Prototype 2

The PAG members were satisfied with the design and purpose of the diary because it allowed patients' different experiences and perspectives on what mattered to them to be captured. They also praised the prompts encouraging consideration of environmental factors, including personal circumstances, as a valuable approach for eliciting a broad range of factors to consider.

> *People would perhaps think about environmental factors, so I think it's good to have all the different ones as a prompt, but it's not compulsory that they fill it in, if it doesn't apply to them. So, I thought that was good.* [PG1]

Regarding the diary content, members proposed the inclusion of 2 additional items to record: the date on which the medication was changed and whether the symptom had resolved. This was considered important given the possibility that medications could change while using the diary and that symptoms could fluctuate or be successfully managed by women. Members also questioned the inclusion of "fatigue" within the mood scale, as they perceived it to be a physical sensation rather than a mood-related state. We agreed to remove fatigue from the scale because the primary purpose of the diary was to facilitate women's reflection rather than serve as a tool for quantitative data collection. Another point raised by the PAG concerned the summary of key points before the medication consultation. Patients suggested replacing "intentions" with "expectations," given that a patient's goals may not necessarily be fulfilled by the pharmacist, whereas "expectations" would empower women to express what they felt they needed and hoped to receive. In addition, participants suggested signposting patients to charity support before the medication consultation. Finally, 1 PAG member reviewed the entire diary for lay comprehension and made additional suggestions, which were incorporated alongside those described above.

All pharmacists were supportive of the use of symptom diaries, and many were familiar with their use for other medicines. Community pharmacists praised the diary for facilitating pharmacists' support and increasing patient awareness.

> *The benefits are huge....I think it would allow patients who are not medically minded to categorise and really work out how bad it is* [CP2-5]

No specific comments were made regarding the diary's set questions other than its general design: "I think make it really simple for patients" [CP2-2]. Some pharmacists interpreted the diary as giving pharmacists a more proactive role (ie, linking side effects to factors recorded for pattern observation in specific brands).

> *So having a symptom diary for, not just in this particular therapeutic area, but in lots of therapeutic areas is useful, to see both the symptoms and side effects and you can track and identify where those are linked to.* [PH2-3]

It was explained that the initial expectation was that pharmacists would not read the diary because of time constraints and privacy concerns, and that they would instead advise patients based on the patients' presentation of key points from the diary (medication expectations valued by patients). Medicine shortages were mentioned as a challenging issue for HT drugs, echoing previous qualitative evidence. There were also comments regarding the optimal time for introducing the diary to patients (ie, at medication initiation or when concerns arise). The research team initially planned to test the intervention among those who had concerns with brands, but its suitability at medication initiation could be explored in future studies, given the reported need for information and self-management support [2,11,13,14]. Feedback was incorporated into a table of changes, with relevant quotes included for illustration (see Table S1 in Multimedia Appendix 6).

Research Team and Nurse Discussion: Prototype 3

A final discussion took place among the research team, the PPIE lead, and a Breast Cancer Now nurse, who received a copy of prototype 2 in advance. The team revisited components of the diary and the underpinning behavioral techniques, such as self-monitoring (confidence in recording symptoms) and problem-solving (identifying key points ahead of consultations), which were reviewed alongside feedback from the PAG and pharmacists. Many of the pharmacists' views informed the training package (e-learning resource reported separately), in which patient-centered and shared decision-making approaches became more central to the intervention's aim of upskilling pharmacists. Current medicine shortages were discussed as a timely opportunity to introduce the intervention, given the additional challenges associated with requesting specific brands. Specific changes related to the diary layout to ensure ease of completion and flexibility in the allocated spaces, following data from workshop contributions. Additionally, the team agreed that pharmacists would be well placed to signpost patients to charity information alongside other recommendations included in the medication consultation guide. Further details and quotes are presented in Table S1 in Multimedia Appendix 6.

Prototype 3 of the symptom diary was designed as a paper copy and in a word-processor format (to support offline completion and adaptability to systems other than Microsoft Word). The diary will be delivered by community pharmacists after training with an e-learning resource and following a medication consultation guide. Eligible patients (identified by pharmacists) are those who have previously expressed or are currently expressing concerns about MBCs. The estimated duration for completing the diary during the testing period is a minimum of 1 month and a maximum of 3 months. Prototype 3 contains the sections outlined in Textbox 1.

XSL•FO
RenderX

Textbox 1. Sections of prototype 3.

1. How to use this diary
- When and what to record; presentation of pain and mood scales
- What happens next (medication consultation appointment)
- Signposting to feedback and information on drug side effects

2. My symptom diary
- Ten questions on drug name, physical and psychological symptoms, impact on daily activities, and anything done to control the symptoms
- Use of other medication(s)
- Personal circumstances and external factors
- Whether symptoms stopped

3. Preparing for your medication consultation with the pharmacist
- Summary of patients' symptoms
- Personal life and external factors that might affect symptoms
- Actions taken to control symptoms
- Medication expectations and plans

4. Feedback notes on your experience of using the diary
- Positive and negative views

5. Common, occasional, and rare side effects of hormone therapy drugs
- Lists of tamoxifen, anastrozole, letrozole, and exemestane side effects

Discussion

Principal Findings

In line with the PBA, our planning stage, supported by data triangulation, was an essential first step given the scarcity of studies specifically focusing on MBCs for HT drugs. Patients' concerns about brand changes first emerged as a potential adherence-related factor in our previous study on the impact of new guidelines extending HT from 5 to 10 years [21]. Online patient forums provide a valuable source of data, as patients perceive them as a space for sharing experiences of coping with the disease and seeking advice on managing side effects [14]. Our findings from 277 patients in the Breast Cancer Now forum, also confirmed through interviews, suggested that patients with MBC concerns encounter 2 types of HCP responses: 1 grounded in the biomedical model, which asserts that generic drugs are bioequivalent (ie, no difference exists), and 1 based on belief in patients' lived experiences, which depends on the HCP's sensitive approach and the patient's ability to advocate for themselves. Data triangulation also highlighted a disconnect in HCP advice (from GPs, oncologists, and nurses) regarding the constraints of the system-level drug distribution process. These studies further deepened our understanding of the complexities inherent in conversations about brand changes. Following patients' and pharmacists' initial proposals regarding the use of a diary (or another method of recording side effects and brands) and a medication consultation with the pharmacist, the research team shared these ideas with PPI, stakeholders, patients, and pharmacists during the development stage. Informed by evidence on behaviors underpinning adherence to HT (self-efficacy and Social Cognitive Theory), we used qualitative research to identify MoAs, barriers and facilitators of MBCs, and candidate BCTs based on existing evidence. Drawing on the PBA for intervention development, a symptom diary was coproduced with patients and pharmacists, prioritizing acceptability and engagement with users' preferences and, as recommended by the MRC [20], its potential implementation within the community pharmacy setting.

During the development stage, the scoping review showed that the use of diaries among patients with cancer is more effective when HCP interaction (social support) is embedded within the intervention. This aligned with a core practice standard of the General Pharmaceutical Council, which places patients' interests and perspectives at the center of patient consultations [80]. However, these principles proved difficult to apply in practice for patients with BC who were concerned about HT generic switching. Therefore, a different approach was needed to engage patients and pharmacists in person-centered and shared decision-making practices. The diary is intended to help patients better understand symptoms and changes in brands, and to reflect on their expectations regarding the medication. This reflective process may better position patients to engage in shared decision-making consultations [81], as recommended by National Institute for Health and Care Excellence (NICE) adherence guidelines [82]. Patients' preferences, needs, and lived experiences can "get a discussion going," as 1 participant described it. As social support from a pharmacist appeared to be a primary motivation for diary engagement, the medication consultation strengthens the potential feasibility and

XSL•FO
RenderX

acceptability of the intervention. This addresses repeated calls for clearer articulation of supported self-management within primary care provision [83], including pharmacy practice [16,34,79,84]. Furthermore, patients with BC in the England-initiated follow-up pathway have already reported a need for medication reviews [85].

Person-centered approaches focus on understanding how users engage with, use, and perceive the relevance of an intervention. This also includes HCPs such as pharmacists, who, in this study, would deliver the intervention in their practice settings. User and stakeholder engagement through coproduction and iterative feedback processes is likely to be critical for future service implementation and may support viability across diverse user groups. For example, concerns raised by ethnically minoritized patients and pharmacists about documenting personal circumstances led to making it optional for patients to decide how much personal information they wished to share in the diary. The use of lay language, readability considerations, and diary translations may further support acceptability among underserved populations. Disagreements arising during stages of the feedback process were evaluated by prioritizing the congruence of proposed modifications with the behavioral determinants and BCTs identified in our theoretical approach. Contributors' input is transparently reported in this paper, as recommended by Slattery et al [86], making it possible to trace changes discussed by this diverse group of contributors and potential users. Further adjustments to the diary will follow completion of the testing phase in community pharmacies, informed by participant feedback.

Limitations

Our approach was strengthened by a recruitment strategy that sought to capture a plurality of views and experiences across different sociodemographic groups. However, we acknowledge the limited inclusion of patients who do not speak English, as well as lesbian and nonbinary women. We also acknowledge that the current version of the diary could be more inclusive for women with other disabilities, including those experiencing drug-related side effects (eg, hand joint pain). In the future, a mobile version with voice-recording input could be developed alongside an offline version.

Conclusions

This study has begun to address the need for support for patients with BC who have concerns about changes to their HT medication brands. The use of different brands causes concern for some patients, lowers QoL, and in some cases may result in poor adherence and treatment discontinuation. We reported on the planning and development of a symptom diary designed to improve patients' self-efficacy in medication-taking and communication with pharmacists. Women perceived the diary as "empowering" because it helped them take control of their symptoms and share this information with their pharmacist, with the expectation of developing agreed-upon medication plans. Importantly, the diary is intended to complement a medication consultation with pharmacists, who will receive training and guidance as part of the ENABLE intervention. The symptom diary demonstrates the theory-, evidence-, and person-based methods used for intervention development, documenting the process used to gather qualitative evidence from patients and pharmacists, from needs identification to diary coproduction and refinement. Behavioral analysis (MoAs and BCTs) underpinned the rationale and decisions made when incorporating feedback from a diverse population group. The use of systematic intervention reporting, particularly regarding intervention components and BCT targets, will allow for further specific adjustments. It may also enhance transferability to other contexts in which generic switching is problematic for patients with other chronic conditions.

We preliminarily developed an intervention that appears engaging, relevant, and acceptable for patients to self-monitor their symptoms. Further revisions informed by user feedback will be incorporated after testing ENABLE in community pharmacies. Large-scale research will be needed to provide evidence of its impact on improving adherence and QoL.

Acknowledgments

We are very grateful to all participants for sharing their experiences in this study. Special thanks to the ENABLE Patient Advisory Group for their invaluable contributions to the study. We also extend our gratitude to all organizations and groups that helped recruit participants, including Breast Cancer Now, South Asian Health Action, Local Pharmaceutical Committee networks, and Community Pharmacy Research Champions at RNDD North London. Our special thanks to Ieva Steinberga for her research assistant work during the planning stage and Jane Murphy, Clinical Nurse Specialist at Breast Cancer Now, for her collaboration during the development stage. The authors declare that no artificial intelligence tools were used in the ideation or writing process.

Funding

This study was funded by the NIHR Research for Patient Benefit (grant NIHR206589). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Data Availability

Study files for the development stage are available from the OSF database: "Co-development of a symptom diary for breast cancer patients (2024-2025)" available online [87].

XSL•FO
RenderX

Authors' Contributions

Conceptualization: YE, DS, RE, BN
Data curation: JC, YE
Formal analysis: YE, JC
Funding acquisition: YE (principal investigator), DS, RE, BN (coapplicants)
Investigation: YE, JC, DS, RE, BN
Methodology: YE
Supervision: YE
Validation: YE, JC, DS, RE
Visualization: YE
Writing – original draft: YE
Writing – review & editing: YE, DS, RE, BN, JC

Conflicts of Interest

None declared.

Multimedia Appendix 1
The Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist.
[PDF File (Adobe PDF File), 459 KB - cancer_v12i1e91234_app1.pdf]

Multimedia Appendix 2
Participant characteristics. Tables S1-S3.
[DOCX File , 20 KB - cancer_v12i1e91234_app2.docx]

Multimedia Appendix 3
The Guidance for the Reporting of Intervention Development (GUIDED) checklist.
[PDF File (Adobe PDF File), 268 KB - cancer_v12i1e91234_app3.pdf]

Multimedia Appendix 4
Planning stage—data analysis.
[DOCX File , 29 KB - cancer_v12i1e91234_app4.docx]

Multimedia Appendix 5
Development stage—data analysis.
[DOCX File , 58 KB - cancer_v12i1e91234_app5.docx]

Multimedia Appendix 6
Table of changes - Table S1.
[DOCX File , 13 KB - cancer_v12i1e91234_app6.docx]

References

1. Breast cancer statistics. Cancer Research UK. 2025. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#breastcs0 [accessed 2026-01-25]
2. Peddie N, Agnew S, Crawford M, Dixon D, MacPherson I, Fleming L. The impact of medication side effects on adherence and persistence to hormone therapy in breast cancer survivors: a qualitative systematic review and thematic synthesis. Breast 2021 Aug;58:147-159 [FREE Full text] [doi: 10.1016/j.breast.2021.05.005] [Medline: 34049260]
3. Ferreira A, Di Meglio A, Pistilli B, Gbenou A, El-Mouhebb M, Dauchy S, et al. Differential impact of endocrine therapy and chemotherapy on quality of life of breast cancer survivors: a prospective patient-reported outcomes analysis. Ann Oncol 2019 Nov 01;30(11):1784-1795 [FREE Full text] [doi: 10.1093/annonc/mdz298] [Medline: 31591636]
4. Yussof I, Mohd Tahir NA, Hatah E, Mohamed Shah N. Factors influencing five-year adherence to adjuvant endocrine therapy in breast cancer patients: a systematic review. Breast 2022 Apr;62:22-35 [FREE Full text] [doi: 10.1016/j.breast.2022.01.012] [Medline: 35121501]
5. National Institute for Health and Care Excellence (NICE). British National Formulary (BNF). NICE BNF. 2025. URL: https://bnf.nice.org.uk/ [accessed 2026-01-25]
6. Blencowe N, Reichl C, Gahir J, Paterson I. The use of Nolvadex in the treatment of generic tamoxifen-associated small joint arthralgia. Breast 2010 Jun;19(3):243-245. [doi: 10.1016/j.breast.2010.02.004] [Medline: 20347307]

XSL•FO
RenderX

7. Zeidan B, Anderson K, Peiris L, Rainsbury D, Laws S. The impact of tamoxifen brand switch on side effects and patient compliance in hormone receptor positive breast cancer patients. Breast 2016 Oct;29:62-67. [doi: 10.1016/j.breast.2016.07.001] [Medline: 27428472]
8. Reker D, Blum SM, Steiger C, Anger KE, Sommer JM, Fanikos J, et al. "Inactive" ingredients in oral medications. Sci Transl Med 2019 Mar 13;11(483):eaau6753 [FREE Full text] [doi: 10.1126/scitranslmed.aau6753] [Medline: 30867323]
9. Haji-Hersi MF, Tilley S, Shelton CA, Lamb N, Kamdem LK. Drug- and patient-related factors are the strongest predictors of endocrine therapy adherence in breast cancer patients. J Oncol Pharm Pract 2022 Jul 30;28(5):1070-1076. [doi: 10.1177/10781552211020805] [Medline: 34053357]
10. Heller MK, Chapman SC, Horne R. Beliefs about medication predict the misattribution of a common symptom as a medication side effect--evidence from an analogue online study. J Psychosom Res 2015 Dec;79(6):519-529. [doi: 10.1016/j.jpsychores.2015.10.003] [Medline: 26519128]
11. Eraso Y, Moon Z, Steinberga I. Patients' experience of medication brand changes during hormone therapy for breast cancer-an interpretative phenomenological analysis. Healthcare (Basel) 2022 Dec 16;10(12):2558 [FREE Full text] [doi: 10.3390/healthcare10122558] [Medline: 36554081]
12. Cahir C, Bennett K, Dombrowski SU, Kelly CM, Wells M, Watson E, et al. Informing interventions to improve uptake of adjuvant endocrine therapy in women with breast cancer: a theoretical-based examination of modifiable influences on non-adherence. Support Care Cancer 2023 Mar 04;31(3):200 [FREE Full text] [doi: 10.1007/s00520-023-07658-x] [Medline: 36869943]
13. Brett J, Boulton M, Fenlon D, Hulbert-Williams NJ, Walter FM, Donnelly P, et al. Adjuvant endocrine therapy after breast cancer: a qualitative study of factors associated with adherence. PPA 2018 Feb;Volume 12:291-300. [doi: 10.2147/ppa.s145784]
14. Clancy C, Lynch J, OConnor P, Dowling M. Breast cancer patients' experiences of adherence and persistence to oral endocrine therapy: a qualitative evidence synthesis. Eur J Oncol Nurs 2020 Feb;44:101706. [doi: 10.1016/j.ejon.2019.101706] [Medline: 32007696]
15. Royal Pharmaceutical Society. Utilising community pharmacists to support people with cancer. Royal Pharmaceutical Society. 2020 Jan 10. URL: https://www.rpharms.com/recognition/all-our-campaigns/policy-a-z/cancer-support [accessed 2025-11-25]
16. Ryan R, Santesso N, Lowe D, Hill S, Grimshaw J, Prictor M. Interventions to improve safe and effective medicines use by consumers: an overview of systematic reviews. Cochrane Library 2014 Apr 29;2014(4):CD007768. [doi: 10.1002/14651858.cd007768.pub3]
17. Patti M, Renfro CP, Posey R, Wu G, Turner K, Ferreri SP. Systematic review of medication synchronization in community pharmacy practice. Res Social Adm Pharm 2019 Nov;15(11):1281-1288. [doi: 10.1016/j.sapharm.2018.11.008] [Medline: 30482457]
18. Nguyen E, Sobieraj DM. The impact of appointment-based medication synchronization on medication taking behaviour and health outcomes: a systematic review. J Clin Pharm Ther 2017 Aug 08;42(4):404-413. [doi: 10.1111/jcpt.12554] [Medline: 28485006]
19. Yardley L, Ainsworth B, Arden-Close E, Muller I. The person-based approach to enhancing the acceptability and feasibility of interventions. Pilot Feasibility Stud 2015 Oct 26;1(1):37 [FREE Full text] [doi: 10.1186/s40814-015-0033-z] [Medline: 27965815]
20. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
21. Eraso Y, Stefler D, Moon Z, Rossi L, Assefa S. Extending adjuvant endocrine therapy for 10 years: a mixed-methods analysis of women's decision making in an online breast cancer forum. Healthcare (Basel) 2021 Jun 07;9(6):688 [FREE Full text] [doi: 10.3390/healthcare9060688] [Medline: 34200326]
22. Smith J, Flowers P, Larkin M. Interpretative Phenomenological Analysis: Theory, Method and Research. London, UK: Sage; 2009.
23. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
24. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implement Sci 2017 Jun 21;12(1):77 [FREE Full text] [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
25. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
26. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. Qual Health Res 2006 Mar;16(3):377-394. [doi: 10.1177/1049732305285708] [Medline: 16449687]
27. Duncan E, O'Cathain A, Rousseau N, Croot L, Sworn K, Turner KM, et al. Guidance for reporting intervention development studies in health research (GUIDED): an evidence-based consensus study. BMJ Open 2020 Apr 08;10(4):e033516 [FREE Full text] [doi: 10.1136/bmjopen-2019-033516] [Medline: 32273313]

XSL•FO
RenderX

28. Bandura A. Social Foundations of Thought and Action. A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall, Inc; 1986.
29. Shelby RA, Edmond SN, Wren AA, Keefe FJ, Peppercorn JM, Marcom PK, et al. Self-efficacy for coping with symptoms moderates the relationship between physical symptoms and well-being in breast cancer survivors taking adjuvant endocrine therapy. Support Care Cancer 2014 Oct 13;22(10):2851-2859. [doi: 10.1007/s00520-014-2269-1] [Medline: 24821365]
30. Toivonen K, Williamson T, Carlson L, Walker L, Campbell T. Potentially modifiable factors associated with adherence to adjuvant endocrine therapy among breast cancer survivors: a systematic review. Cancers (Basel) 2020 Dec 31;13(1):107 [FREE Full text] [doi: 10.3390/cancers13010107] [Medline: 33561076]
31. Bandura A. The explanatory and predictive scope of self-efficacy theory. Journal of Social and Clinical Psychology 1986 Sep;4(3):359-373. [doi: 10.1521/jscp.1986.4.3.359]
32. Kimmick G, Edmond SN, Bosworth HB, Peppercorn J, Marcom PK, Blackwell K, et al. Medication taking behaviors among breast cancer patients on adjuvant endocrine therapy. Breast 2015 Oct;24(5):630-636 [FREE Full text] [doi: 10.1016/j.breast.2015.06.010] [Medline: 26189978]
33. Liu Y, Malin JL, Diamant AL, Thind A, Maly RC. Adherence to adjuvant hormone therapy in low-income women with breast cancer: the role of provider-patient communication. Breast Cancer Res Treat 2013 Feb;137(3):829-836 [FREE Full text] [doi: 10.1007/s10549-012-2387-8] [Medline: 23263740]
34. Smits M, Mammatas L, Schoonhoven L, Vervoort S. Improving medication adherence to endocrine therapy in breast cancer patients: a mixed-methods systematic review of effective communication strategies for healthcare providers. Breast 2025 Aug;82:104510 [FREE Full text] [doi: 10.1016/j.breast.2025.104510] [Medline: 40450986]
35. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Education and Counseling 2002 Oct;48(2):177-187. [doi: 10.1016/S0738-3991(02)00032-0]
36. Wiecek E, Tonin FS, Torres-Robles A, Benrimoj SI, Fernandez-Llimos F, Garcia-Cardenas V. Temporal effectiveness of interventions to improve medication adherence: a network meta-analysis. PLoS One 2019;14(3):e0213432 [FREE Full text] [doi: 10.1371/journal.pone.0213432] [Medline: 30861014]
37. Finitsis DJ, Vose BA, Mahalak JG, Salner AL. Interventions to promote adherence to endocrine therapy among breast cancer survivors: a meta-analysis. Psychooncology 2019 Feb 28;28(2):255-263. [doi: 10.1002/pon.4959] [Medline: 30511789]
38. McGuinness S, Hughes L, Moss-Morris R, Hunter M, Norton S, Moon Z. Adherence to adjuvant endocrine therapy among White British and ethnic minority breast cancer survivors in the United Kingdom. Eur J Cancer Care (Engl) 2022 Nov 18;31(6):e13722 [FREE Full text] [doi: 10.1111/ecc.13722] [Medline: 36255032]
39. Cancer in the UK 2025: socioeconomic deprivation. Cancer Research UK. 2025. URL: https://www.cancerresearchuk.org/sites/default/files/cancer_in_the_uk_2025_socioeconomic_deprivation.pdf [accessed 2025-11-27]
40. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z. Scoping reviews. In: JBI Manual for Evidence Synthesis. Adelaide, SA, Australia: JBI; 2024.
41. Eraso Y, Cochrane J, Stewart S, Edwards R, Nanray B. Diary interventions and self-monitoring behaviours in cancer survivors. A rapid scoping review protocol. London Metropolitan University. 2025 Jan. URL: https://repository.londonmet.ac.uk/id/eprint/10517 [accessed 2025-11-15]
42. Farr M. Power dynamics and collaborative mechanisms in co-production and co-design processes. Critical Social Policy 2017 Dec 13;38(4):623-644. [doi: 10.1177/0261018317747444]
43. O'Cathain A. Workshops as a qualitative research method in health research. BMJ Open 2025 Sep 26;15(9):e106459 [FREE Full text] [doi: 10.1136/bmjopen-2025-106459] [Medline: 41005775]
44. Guidance on co-producing a research project. NIHR. 2021 Apr. URL: https://www.learningforinvolvement.org.uk/wp-content/uploads/2021/04/NIHR-Guidance-on-co-producing-a-research-project-April-2021.pdf [accessed 2025-11-20]
45. Theory and technique tool. Human Behaviour Change Project. 2022. URL: https://theoryandtechniquetool.humanbehaviourchange.org/ [accessed 2026-01-15]
46. Muller I, Santer M, Morrison L, Morton K, Roberts A, Rice C, et al. Combining qualitative research with PPI: reflections on using the person-based approach for developing behavioural interventions. Res Involv Engagem 2019;5:34 [FREE Full text] [doi: 10.1186/s40900-019-0169-8] [Medline: 31807316]
47. World Health Organization (WHO). Adherence to long-term therapies: evidence for action. WHO. 2003. URL: https://iris.who.int/handle/10665/42682 [accessed 2026-01-13]
48. Advanced Service Specification - NHS New Medicine Service (NMS). National Health Service (NHS). 2021. URL: https://www.england.nhs.uk/publication/advanced-service-specification-nhs-new-medicine-service-nms/ [accessed 2025-09-13]
49. NHS Prescription Services. National Health Service (NHS). URL: https://www.nhsbsa.nhs.uk/pharmacies-gp-practices-and-appliance-contractors/prescribing-and-dispensing [accessed 2025-09-13]
50. Lupton D. The digitally engaged patient: self-monitoring and self-care in the digital health era. Soc Theory Health 2013 Jun 19;11(3):256-270. [doi: 10.1057/sth.2013.10]
51. Vallance J, Lesniak S, Belanger L, Courneya K. Development and assessment of a physical activity guidebook for the Colon Health and Life-Long Exercise Change (CHALLENGE) trial (NCIC CO.21). J Phys Act Health 2010 Nov;7(6):794-801. [doi: 10.1123/jpah.7.6.794] [Medline: 21088311]

XSL•FO
RenderX

52. Olesen ML, Duun-Henriksen A, Hansson H, Ottesen B, Andersen KK, Zoffmann V. A person-centered intervention targeting the psychosocial needs of gynecological cancer survivors: a randomized clinical trial. J Cancer Surviv 2016 Oct 22;10(5):832-841. [doi: 10.1007/s11764-016-0528-5] [Medline: 26902366]
53. Suh EE, Kim H, Kang J, Kim H, Park KO, Jeong BL, et al. Outcomes of a culturally responsive health promotion program for elderly Korean survivors of gastrointestinal cancers: a randomized controlled trial. Geriatr Nurs 2013 Nov;34(6):445-452. [doi: 10.1016/j.gerinurse.2013.09.007] [Medline: 24156925]
54. Kekäle M, Söderlund T, Koskenvesa P, Talvensaari K, Airaksinen M. Impact of tailored patient education on adherence of patients with chronic myeloid leukaemia to tyrosine kinase inhibitors: a randomized multicentre intervention study. J Adv Nurs 2016 Sep;72(9):2196-2206. [doi: 10.1111/jan.12978] [Medline: 27113362]
55. Edbrooke L, Khaw P, Freimund A, Carpenter D, McNally O, Joubert L, et al. ENhAncing Lifestyle Behaviors in EndometriaL CancEr (ENABLE): a pilot randomized controlled trial. Integr Cancer Ther 2022 Jan 19;21:15347354211069885 [FREE Full text] [doi: 10.1177/15347354211069885] [Medline: 35045735]
56. Ainsworth MC, Pekmezi D, Bowles H, Ehlers D, McAuley E, Courneya KS, et al. Acceptability of a mobile phone app for measuring time use in breast cancer survivors (Life in a Day): mixed-methods study. JMIR Cancer 2018 May 14;4(1):e9. [doi: 10.2196/cancer.8951] [Medline: 29759953]
57. Sun V, Guthrie KA, Crane TE, Arnold KB, Colby S, Freylersythe SG, et al. SWOG S1820: a pilot randomized trial of the altering intake, managing bowel symptoms intervention in survivors of rectal cancer. Cancer 2024 Jul 01;130(13):2384-2394 [FREE Full text] [doi: 10.1002/cncr.35264] [Medline: 38386696]
58. Koutoukidis DA, Beeken RJ, Manchanda R, Michalopoulou M, Burnell M, Knobf MT, et al. Recruitment, adherence, and retention of endometrial cancer survivors in a behavioural lifestyle programme: the Diet and Exercise in Uterine Cancer Survivors (DEUS) parallel randomised pilot trial. BMJ Open 2017 Oct 08;7(10):e018015 [FREE Full text] [doi: 10.1136/bmjopen-2017-018015] [Medline: 28993394]
59. Sun V, Crane TE, Freylersythe S, Slack SD, Yung A, Krouse RS, et al. Altering intake and managing symptoms: feasibility of a diet modification intervention for post-treatment bowel dysfunction in rectal cancer. Clin J Oncol Nurs 2022 Jun 01;26(3):283-292. [doi: 10.1188/22.CJON.283-292] [Medline: 35604741]
60. Koutoukidis DA, Beeken RJ, Manchanda R, Burnell M, Ziauddeen N, Michalopoulou M, et al. Diet, physical activity, and health-related outcomes of endometrial cancer survivors in a behavioral lifestyle program: the Diet and Exercise in Uterine Cancer Survivors (DEUS) parallel randomized controlled pilot trial. Int J Gynecol Cancer 2019 Mar;29(3):531-540 [FREE Full text] [doi: 10.1136/ijgc-2018-000039] [Medline: 30723098]
61. Lee MK, Yun YH, Park H, Lee ES, Jung KH, Noh D. A Web-based self-management exercise and diet intervention for breast cancer survivors: pilot randomized controlled trial. Int J Nurs Stud 2014 Dec;51(12):1557-1567. [doi: 10.1016/j.ijnurstu.2014.04.012] [Medline: 24856854]
62. van Cappellen-van Maldegem SJM, Mols F, Seidell JC, de Kruif A, van de Poll-Franse LV, Hoedjes M. Using the behavior change wheel to identify and understand key facilitators and barriers for lifestyle care for postmenopausal breast cancer survivors: a Delphi-study. Ann Behav Med 2024 Jan 01;58(1):22-36 [FREE Full text] [doi: 10.1093/abm/kaad049] [Medline: 37672352]
63. Nelson V, Cross AJ, Powell J, Shaw C. Can people living with and beyond colorectal cancer make lifestyle changes with the support of health technology: a feasibility study. J Hum Nutr Diet 2023 Apr 09;36(2):554-565. [doi: 10.1111/jhn.13008] [Medline: 35320595]
64. Courneya KS, Vardy JL, O'Callaghan CJ, Friedenreich CM, Campbell KL, Prapavessis H, et al. Effects of a structured exercise program on physical activity and fitness in colon cancer survivors: one year feasibility results from the CHALLENGE trial. Cancer Epidemiol Biomarkers Prev 2016 Jun;25(6):969-977. [doi: 10.1158/1055-9965.EPI-15-1267] [Medline: 27197271]
65. Oldenmenger WH, Echteld MA, de Wit R, Sillevis Smitt PAE, Stronks DL, Stoter G, et al. Analgesic adherence measurement in cancer patients: comparison between electronic monitoring and diary. J Pain Symptom Manage 2007 Dec;34(6):639-647 [FREE Full text] [doi: 10.1016/j.jpainsymman.2007.01.015] [Medline: 17703909]
66. Brett J, Boulton M, Watson E. Development of an e-health app to support women prescribed adjuvant endocrine therapy after treatment for breast cancer. Patient Prefer Adherence 2018;12:2639-2647 [FREE Full text] [doi: 10.2147/PPA.S187692] [Medline: 30587936]
67. Sun V, Crane TE, Arnold KB, Guthrie K, Freylersythe S, Braun-Inglis C, et al. SWOG S1820: altering intake, managing symptoms for bowel dysfunction in survivors of rectal cancer (The AIMS-RC intervention trial). Contemp Clin Trials Commun 2021 Jun;22:100768. [doi: 10.1016/j.conctc.2021.100768] [Medline: 33997460]
68. Paterson C, Armitage L, Turner M. Current landscape of ecological momentary assessment (real-time data) methodology in cancer research: a systematic review. Semin Oncol Nurs 2023 Dec;39(6):151514 [FREE Full text] [doi: 10.1016/j.soncn.2023.151514] [Medline: 37865555]
69. Short CE, James EL, Plotnikoff RC. Theory-and evidence-based development and process evaluation of the Move More for Life program: a tailored-print intervention designed to promote physical activity among post-treatment breast cancer survivors. Int J Behav Nutr Phys Act 2013 Nov 05;10:124 [FREE Full text] [doi: 10.1186/1479-5868-10-124] [Medline: 24192320]

XSL•FO
RenderX

70. Henshall C, Davey Z. Development of an app for lung cancer survivors (iEXHALE) to increase exercise activity and improve symptoms of fatigue, breathlessness and depression. Psychooncology 2020 Jan;29(1):139-147. [doi: 10.1002/pon.5252] [Medline: 31773808]
71. Koutoukidis DA, Lopes S, Atkins L, Croker H, Knobf MT, Lanceley A, et al. Use of intervention mapping to adapt a health behavior change intervention for endometrial cancer survivors: the shape-up following cancer treatment program. BMC Public Health 2018 Mar 27;18(1):415 [FREE Full text] [doi: 10.1186/s12889-018-5329-5] [Medline: 29587699]
72. Paterson C, Primeau C, Lauder W. What are the experiences of men affected by prostate cancer participating in an ecological momentary assessment study? Cancer Nurs 2019 Mar 19;43(4):300-310. [doi: 10.1097/ncc.0000000000000699]
73. Courneya KS, Booth CM, Gill S, O'Brien P, Vardy J, Friedenreich CM, et al. The Colon Health and Life-Long Exercise Change trial: a randomized trial of the National Cancer Institute of Canada Clinical Trials Group. Curr Oncol 2008 Dec 01;15(6):279-285 [FREE Full text] [doi: 10.3747/co.v15i6.378] [Medline: 19079628]
74. Sun V, Crane TE, Slack SD, Yung A, Wright S, Sentovich S, et al. Rationale, development, and design of the Altering Intake, Managing Symptoms (AIMS) dietary intervention for bowel dysfunction in rectal cancer survivors. Contemp Clin Trials 2018 May;68:61-66 [FREE Full text] [doi: 10.1016/j.cct.2018.03.010] [Medline: 29567283]
75. Koutoukidis DA, Beeken RJ, Manchanda R, Burnell M, Knobf MT, Lanceley A. Diet and exercise in uterine cancer survivors (DEUS pilot) - piloting a healthy eating and physical activity program: study protocol for a randomized controlled trial. Trials 2016 Mar 10;17(1):130. [doi: 10.1186/s13063-016-1260-1] [Medline: 26965165]
76. Olesen ML, Hansson H, Ottesen B, Thranov IR, Thisted LB, Zoffmann V. The psychosocial needs of gynaecological cancer survivors: a framework for the development of a complex intervention. Eur J Oncol Nurs 2015 Aug;19(4):349-358. [doi: 10.1016/j.ejon.2015.01.007] [Medline: 25697547]
77. van Cappellen-van Maldegem SJM, Mols F, Horevoorts N, de Kruif A, Buffart LM, Schoormans D, OPTIMUM research team. Towards OPtimal TIming and Method for promoting sUstained adherence to lifestyle and body weight recommendations in postMenopausal breast cancer survivors (the OPTIMUM-study): protocol for a longitudinal mixed-method study. BMC Womens Health 2021 Jul 06;21(1):268 [FREE Full text] [doi: 10.1186/s12905-021-01406-1] [Medline: 34229690]
78. Huang Y, Li Q, Zhou F, Song J. Effectiveness of internet-based support interventions on patients with breast cancer: a systematic review and narrative synthesis. BMJ Open 2022 May 31;12(5):e057664 [FREE Full text] [doi: 10.1136/bmjopen-2021-057664] [Medline: 35641011]
79. Dang TH, Forkan ARM, Wickramasinghe N, Jayaraman PP, Alexander M, Burbury K, et al. Investigation of intervention solutions to enhance adherence to oral anticancer medicines in adults: overview of reviews. JMIR Cancer 2022 Apr 27;8(2):e34833 [FREE Full text] [doi: 10.2196/34833] [Medline: 35475978]
80. Standards for pharmacy professionals. General Pharmaceutical Council. 2017 May. URL: https://assets.pharmacyregulation.org/files/standards_for_pharmacy_professionals_may_2017_0.pdf?VersionId=C8dRrU1opDLdsuveSss5cKsPSwKObTi2 [accessed 2025-07-30]
81. Shared decision making. National Institute for Health and Care Excellence (NICE). URL: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-guidelines/shared-decision-making [accessed 2026-04-20]
82. Medicines adherence: involving patients in decisions about prescribed medicines and supporting adherence. National Institute for Health and Care Excellence (NICE). 2009. URL: https://www.nice.org.uk/guidance/cg76/chapter/Recommendations#reviewing-medicines [accessed 2025-08-27]
83. Howell D, Mayer D, Fielding R, Eicher M, Verdonck-de Leeuw IM, Johansen C, Global Partners for Self-Management in Cancer. Management of cancer and health after the clinic visit: a call to action for self-management in cancer care. J Natl Cancer Inst 2021 May 04;113(5):523-531 [FREE Full text] [doi: 10.1093/jnci/djaa083] [Medline: 32525530]
84. Olsen NL, Jacobsen R, Thomsen LA, Nørgaard LS. Involving community pharmacies in management of late effects of cancer treatment: opinions from cancer survivors. Explor Res Clin Soc Pharm 2024 Dec;16:100514 [FREE Full text] [doi: 10.1016/j.rcsop.2024.100514] [Medline: 39399766]
85. Moore L, Matheson L, Brett J, Lavender V, Kendall A, Lavery B, et al. Optimising patient-initiated follow-up care - a qualitative analysis of women with breast cancer in the UK. Eur J Oncol Nurs 2022 Oct;60:102183 [FREE Full text] [doi: 10.1016/j.ejon.2022.102183] [Medline: 35932754]
86. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
87. Open Science Framework (OSF). URL: https://10.17605/OSF.IO/3TMJ4 [accessed 2026-05-14]

Abbreviations

BC: breast cancer
BCT: behavior change technique
COREQ: Consolidated Criteria for Reporting Qualitative Research
ENABLE: Medication Brand Changes in Hormone Therapy for Breast Cancer: A Community Pharmacy Intervention Development to Improve Patients' Adherence and Quality of Life
GP: general physician

XSL•FO
RenderX

GUIDED: Guidance for the Reporting of Intervention Development
HCP: health care professional
HT: Hormone therapy
IPA: interpretative phenomenological analysis
ISRCTN: International Standard Randomised Controlled Trial Number
MBC: medication brand change
MeSH: Medical Subject Headings
MoA: mechanisms of action
NHS: National Health Service
NICE: National Institute for Health and Care Excellence
PAG: patient advisory group
PBA: person-based approach
PIS: participant information sheet
PPIE: patient and public involvement and engagement
QoL: quality of life

Edited by M Balcarras; submitted 11.Jan.2026; peer-reviewed by C Lei, C Odo; comments to author 14.Apr.2026; accepted 08.May.2026; published 26.May.2026.

Please cite as:
Eraso Y, Stewart D, Edwards R, Nanray B, Cochrane J
Planning and Developing a Symptom Diary Intervention for Breast Cancer Survivors With Concerns About Medication Brands (ENABLE Study): User-Centered Design Approach
JMIR Cancer 2026;12:e91234
URL: https://cancer.jmir.org/2026/1/e91234
doi:10.2196/91234
PMID:

©Yolanda Eraso, Duncan Stewart, Ruth Edwards, Bal Nanray, Jacqueline Cochrane. Originally published in JMIR Cancer (https://cancer.jmir.org), 26.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient-Reported Outcomes on Quality of Life and Psychological Distress After Focal High-Intensity Focused Ultrasound for Prostate Cancer: Prospective Multicenter Study

Toni Franz[1], Dr med; Cornelia Hartung[2], Dr rer med; Roman Ganzer[3], Prof Dr med; Sigrun Holze[1], Dr rer med; Jens-Uwe Stolzenburg[1], Prof Dr med

[1]Department of Urology, University Hospital Leipzig, Liebigstraße 20, Leipzig, Saxony, Germany

[2]Leipzig Network for Integrative Psychotherapy, Leipzig, Saxony, Germany

[3]Department of Urology, Asklepios Hospital Bad Tölz, Bad Tölz, Bavaria, Germany

Corresponding Author:
Toni Franz, Dr med
Department of Urology, University Hospital Leipzig, Liebigstraße 20, Leipzig, Saxony, Germany

Abstract

Background: Focal high-intensity focused ultrasound (HIFU) is an emerging tissue-preserving treatment for localized prostate cancer (PCa) that aims to reduce functional impairment and psychological burden while maintaining oncological safety. Although its clinical use is increasing, prospective data on health-related quality of life (HRQoL) and psychological distress after focal hemiablation remain limited.

Objective: This study aims to evaluate longitudinal HRQoL trajectories and psychological outcomes in patients eligible for active surveillance who underwent focal HIFU following shared decision-making.

Methods: This prospective, multicenter, single-arm longitudinal study included men with unilateral, organ-confined PCa treated with focal HIFU hemiablation within the HEMI study (AP study protocol/reference number: 68/11, AUO). All patients had concordant multiparametric magnetic resonance imaging and biopsy findings, and underwent standardized unilateral HIFU. HRQoL was assessed at baseline and at 1, 3, 6, 9, and 12 months using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire—Core 30 (EORTC QLQ-C30, version 3.0). Psychological distress was measured with the Hospital Anxiety and Depression Scale. Longitudinal changes were analyzed using linear mixed-effects models with false discovery rate correction. In addition, univariate regression analyses were performed to assess associations between clinical and functional parameters, and HRQoL outcomes. Outcomes were compared with German and Norwegian normative populations, and clinical relevance was interpreted using Osoba criteria.

Results: Fifty-four patients were included, of whom 51 (94.4%) completed 12-month follow-up (mean age 63.4, SD 8.3 y). Baseline HRQoL, functional status, and symptom burden were favorable. One month after hemi-HIFU, a statistically significant and clinically moderate decline in global health status was observed (mean 75 [SD 17.8] vs 63 [SD 17.9]; P=.002; Cohen d=0.58). This transient deterioration was mainly driven by increased fatigue, pain, insomnia, appetite loss, and reduced physical, role, and social functioning, with social functioning most affected. Emotional and cognitive functioning remained stable. From 3 months onward, HRQoL and all functional domains recovered to baseline or higher levels, with no clinically relevant impairments at 12 months. Symptom burden normalized across most domains after 3 months, although mild constipation remained slightly elevated. Depression scores showed a transient increase at 1 and 3 months but remained below clinical thresholds, whereas anxiety scores were stable at all time points. In univariate regression analyses, no clinical or functional parameters were identified as independent predictors of HRQoL. Overall, recovery was rapid, long-term HRQoL was preserved, and treatment-related morbidity remained low.

Conclusions: Focal HIFU hemiablation preserves long-term HRQoL and psychological well-being in carefully selected patients with localized PCa. Early postoperative impairment is temporary, with recovery to baseline and convergence with normative population values from 3 months onward. Clinical and functional parameters were not significant predictors of HRQoL. These findings support focal HIFU as a minimally invasive quality-of-life–preserving treatment option.

(JMIR Cancer 2026;12:e91984) doi:10.2196/91984

KEYWORDS

prostate cancer; focal HIFU; high-intensity focused ultrasound; hemiablation; quality of life; psychological distress; patient-reported outcomes

XSL•FO
RenderX

Introduction

Prostate cancer (PCa) is one of the most commonly diagnosed malignancies among men worldwide and represents a major global health burden [1-3]. Although radical prostatectomy (RP) and radiotherapy (RT) provide excellent oncological control, both are associated with substantial risks of urinary incontinence and erectile dysfunction, which can significantly impair health-related quality of life (HRQoL) [4-7]. Concurrently, widespread early detection has increased the diagnosis of low-risk and potentially indolent tumors, intensifying concerns about overtreatment [8-10].

Active surveillance (AS) is an established strategy to avoid unnecessary intervention in low-risk PCa. However, it is limited by uncertainties in tumor staging and biological behavior. In addition, the psychological burden of living with an untreated malignancy leads up to 20% of patients to choose definitive treatment despite low oncological risk, and approximately one-third discontinue AS within 2 years [11,12]. Repeated prostate biopsies may further impair erectile function, prompting some eligible patients to pursue radical therapy [13,14].

These limitations highlight the need for treatment approaches that preserve oncological safety while minimizing functional and psychological morbidity. In this context, focal therapy has emerged as a promising alternative, aiming to eradicate the index lesion while sparing uninvolved prostate tissue. High-intensity focused ultrasound (HIFU) is one of the most extensively studied focal modalities and has been shown to reduce treatment-related side effects compared with whole-gland approaches [15-20]. Hemiablation—often referred to as “male lumpectomy”—provides lobe-directed tumor control in carefully selected patients [21-24]. Accurate tumor localization is essential for safe and effective focal therapy planning and is optimally achieved using multiparametric magnetic resonance imaging (mpMRI) combined with systematic and/or targeted biopsy [25-29].

Across all PCa treatment strategies, preservation of HRQoL has become a central outcome measure and an important component of shared decision-making. Quality of life, originally introduced as a sociopolitical concept in the 1960s [30,31], is defined by the World Health Organization as a subjective, multidimensional construct encompassing physical, psychological, social, and environmental domains [32]. Owing to its personal and nondirectly measurable nature, HRQoL assessment relies on validated patient-reported outcome measures [33,34].

In oncology, HRQoL reflects a paradigm shift toward a biopsychosocial understanding of disease and plays a growing role in clinical evaluation, health-economic decision-making, and outcomes research [35,36]. Given that oncological success may be offset by functional impairments, HRQoL often becomes a decisive factor in individual treatment selection [37]. Despite the increasing adoption of focal HIFU, prospective data on HRQoL and psychological outcomes following hemiablation remain limited.

This study therefore aimed to evaluate longitudinal changes in HRQoL and psychological distress in patients undergoing focal HIFU hemiablation for unilateral PCa who met the criteria for AS but opted for focal therapy after shared decision-making. In particular, this study examined short-term and 12-month patient-reported outcomes to assess whether focal HIFU preserves quality of life and psychological well-being in this carefully selected patient population.

Methods

Patient Cohort and Study Design

This study represents a secondary analysis of prospectively collected patient-reported outcome data from the multicenter HEMI study by Ganzer et al [38], a prospective single-arm clinical trial evaluating focal HIFU hemiablation in men with unilateral, organ-confined PCa. Key functional and oncological outcomes were predefined primary endpoints of the parent HEMI study and have been published previously [38]. The study was registered with the German Working Group on Urological Oncology (AUO) under study protocol AP 68/11.

Patient-reported outcomes, including HRQoL and psychological distress, were prospectively collected according to the study protocol. The present analysis was conducted after the completion of data collection and focuses specifically on longitudinal changes in these outcomes over a 12-month follow-up period.

The study population comprised 54 consecutively enrolled patients who fulfilled the criteria for AS [39] and presented with histologically confirmed unilateral disease and concordant findings on mpMRI. At the time of study conduct, focal therapy was recommended to be performed only within clinical trials. All patients were therefore recruited within the HEMI study after being informed about guideline-conforming treatment options, including AS, RP, and RT, within a shared decision-making process. The final sample size reflects the strictly selected patient population, which applied predefined inclusion and exclusion criteria to identify men suitable for focal hemiablation. All patients were consecutively enrolled at the participating centers during the study period. The present analysis included all patients with available baseline and follow-up patient-reported outcome data. No additional selection or sampling was performed. Therefore, the sample size represents the complete cohort of eligible patients within the study framework. Patients who developed secondary malignancies during follow-up were excluded from longitudinal analyses, as these represent competing health events that would confound the assessment of HRQoL outcomes.Textbox 1

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria (all must be fulfilled)

- Male patients aged 18 years and older
- Prostate-specific antigen (PSA) ≤10 ng/mL
- Clinical stage T1c-T2a
- Biopsy-proven unilateral prostate cancer detected by magnetic resonance imaging (MRI) or transrectal ultrasound scan (TRUS) fusion biopsy
- 30% or fewer positive biopsies
- Gleason score of 3+4=7 or less
- Peripheral zone height to be 30 mm or less and 40 mm or less on TRUS in patients treated with Ablatherm integrated imaging and the focal one device

Exclusion criteria (excluded if *any* of the following criteria applied)

- Significant cancer of the contralateral side on multiparametric MRI as defined by a score of 4 or greater on PIRADS (Prostate Imaging Reporting and Data System)
- Previous prostatic or urethral surgery
- Intake of 5-α reductase inhibitor
- Previous androgen deprivation therapy
- Incomplete clinical or imaging data (eg, missing PSA)

The therapeutic strategy and technical execution of focal HIFU therapy followed the standardized protocol described by Ganzer et al [38], which served as the clinical reference framework for the present analysis. Baseline assessments were performed on the day prior to treatment. Patient enrollment took place between April 2013 and March 2016, corresponding to a recruitment period of 2 years and 11 months. Follow-up evaluations were conducted at 1, 3, 6, and 9 months post intervention using standardized mailed questionnaires. The final 12-month assessment was performed in conjunction with a scheduled rebiopsy at the study center. Time was included as a fixed effect in the longitudinal models without specification of random slopes. The study protocol was initially approved by the institutional review board of the University of Regensburg (Ethics Committee reference: 11-101-0156; approval date: December 29, 2011) and thereafter by the other participating study centers. All participants provided written informed consent prior to inclusion. The inclusion and exclusion criteria applied during study selection are summarized in Textbox 1.

HRQoL Assessment (EORTC QLQ-C30)

HRQoL was assessed using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire—Core 30 (EORTC QLQ-C30, version 3.0), a validated and internationally established instrument extensively used in oncological research [40]. The questionnaire comprises 30 items assessing:

- Global health status and quality of life
- Five functional domains (physical, role, emotional, cognitive, and social)
- Three multi-item symptom scales (fatigue, pain, and nausea and vomiting)
- Six single-item symptom dimensions (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties)

Items 1 to 28 are rated on a 4-point Likert scale, ranging from “not at all” to “very much,” whereas items 29 and 30 are scored on a 7-point scale assessing global health and overall quality of life.

All scales were scored and transformed according to the EORTC scoring manual [41], resulting in standardized scores ranging from 0 to 100. Higher scores on global and functional scales reflect better functioning and quality of life, whereas higher scores on symptom scales indicate greater symptom burden (Table 1). Raw scores were calculated as the arithmetic mean of completed items per scale and subsequently linearly transformed based on each scale’s scoring range.

Longitudinal changes in HRQoL were analyzed using linear mixed-effects models with random intercepts for participants to account for within-patient correlations over time. Time was included as a fixed effect. Analyses were conducted using restricted maximum likelihood estimation. To control for type I error inflation resulting from multiple testing, false discovery rate (FDR) adjustment according to Benjamini-Hochberg was applied.

XSL•FO
RenderX

Table . Evaluation of global health status and quality of life, functional scales, general symptom and problem scales, and symptom scales.

Calculated scale symptoms score	Global health status and functional scales	Symptom scales
100	No impairment	Severe symptoms
76 - 99	Slight impairment	Moderate to severe symptoms
51 - 75	Mild impairment	Moderate symptoms
26 - 50	Moderate impairment	Few symptoms
1 - 25	Moderate to severe impairment	Slight symptoms
0	Severe impairment	No symptoms

Comparison With Normative Reference Data and Interpretation of Clinical Relevance

EORTC QLQ-C30 scores of the study cohort were compared with male normative reference values from 2 population-based datasets of the EORTC validation program: the German cohort reported by Schwarz and Hinz and the Norwegian reference population published by Hjermstad et al [42,43]. Inferential analyses were performed exclusively using the German normative dataset, while Norwegian reference values were used for descriptive comparison only. Statistical comparisons with the German population were conducted using Welch two-sided *t* test, with *P* values adjusted for multiple testing using the Benjamini-Hochberg FDR procedure.

Clinical relevance was interpreted according to the classification system proposed by Osoba et al [44], which defines mean score differences of 5 to 10 points as small, 10 to 20 points as moderate, and >20 points as clinically large. As these thresholds were originally derived from breast and lung cancer populations, their applicability to PCa should be interpreted with caution. Only differences meeting predefined criteria for clinical relevance were included in the comparative summary table. All comparisons with normative populations were performed without adjustment for age or comorbidity and are therefore interpreted descriptively rather than causally.

Psychological Assessment

Psychological distress was assessed using the Hospital Anxiety and Depression Scale (HADS), a validated self-report instrument specifically developed for use in medically ill populations [45,46]. The questionnaire comprises 14 items, with 7 items each evaluating anxiety and depressive symptoms. Each item is rated on a 4-point scale (0 - 3), yielding subscale scores ranging from 0 to 21. Clinically relevant cutoff values were set at ≥11, in accordance with published validation studies [47]. Scoring followed the recommendations of the official HADS manual.

Statistical Analysis

All questionnaire data were entered into an electronic database and analyzed using DATAtab (Seiersberg, Austria). Continuous variables are reported as mean (SD). Although individual questionnaire items are ordinal in nature, aggregated scale and subscale scores were treated as approximately interval-level variables to enable parametric statistical analyses.

Longitudinal changes in HRQoL were analyzed using linear mixed-effects models with random intercepts for participants to account for within-patient correlations over time. Time was included as a fixed effect, and analyses were conducted using restricted maximum likelihood estimation. Comparisons between the study cohort and normative reference populations were conducted using Welch *t* test to account for unequal variances. To control for type I error inflation resulting from multiple testing, FDR adjustment according to Benjamini-Hochberg was applied to all longitudinal analyses and normative comparisons across EORTC QLQ-C30 scales and time points. All statistical tests were 2-sided, and statistical significance was defined as $P<.05$.

Receiver operating characteristic (ROC) curve analysis was performed to assess the discriminatory ability of baseline clinical and functional parameters for HRQoL at 12 months. The EORTC QLQ-C30 global health score was dichotomized at the median. Univariate ROC curves were calculated, and the area under the curve (AUC) was used as a measure of discrimination. ROC curves were visually smoothed for presentation.

Auditing, Data Monitoring, and Confidentiality

Semiannual audits and monitoring were conducted for data oversight at the study center. These procedures ensured data quality, protocol compliance, and proper conduct of the trial. Participant data were stored in a password-protected database accessible only to authorized study investigators. All personal information was handled in accordance with data protection regulations to ensure confidentiality. Within the framework of the study, patient insurance was in place to cover both the treatment and the entire follow-up period.

Ethical Considerations

The submitted manuscript meets all criteria of the ICMJE (International Committee of Medical Journal Editors) recommendations and COPE (Committee on Publication Ethics) guidelines. All procedures performed in this study were conducted in accordance with the ethical standards outlined in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards of the institutional and national research committees. Ethical approval was granted by the Ethics Committee of the Medical Faculty of the University of Regensburg (Ethics Committee reference: 11-101-0156; approval date: December 29, 2011) and subsequently by the ethics committees of the other participating study centers. The committees confirmed that there were no ethical or scientific concerns. Written informed consent was obtained from all participants prior to their inclusion in the study. All data were

XSL•FO
RenderX

anonymized, and no information enabling individual identification is presented in this manuscript.

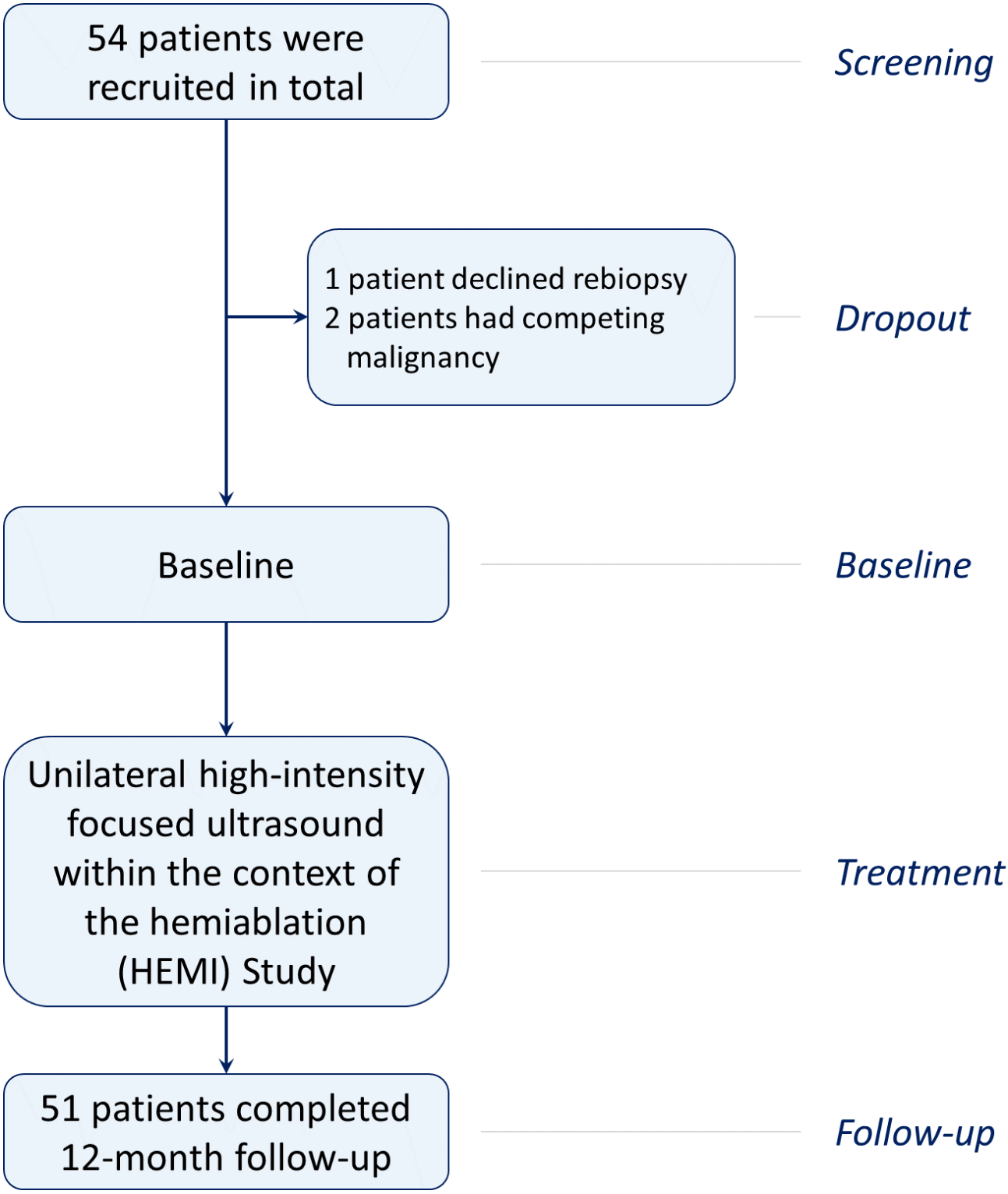
Results

Study Population and Follow-Up

Fifty-four patients were included in the analysis. Of these, 51 (94.4%) completed the planned 12-month assessment and were included in the final analysis. One patient declined the scheduled rebiopsy, and 2 patients discontinued follow-up due to secondary malignancies. Ultimately, data from 51 patients were available for the final analysis. The study selection process is illustrated in the CONSORT (Consolidated Standards of Reporting Trials) flow diagram (Figure 1). The mean age at treatment was 63.4 (SD 8.3) years, reflecting a typical PCa patient cohort, predominantly comprising a middle-aged population considered suitable for focal therapy. Patient characteristics are summarized in Table 2. Hemi-HIFU was associated with a significant but partial decline in erectile function, while urinary continence and lower urinary tract symptoms were largely preserved. These functional outcomes constituted the primary endpoints of the original study by Ganzer et al [38].

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials)-style flow diagram illustrating patient selection.



XSL•FO
RenderX

Table . Patients' characteristics.

	All patients (N=51)
Age (y), mean (SD)	63.4 (8.3)
PSA[a] (ng/mL), median (IQR)	6.1 (5.1 - 7.9)
Number of positive cores, median (IQR)	2 (1-2)
Prostate volume (mL), median (IQR)	32 (27 - 41)
Body mass index (kg/m^2), mean (SD)	28.6 (4.3)
ISUP 2[b], n (%)	8 (15.7)

[a]PSA: prostate-specific antigen.

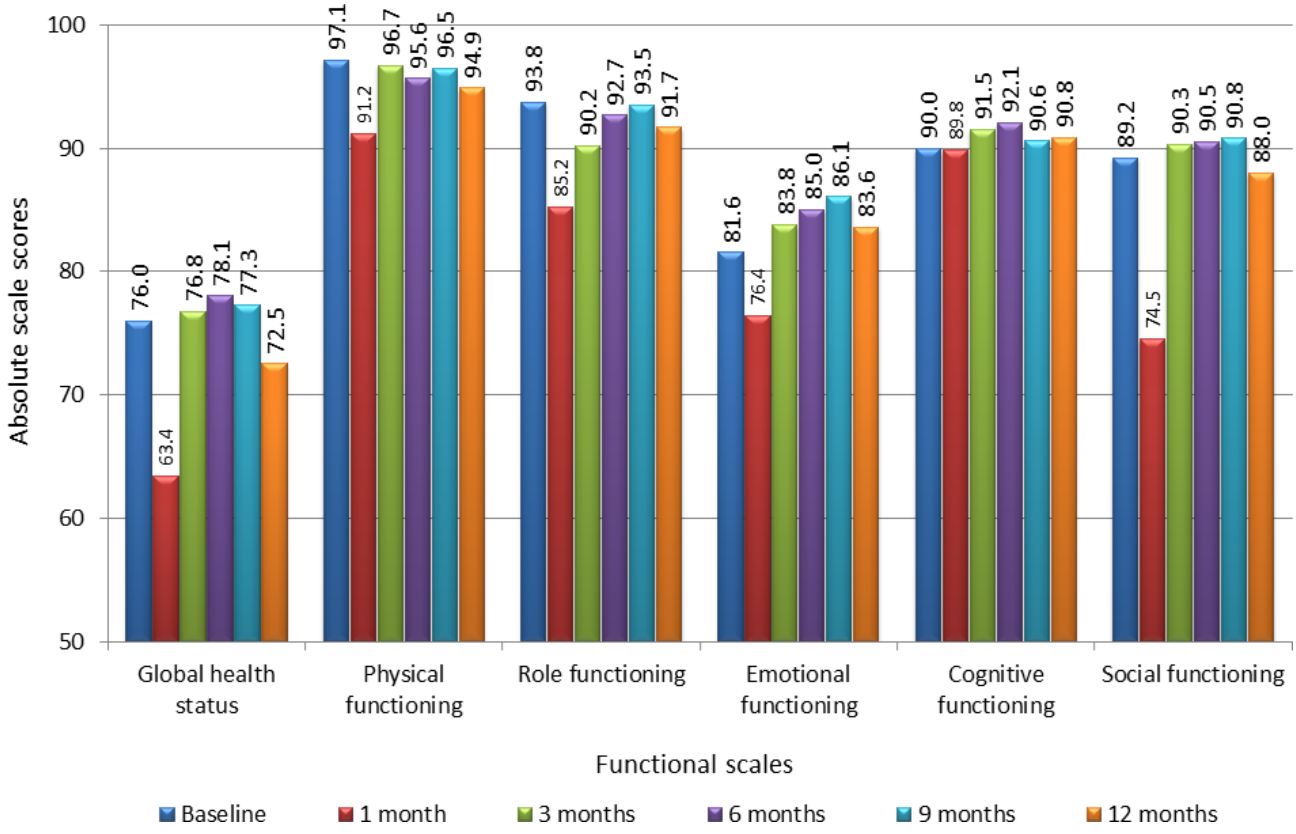
[b]ISUP 2: International Society of Urological Pathology Grade 2.

Health-Related Quality of Life (EORTC QLQ-C30)

Global Health Status and Quality of Life

At baseline, the mean global health status was 75 (SD 17.8), indicating only minor impairment. At 1 month posttreatment, a significant decline was observed (mean 63, SD 17.9; FDR-adjusted *P*=.002), corresponding to a clinically moderate deterioration. From 3 months onward, scores recovered to baseline or higher levels (3 mo: 78; 6 mo: 78; 9 mo: 77; and 12 mo: 74). Scores at 3 to 9 months were significantly higher than at 1 month (all FDR-adjusted *P*=.004, *P*=.003, and *P*=.004). By month 12, no significant difference from baseline was detected (FDR-adjusted *P*=.39), indicating full recovery of global quality of life (Figure 2; Table 3). The decline in global health status at 1 month corresponded to a medium effect size (Cohen *d*=0.58), and recovery between 1 and 3 months showed a comparable magnitude (*d*=0.61).

Figure 2. Overview of the absolute mean scale scores for global health status and overall quality of life and the functional parameters across the entire observation period. Higher scores reflect better functioning and a higher quality of life.



XSL•FO
RenderX

Table . A comprehensive overview of all scores and the corresponding *P* values.

Variable	Baseline	1 month	3 months	6 months	9 months	12 months
Global health status and quality of life, mean (SD)	74.6 (17.8)	63.4 (17.9)	77.5 (18.2)	78.0 (18.8)	77.5 (21.0)	73.9 (22.0)
P value	—[a]	.002[b]	.38	.31	.44	.39
Physical functioning, mean (SD)	97.1 (9.0)	91.2 (18.2)	96.7 (7.5)	95.6 (11.1)	96.5 (8.7)	94.4 (10.6)
P value	—	.004[b]	.38	.10	.45	.05
Role functioning, mean (SD)	93.8 (17.9)	85.2 (21.3)	90.2 (22.3)	92.7 (20.1)	93.5 (18.8)	91.7 (19.6)
P value	—	.04[b]	.43	.57	.72	.29
Emotional functioning, mean (SD)	81.6 (19.8)	76.4 (22.0)	83.8 (20.3)	85.0 (19.3)	86.1 (17.7)	83.6 (21.4)
P value	—	.20	.58	.38	.24	.63
Cognitive functioning, mean (SD)	90.0 (15.5)	89.8 (16.0)	91.5 (15.1)	92.1 (15.4)	90.6 (16.8)	90.8 (14.6)
P value	—	.76	.21	.67	.84	.82
Social functioning, mean (SD)	89.2 (17.0)	74.5 (25.0)	90.3 (16.9)	90.5 (18.1)	90.8 (19.0)	88.0 (20.2)
P value	—	<.001[b]	.74	.69	.60	.73
Fatigue, mean (SD)	13.7 (22.6)	28.5 (30.4)	13.4 (20.2)	12.9 (20.4)	12.5 (19.5)	14.7 (21.8)
P value	—	.007[b]	.94	.86	.78	.82
Nausea and vomiting, mean (SD)	2.1 (9.1)	0.9 (3.8)	0.3 (2.1)	1.4 (4.6)	0.6 (3.0)	1.3 (4.4)
P value	—	.66	.18	.66	.32	.71
Pain, mean (SD)	5.2 (17.7)	16.2 (17.7)	7.8 (18.5)	11.0 (22.7)	8.2 (19.3)	11.5 (20.0)
P value	—	.008[b]	.14	.03[b]	.053	.06
Dyspnea, mean (SD)	5.2 (16.3)	5.6 (14.9)	5.9 (16.5)	5.9 (15.3)	6.2 (15.6)	10.6 (15.2)
P value	—	.34	.08	.32	.53	.06
Insomnia, mean (SD)	17.7 (26.1)	29.6 (32.1)	17.5 (26.9)	14.7 (24.0)	15.5 (25.9)	18.0 (26.9)
P value	—	.04[b]	.65	.25	.11	.19
Appetite loss, mean (SD)	2.4 (8.7)	6.5 (19.0)	1.1 (6.0)	2.2 (8.3)	2.9 (11.3)	2.3 (8.2)
P value	—	.002[b]	.16	.56	.26	.89
Constipation, mean (SD)	4.2 (14.4)	13.5 (25.4)	4.8 (15.3)	4.4 (12.5)	4.0 (12.1)	6.1 (14.3)
P value	—	.16	.72	.59	.59	.83
Diarrhea, mean (SD)	5.4 (17.9)	8.1 (21.4)	6.4 (15.9)	6.6 (21.5)	8.0 (19.6)	4.2 (10.4)
P value	—	.07	.56	.30	.23	.33
Financial difficulties, mean (SD)	4.2 (14.7)	8.1 (20.1)	3.8 (10.6)	4.4 (13.1)	3.5 (12.4)	4.9 (11.9)
P value	—	.10	.17	.59	.32	.47
Depression, mean (SD)	8.8 (4.7)	10.8 (1.9)	10.7 (2.1)	10.6 (2.5)	10.4 (3.0)	10.1 (3.2)

XSL•FO
RenderX

Variable	Baseline	1 month	3 months	6 months	9 months	12 months
P value	—	.02[b]	.02[b]	.07	.23	.28
Anxiety, mean (SD)	6.5 (3.8)	7.8 (1.8)	7.5 (1.7)	6.8 (2.0)	7.0 (1.7)	6.8 (1.8)
P value	—	.30	.43	.70	.81	.80

[a]Not applicable.

[b]Statistically significant values.

Functional Scales

Physical, Role, Emotional, Cognitive, and Social Functioning

At baseline, physical, role, and cognitive functioning were high (means ≥90, SD 9 and 15), while emotional and social functioning were slightly lower (means 89, SD 19 and 82, SD 17, respectively). At 1 month, physical, role, and social functioning showed transient deterioration, with social functioning demonstrating the most pronounced decline (mean reduction of 14 points, from 89 to 75, SD 25; *P*=.004, *P*=.04, and *P*<.001, respectively). Emotional and cognitive functioning remained stable throughout. From 3 months onward, all functional domains improved. By month 12, no significant differences from baseline persisted across any functional scale, and clinically relevant impairments were rare (Figure 2; Table 3).

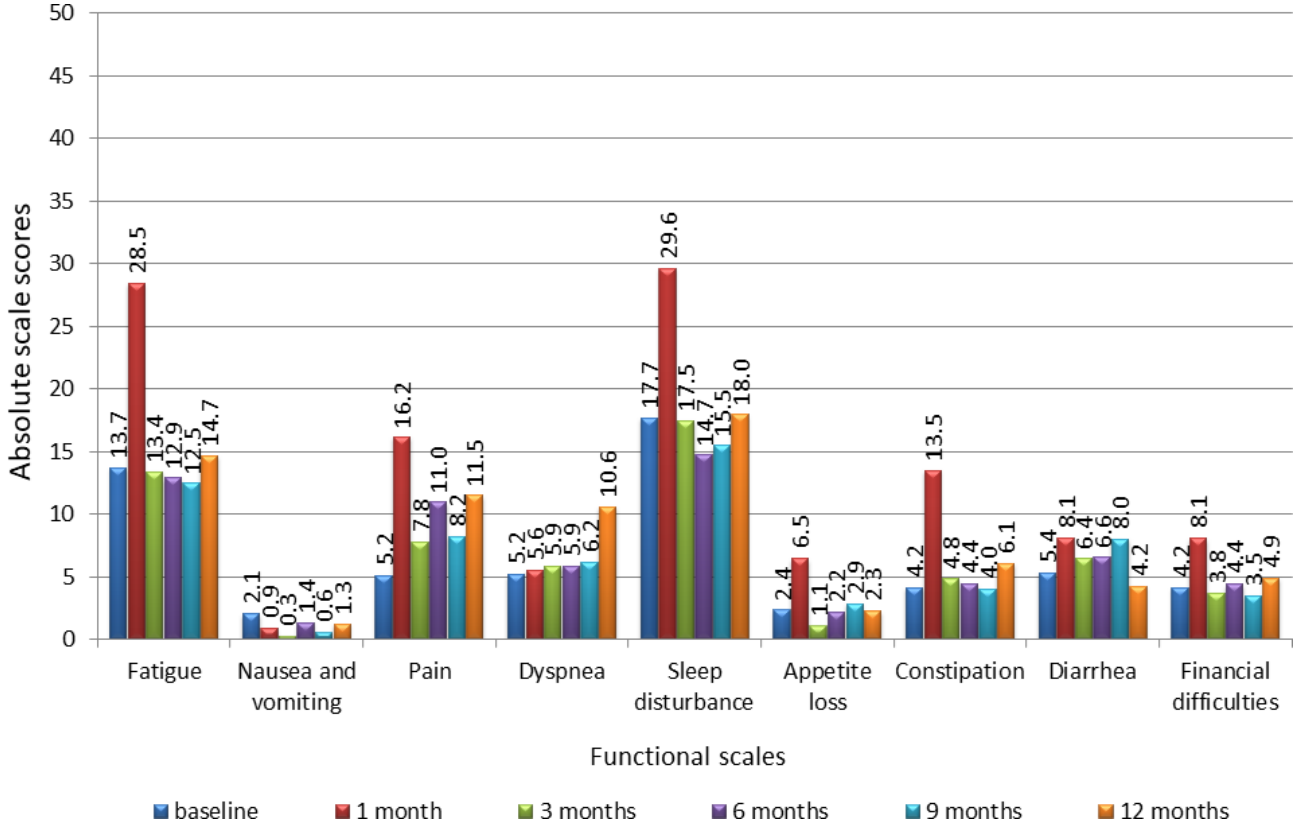
Symptom Scales—General Symptoms and Problem Burden

Fatigue, Nausea and Vomiting, Pain, Dyspnea, Insomnia, Appetite Loss, Constipation, Diarrhea, and Financial Difficulties

Baseline symptom burden was low across all domains. At 1 month, fatigue, pain, and insomnia increased transiently. Fatigue (mean increase of 15 points, SD 8.5; *P*=.007), insomnia (mean increase of 12 points, SD 6.9; *P*=.04), and pain (mean increase of 11 points, SD 4.4; *P*=.008) reached clinically moderate severity, while all other symptoms remained in the mild range.

From 3 months onward, symptoms normalized across all domains except for pain. In this domain, scores remained significantly different from baseline at 6 months (*P*=.03), but no differences compared to baseline were observed at 9 or 12 months. At no time point was the cumulative symptom burden suggestive of persistent treatment-related morbidity (Figure 3; Table 3).

Figure 3. Overview of the absolute mean scale scores for the symptom scales across the entire observation period. Higher scores indicate a greater symptom burden.

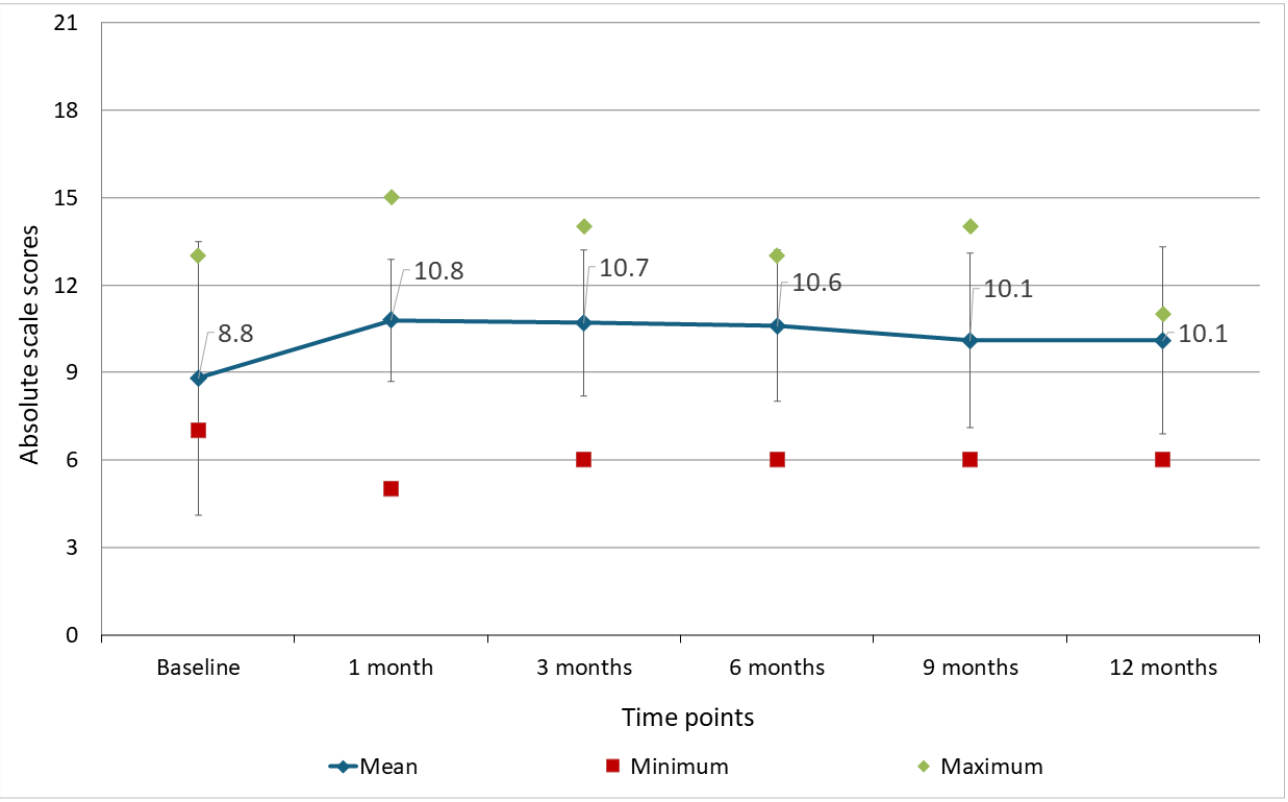


XSL•FO
RenderX

HADS—Depression

Mean depression scores remained below the clinical cutoff value (≥11) at all assessment points. Statistically significant changes were observed at 1 and 3 months (*P*=.02 and *P*=.02, respectively), indicating a transient but statistically detectable worsening in depressive symptom burden. From the 6-month assessment onward, no differences compared with baseline were detected (all FDR-adjusted *P*>.05 for pairwise comparisons). These findings indicate that clinically relevant depressive symptoms were not present at baseline, increased significantly at 1 and 3 months without exceeding the clinical threshold, and subsequently returned to baseline levels and remained stable throughout follow-up, suggesting recovery from a short-term postinterventional effect rather than persistent psychological deterioration attributable to HIFU therapy (Figure 4; Table 3).

Figure 4. Hospital Anxiety and Depression Scale depression scores over time.

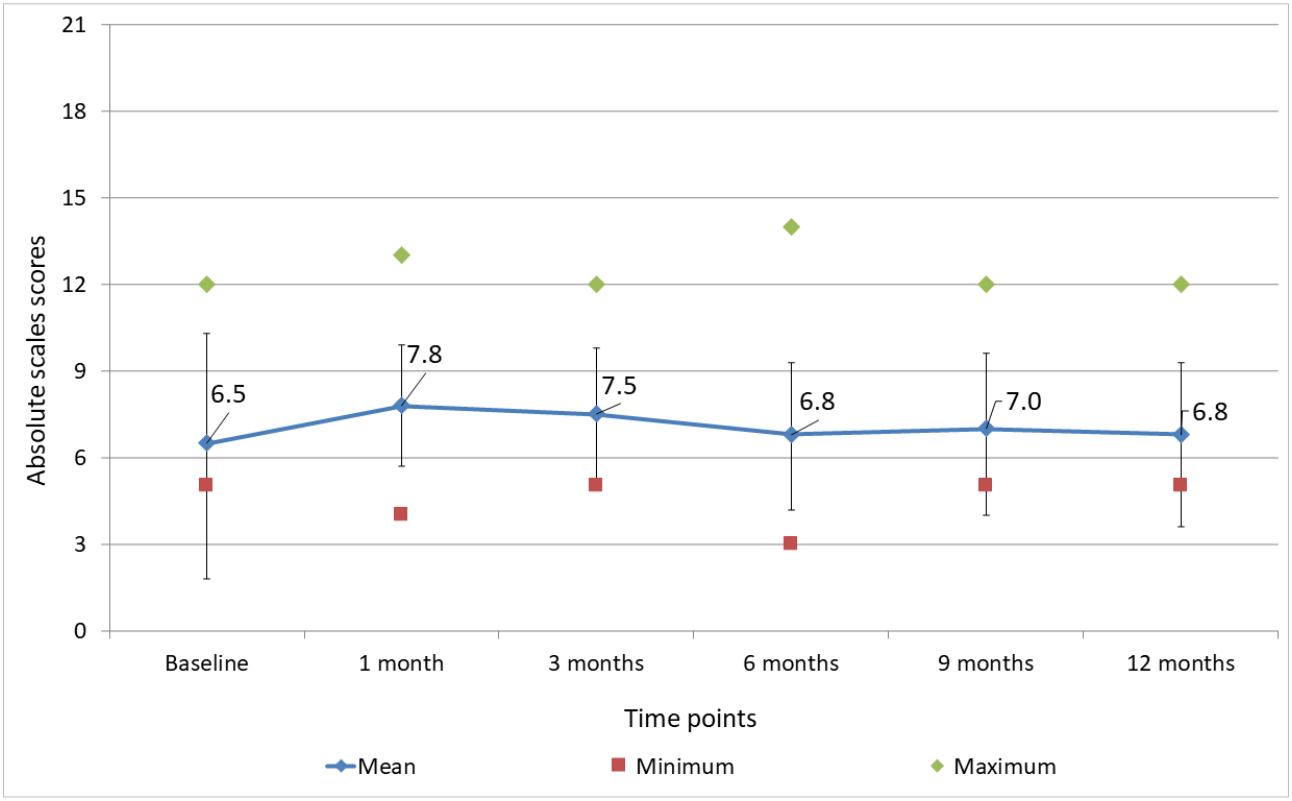


HADS-D—Anxiety

Mean anxiety scores remained below the threshold throughout, with no significant changes over time (FDR-adjusted *P*>.05). The proportion of patients within the nonpathological range increased slightly during follow-up (Figure 5; Table 3).

XSL•FO
RenderX

Figure 5. Hospital Anxiety and Depression Scale (HADS) anxiety scores over time.



An overview of all scale scores from both the functional and symptom domains of the EORTC QLQ-C30, as well as HADS values, including all baseline-referenced *P* values, is presented in Table 3.

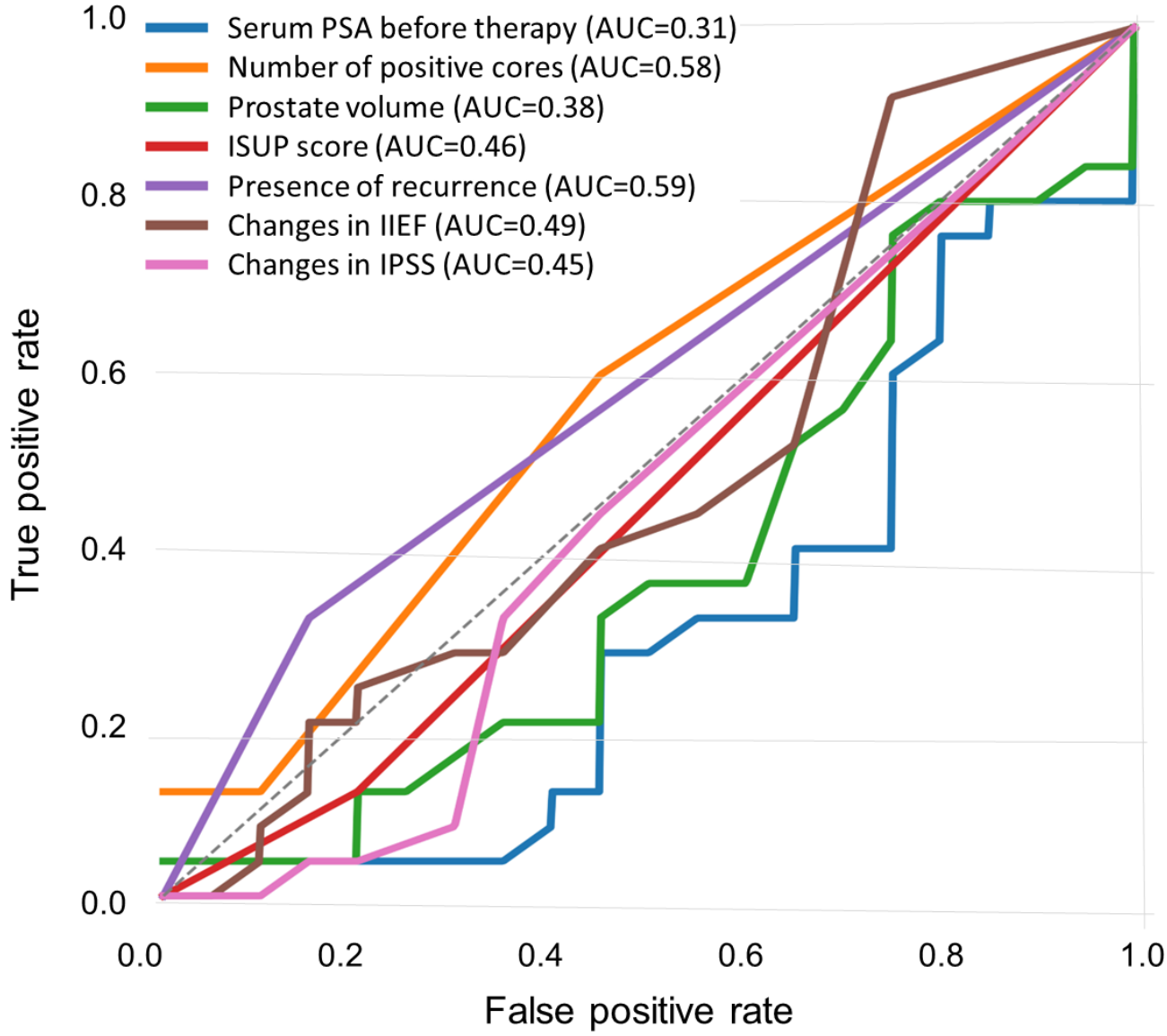
Exploratory subgroup analyses comparing patients with and without histologically confirmed recurrence at 12 months revealed no statistically significant differences in any EORTC QLQ-C30 domain or HADS subscale (all *P*>.05). No clinically relevant differences in anxiety or depression scores were observed between the groups.

Exploratory ROC Analysis of Predictors of 12-Month Global Quality of Life

To further explore the discriminatory ability of selected clinical and functional parameters for 12-month global HRQoL, univariate ROC curve analyses were performed using dichotomized EORTC QLQ-C30 global health scores at 12 months. Overall, discriminatory performance was low across all variables. The highest AUC values were observed for the presence of recurrence (AUC=0.59) and the number of positive biopsy cores (AUC=0.58), indicating only weak predictive ability. Serum prostate-specific antigen before therapy showed an AUC of 0.31, suggesting an inverse association with the outcome. Prostate volume (AUC=0.38), International Society of Urological Pathology score (AUC=0.46), changes in IIEF (International Index of Erectile Function, AUC=0.49), and changes in IPSS (International Prostate Symptom Score, AUC=0.45) also demonstrated poor discrimination. Overall, none of the analyzed parameters showed clinically meaningful predictive performance for 12-month global quality of life (Figure 6).

XSL•FO
RenderX

Figure 6. Receiver operating characteristic curves illustrating the discriminatory performance of clinical parameters for predicting European Organization for Research and Treatment of Cancer Quality of Life Questionnaire—Core 30 (EORTC QLQ-C30) scores at 12 months. Univariate models were constructed for prostate volume, serum prostate-specific antigen (PSA), number of positive cores, International Society of Urological Pathology (ISUP) score, presence of recurrence, changes in IPSS (International Prostate Symptom Score), and changes in IIEF (International Index of Erectile Function). The dashed diagonal line indicates no discriminative ability (AUC=0.5). AUC: area under the curve.



Comparisons of the Individual Dimensions With the General Population

The present analysis compares EORTC QLQ-C30 scores of the study cohort with normative data from healthy male populations. Reference values were obtained from the German general population dataset published by Schwarz and Hinz [42] and from the Norwegian normative dataset reported by Hjermstad et al [43], the latter lacking SD values. Both datasets were originally generated as part of the EORTC validation program to assess the reliability and validity of the QLQ-C30. In the following figures, the study cohort is contrasted with these reference populations (Figure 7; Figure 8; Figure 9). Differences between the cohort and the German normative sample were examined using independent 2-sided *t* tests. To distinguish statistical differences from clinically meaningful differences, the interpretation framework proposed by Osoba et al [44] was applied, defining changes of 5 to 10 points as small, 10 to 20 points as moderate, and >20 points as large clinical differences.

Global Health Status and Quality of Life

After 1 month, patients reported significantly lower global health compared with the German reference population (P=.01), corresponding to a clinically moderate difference according to the Osoba classification. At 6 months, scores exceeded normative values (P=.02), although the magnitude of the difference remained clinically small. No clinically relevant deviation was present at 12 months (Figure 7; Table 4).

XSL•FO
RenderX

Figure 7. Global health status of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire—Core 30 (EORTC QLQ-C30) in comparison with the German male normative population. The asterisk (*) indicates values that differ significantly.

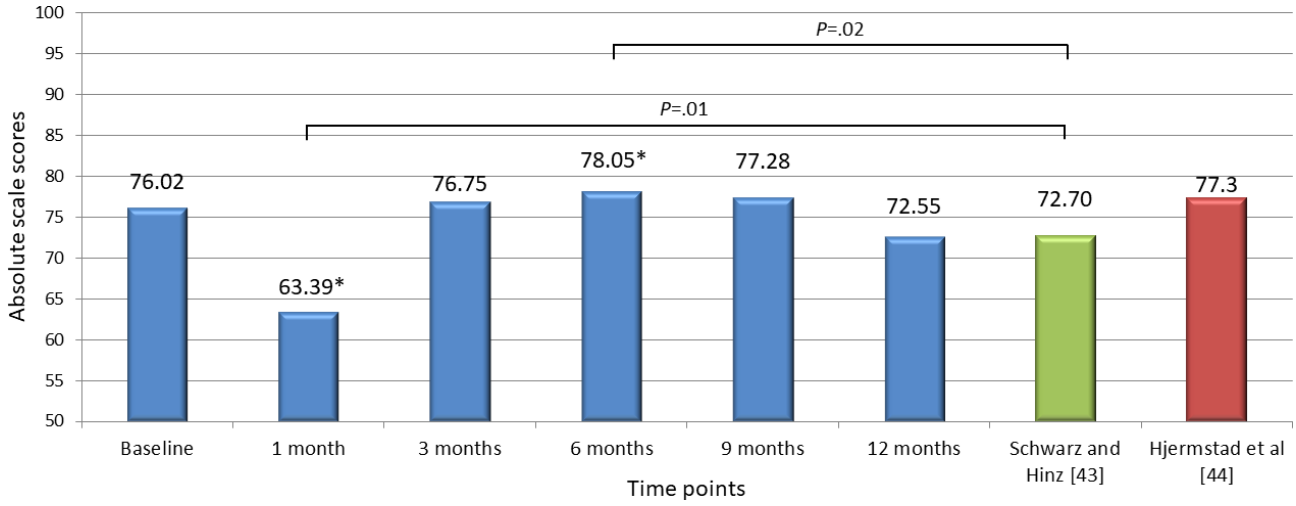


Table . Overview of functional and symptom scales showing clinical significance according to the Osoba criteria at each assessment time point.

Scale and time point	Scale score of the patient cohort	Reference group score	Difference between groups	Clinical relevance
Baseline				
Pain	5.15	13.0	7.85	Low
1 month				
Global health status	63.39	72.7	9.31	Low
Physical functioning	97.11	92.0	5.11	Low
Emotional functioning	76.39	81.8	5.41	Low
Social functioning	74.52	92.0	17.48	Moderate
Fatigue	28.47	14.0	14.47	Moderate
Insomnia	29.63	13.0	16.63	Moderate
Constipation	13.51	2.5	11.01	Moderate
Diarrhea	8.9	2.5	6.4	Low
3 months				
Pain	7.76	13.0	5.24	Low
6 months				
Global health status	78.05	72.7	5.35	Low
9 months				
Diarrhea	8.04	2.5	5.54	Low
12 months				
None	—[a]	—	—	—

[a]Not applicable.

Functional Subscales

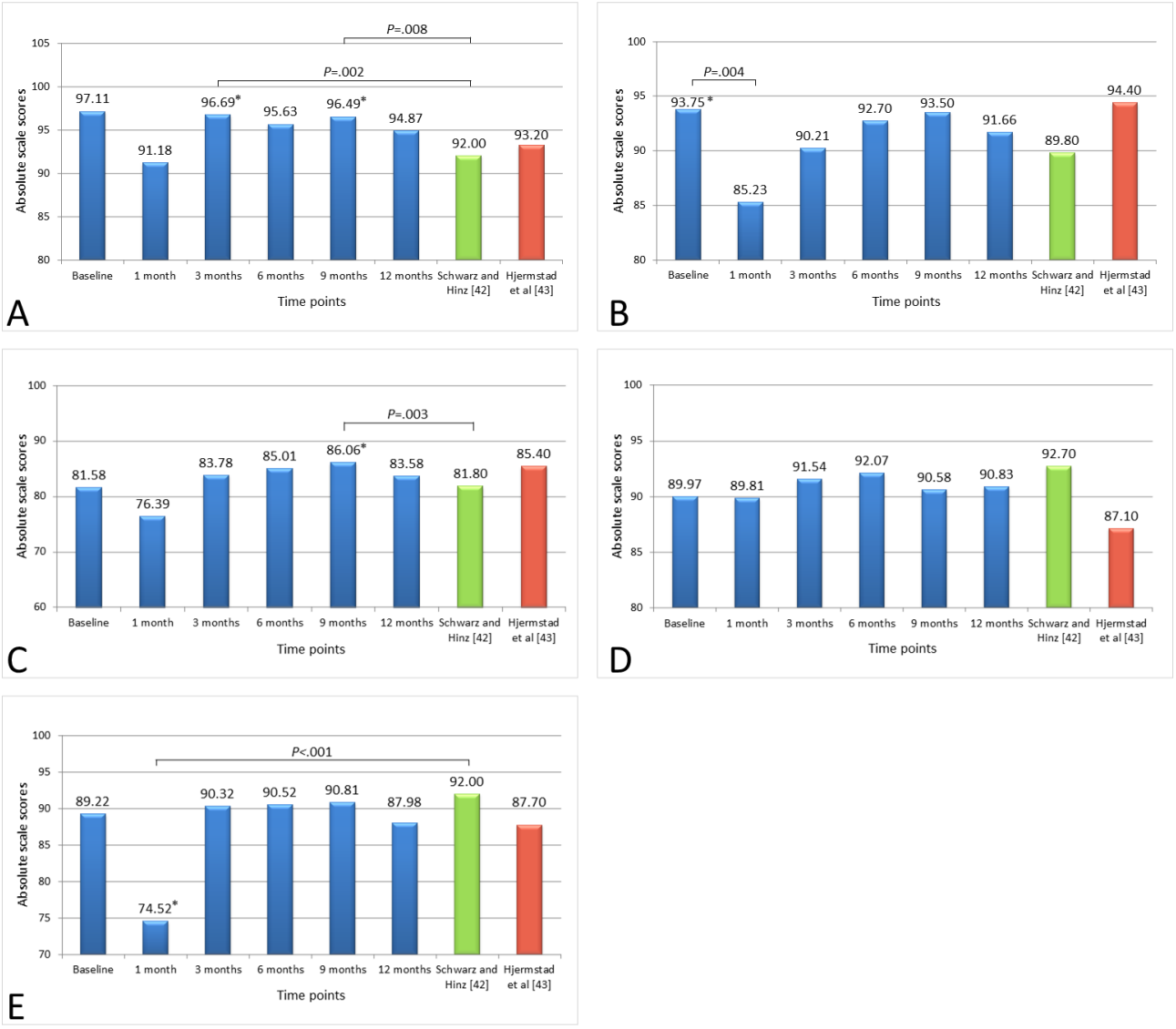
Physical, Role, Emotional, Cognitive, and Social Functioning

Physical functioning differed significantly from normative values at 3, 6, and 9 months (all FDR-adjusted P=.002, P=.03, P=.008, respectively); however, all differences remained below the threshold for clinical relevance. A small, clinically relevant advantage of the study cohort was observed at baseline. Role functioning showed a significant decline at 1 month (P=.04), while no difference from baseline was observed at 3 months, and values subsequently remained comparable to the normative population across all assessments with no significant differences. Emotional functioning showed a statistically significant improvement at 9 months compared with population data (P=.03) but the difference was clinically negligible. Cognitive functioning showed no deviations from reference values at any time point.

XSL•FO
RenderX

Social functioning demonstrated the most pronounced deviation. At 1 month, a moderate clinical deterioration of 15 points relative to the normative population was observed (*P*<.001). From 3 months onward, scores normalized completely (Figure 8; Table 4).

Figure 8. Functional scales of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire—Core 30 (EORTC QLQ-C30): physical (A), role (B), emotional (C), cognitive (D), and social functioning (E) in comparison with the German male normative population. The asterisk (*) indicates values that differ significantly.



Symptom Scales—General Symptoms and Problem Burden

Fatigue, Nausea and Vomiting, Pain, Dyspnea, Insomnia, Appetite Loss, Constipation, Diarrhea, and Financial Difficulties

Baseline symptom burden was low across all domains compared with the German normative population. At 1 month posttreatment, fatigue increased significantly (items 10, 12, and 18; *P*<.001), corresponding to a clinically moderate deterioration (mean difference of 14 points, from 14 to 28, SD 23). Fatigue scores returned to baseline by 3 months and remained stable thereafter.

Nausea and vomiting (items 14 and 15) differed significantly from normative values at 3 and 9 months in favor of the study cohort (*P*=.004 and *P*=.03, respectively); however, absolute symptom levels remained negligible and were not clinically meaningful (maximum mean score of 2.08, SD 9.1). Pain (items 9 and 19) was significantly lower in the study cohort at baseline and 3 months compared with normative data (*P*=.002 and *P*=.31, respectively), corresponding to small-to-moderate clinical effects.

No significant differences were observed for dyspnea at any time point (all *P*>.05), with all deviations falling below clinical relevance thresholds. Insomnia (item 11) demonstrated a clinically moderate short-term deterioration at 1 month (mean difference of 17, SD 32 points; FDR-adjusted *P*=.005), which fully resolved by month 3.

Appetite loss (item 13) differed significantly from normative values only at 3 months (*P*=.002), but the magnitude of the difference was clinically negligible (<5 points). Constipation (item 16) was significantly elevated at 1 and 12 months

XSL•FO
RenderX

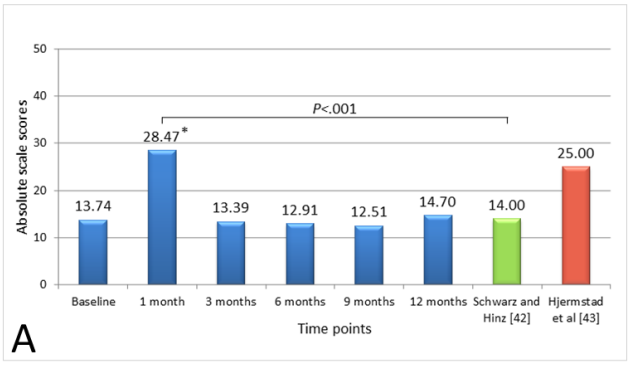
(FDR-adjusted P=.01 and P=.048), representing the only symptom with persistent clinically relevant impairment. Diarrhea (item 17) increased at 9 months (P=.02), with small but clinically meaningful differences.

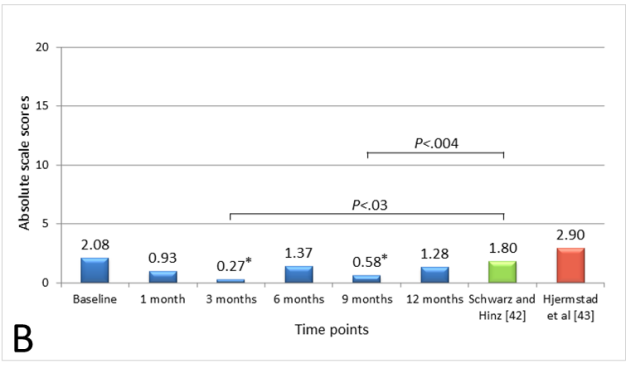
No differences were detected for financial difficulties (item 28). Overall, symptom trajectories were characterized by transient early postoperative worsening—most notably for fatigue and insomnia—followed by recovery by month 3 across most domains. Persistent morbidity was limited to constipation, supporting a favorable tolerability profile of focal HIFU therapy (Figure 9; Table 4).

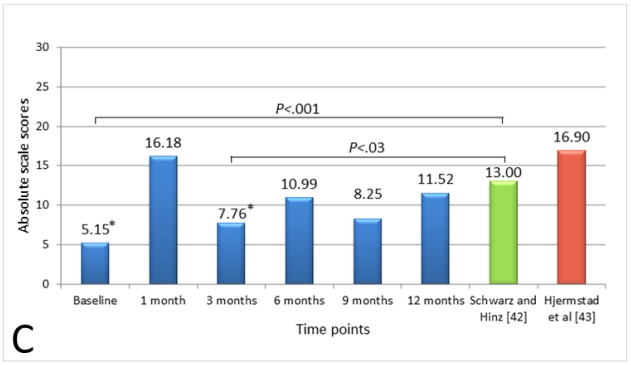
XSL•FO
RenderX

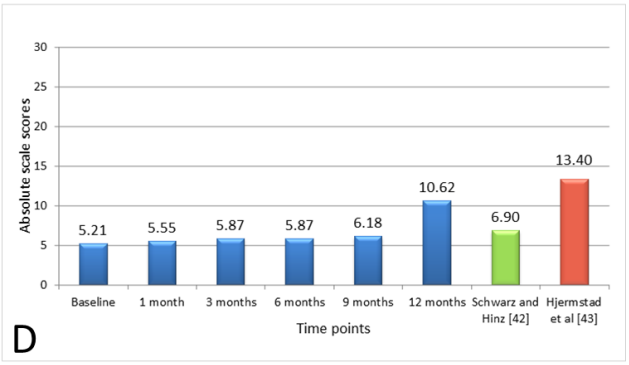
Figure 9. Symptom scales of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire—Core 30 (EORTC QLQ-C30): fatigue (A), nausea and vomiting (B), pain (C), dyspnea (D), insomnia (E), appetite loss (F), constipation (G), diarrhea (H), and financial difficulties (I) in comparison with male normative population. The asterisk (*) indicates values that differ significantly.

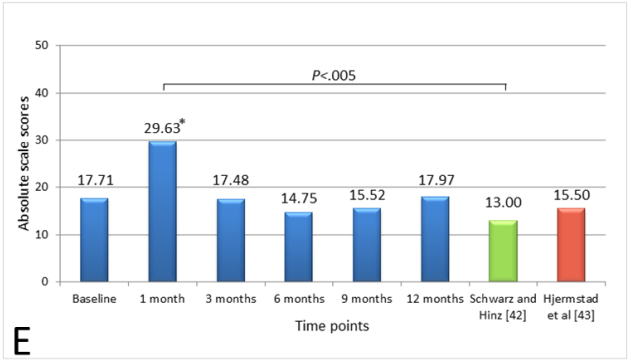
XSL•FO
RenderX

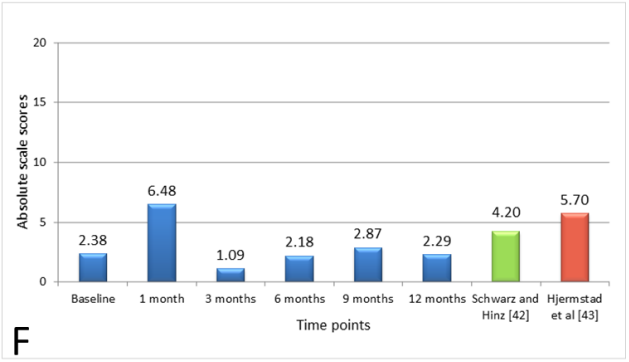


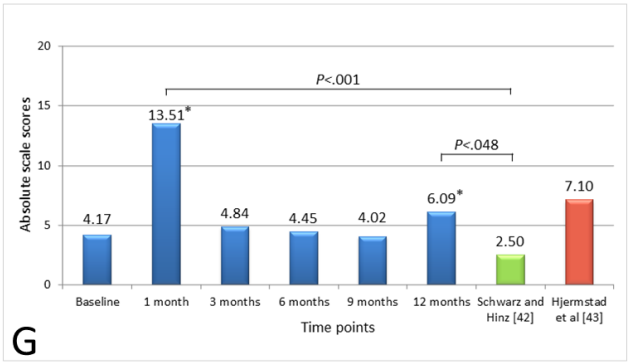


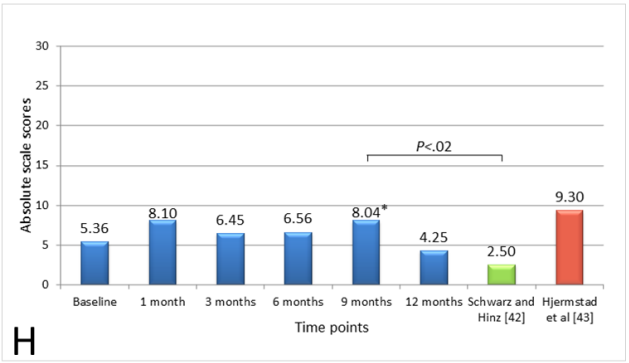


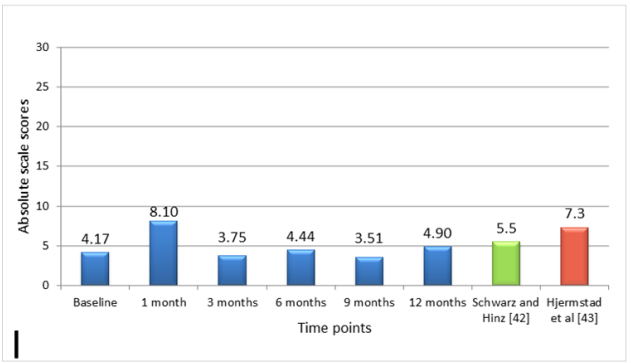












XSL•FO
RenderX

Discussion

Summary of Main Findings

This study evaluated psychological distress and HRQoL in patients undergoing focal HIFU hemiablation for localized PCa over a 12-month follow-up period. The primary objective was to characterize the temporal evolution of patient-reported outcomes from baseline through long-term follow-up. Overall, focal therapy was associated with low levels of psychological distress and preserved HRQoL. A statistically significant but transient decline in HRQoL and selected symptom domains was observed at 1 month, followed by recovery to baseline levels by 3 months. At 12 months, no clinically relevant impairments persisted. Depression scores showed a temporary increase without exceeding clinical thresholds, while anxiety levels remained stable throughout follow-up.

Contextualization

Across all psychological instruments and subscales, focal therapy demonstrated a favorable profile with respect to quality-of-life preservation. Mean scores consistently indicated a low overall psychological burden, with most patients reporting minimal or no impairment. Notably, no statistically significant differences were observed between baseline and 12-month values in any psychological domain, supporting the assumption that focal therapy is associated with stable long-term psychosocial outcomes [48,49].

The transient decline in HRQoL and psychological well-being observed at 1 month is most likely attributable to procedure-related factors, including peri-interventional stress, temporary functional impairment, and recovery from treatment. By 3 months, baseline values were restored and remained stable thereafter. Application of Osoba criteria suggests that clinically meaningful deterioration was largely confined to the early postoperative phase. Similar transient deteriorations have been described following focal ablative therapies and are attributed to perioperative stressors, including hospitalization, anesthesia, and temporary functional impairment [49]. Recent longitudinal studies using PCa-specific outcome measures (eg, IPSS, IIEF-5, EPIC [Expanded Prostate Cancer Index Composite]) similarly demonstrate a transient decline in urinary and sexual function at 1 month following focal HIFU, with recovery to baseline levels from 3 months onward, supporting the functional safety of this approach [50]. Overall, the early postoperative symptom increase appears primarily related to procedural stress rather than oncological anxiety.

Functional and oncological outcomes of the underlying HEMI study have been reported in detail elsewhere [38]. In brief, focal HIFU hemiablation was associated with preserved urinary continence and largely maintained erectile function, with no increase in incontinence rates and potency retained in the majority of preoperatively potent patients. These findings provide important clinical context for the present analysis and support the interpretation of preserved quality-of-life outcomes.

Comparison with normative population data revealed significant deviations primarily at the early postoperative time point, whereas later assessments showed convergence with population-based reference values. This pattern indicates that focal therapy does not result in sustained impairment relative to the general population. Similar recovery trajectories have been reported in previous studies of focal HIFU and other focal treatment modalities [51-53].

Psychological outcomes in this cohort were characterized by low baseline distress and stable trajectories over time. The predominance of patients with early-stage PCa may partly account for the low prevalence of clinically relevant psychological impairment. The temporary increase in depressive symptoms likely reflects a situational stress response rather than a persistent affective disorder. Importantly, no clinically relevant anxiety or depression was observed throughout follow-up. These findings are consistent with previous reports demonstrating that anxiety and depressive symptoms in patients undergoing focal therapy remain close to population norms and are associated with low levels of psychological burden and decision regret [49,54].

When compared with other established treatment modalities for localized PCa, these findings further support the favorable quality-of-life profile of focal therapy. RP and RT have been consistently associated with higher rates of urinary incontinence and sexual dysfunction, which are key determinants of long-term HRQoL [4-7,37]. In contrast, AS is generally associated with preserved functional outcomes but may be accompanied by an increased psychological burden due to ongoing uncertainty regarding disease progression. In this context, focal HIFU may represent a balanced approach, combining the low psychological burden observed in AS with the benefits of active oncological treatment while maintaining favorable functional outcomes.

Exploratory analyses did not demonstrate an association between short-term oncological outcomes and patient-reported quality of life or psychological distress. No statistically significant or clinically relevant differences in EORTC QLQ-C30 domains or HADS scores were observed between patients with and without histologically confirmed recurrence at 12 months. These findings suggest that short-term oncological outcomes, including the presence of residual or recurrent disease, did not translate into measurable impairments in HRQoL or increased psychological burden within the observation period. This may be explained by the low overall tumor burden and favorable disease characteristics of the study population, as well as structured follow-up and patient counseling within the study setting.

However, these results should be interpreted with caution. The number of recurrence events was limited, and the study was not powered to detect subtle subgroup differences. In addition, the relatively short follow-up period may not capture delayed psychological effects related to oncological uncertainty or disease progression.

Additional factors may include patient selection and individual coping mechanisms. Patients with localized, low-risk disease generally exhibit a lower baseline psychological burden than those with advanced cancer. Furthermore, the active decision to undergo focal therapy may enhance perceived control and facilitate psychological adaptation.

XSL•FO
RenderX

The deliberate decision to undergo focal HIFU may represent an additional psychological factor influencing patient-reported outcomes. By selecting a minimally invasive, organ-preserving treatment approach, patients may experience a greater sense of control and involvement in their care. Such perceptions have been associated with reduced feelings of helplessness and improved psychological adjustment. In this context, previously reported low levels of decision regret following focal therapy further support this interpretation [54].

Conclusion and Clinical Implications

Focal HIFU hemiablation is associated with preserved psychological well-being and sustained HRQoL in patients with localized PCa. Although transient postoperative distress may occur, recovery is rapid, and long-term outcomes remain comparable to population norms.

Taken together, these findings suggest that psychological stabilization following focal therapy is achievable in the context of preserved coping capacity and patient empowerment. Both confrontation with the diagnosis and postinterventional adjustment appear manageable within this treatment framework, supporting focal therapy as a patient-centered approach that balances disease control with quality-of-life preservation.

From a clinical perspective, routine psycho-oncological follow-up may be most relevant for vulnerable subgroups, whereas broader interventions may not be necessary in psychologically stable patients. Future research should focus on extended longitudinal follow-up, the use of standardized patient-reported outcome measures, integration of systematic psycho-oncological screening, identification of vulnerability profiles, and the parallel assessment of oncological and psychological endpoints.

Limitations

Several limitations should be considered. The predominance of patients eligible for AS reflects historical indications for focal therapy at the time of study initiation (2013 - 2016) and may limit the applicability of the findings to contemporary patient populations with clinically significant disease. Since study initiation, indications for focal therapy have evolved toward carefully selected patients with clinically significant localized disease, including favorable intermediate-risk PCa. At the same time, the use of a low-risk, AS-eligible population allowed for an isolated assessment of treatment-related quality-of-life outcomes with minimal confounding by advanced disease burden and contributed to the early evidence base supporting the broader application of focal therapy. Detailed data on the number of screened patients and reasons for nonparticipation were not systematically recorded due to the study design and shared decision-making process.

First, the sample size was relatively small (N=54), and attrition over the 12-month follow-up reduced statistical power, particularly for subgroup analyses and small effect sizes. The exclusion of patients due to protocol-defined dropouts, including competing health events and incomplete follow-up, may introduce a potential risk of selection bias in the longitudinal analyses. In addition, the highly selected study cohort limits generalizability to the broader PCa population.

Second, patient recruitment occurred between 2013 and 2016, and subsequent advances in imaging, biopsy techniques, and treatment guidelines (eg, European Association of Urology, 2024) may limit direct applicability to contemporary clinical practice. Third, the study lacked a randomized control group, relying instead on external normative reference populations, which restricts causal inference and increases susceptibility to unmeasured confounding.

Fourth, although validated instruments were used, patient-reported outcomes are inherently subjective and may be influenced by response and recall bias. In addition, the use of mailed questionnaires may have introduced additional selection bias. Fifth, clinical relevance was interpreted using the Osoba classification, which was originally developed in breast and lung cancer populations, and its applicability to PCa remains uncertain.

Finally, despite the use of mpMRI and targeted biopsies, focal hemiablation carries a risk of undertreatment due to unrecognized multifocal disease. Detailed oncological outcomes were not assessed, and follow-up was limited to 12 months, precluding the evaluation of late functional or psychological effects.

Acknowledgments

The authors would like to express their sincere gratitude to their colleagues and collaborators for their invaluable support and contributions throughout this study. Their expertise and dedication have been essential in advancing this research. The authors also extend their thanks to the University Hospital Leipzig for providing the opportunity and resources to conduct this study. Their commitment to fostering a collaborative research environment has been instrumental in the success of this work.

The authors declare that generative artificial intelligence (GAI) tools were not used for conceptualization, literature review, study design, methodology development, software development, data management, data analysis, ethics review, or supervision.

GAI (ChatGPT 5.3) was used solely for language-related support during the preparation of the manuscript, including refinement, correction, editing, and formatting to improve clarity and readability. This usage was comparable to that of spell-checking and grammar-checking tools.

All intellectual content, study design, data analysis, interpretation of results, and primary manuscript writing were performed independently by the authors. The authors take full responsibility for the content of the manuscript. Responsibility for the final manuscript lies entirely with the authors.

XSL•FO
RenderX

Funding

The authors declare that no financial support was received for this work.

Data Availability

The datasets generated and analyzed during this study will be made available upon reasonable request to the corresponding author.

Authors' Contributions

Conceptualization: TF, CH
Data acquisition: TF, CH, RG
Data analysis and interpretation: TF, CH
Figure design: TF
Statistical analysis: CH
Supervision: TF, JUS
Writing – original draft: TF, CH
Writing – review & editing: RG, SH, JUS

Conflicts of Interest

None declared.

References

1. Wang L, Lu B, He M, Wang Y, Wang Z, Du L. Prostate cancer incidence and mortality: global status and temporal trends in 89 countries from 2000 to 2019. Front Public Health 2022;10:811044. [doi: 10.3389/fpubh.2022.811044] [Medline: 35252092]
2. Cancer stat facts: prostate cancer. National Cancer Institute SEER. URL: https://seer.cancer.gov/statfacts/html/prost.html [accessed 2025-12-01]
3. European Cancer Information System. URL: https://ecis.jrc.ec.europa.eu [accessed 2025-12-01]
4. Resnick MJ, Koyama T, Fan KH, et al. Long-term functional outcomes after treatment for localized prostate cancer. N Engl J Med 2013 Jan 31;368(5):436-445. [doi: 10.1056/NEJMoa1209978] [Medline: 23363497]
5. Sanda MG, Dunn RL, Michalski J, et al. Quality of life and satisfaction with outcome among prostate-cancer survivors. N Engl J Med 2008 Mar 20;358(12):1250-1261. [doi: 10.1056/NEJMoa074311] [Medline: 18354103]
6. Seputra KP, Daryanto B, Budaya TN, et al. Quality of life in prostate cancer patients following radical prostatectomy: a systematic review and meta-analysis. Asian Pac J Cancer Prev 2025 Sep 1;26(9):3185-3192. [doi: 10.31557/APJCP.2025.26.9.3185] [Medline: 40952272]
7. Liu Y, Wang Y, Xiao X, et al. Postoperative functional complications and quality of life following robot-assisted prostatectomy and radiotherapy in localized prostate cancer: evidence from a systematic review and meta-analysis. J Robot Surg 2025 Jun 22;19(1):314. [doi: 10.1007/s11701-025-02496-x] [Medline: 40544416]
8. Draisma G, Boer R, Otto SJ, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. J Natl Cancer Inst 2003 Jun 18;95(12):868-878. [doi: 10.1093/jnci/95.12.868] [Medline: 12813170]
9. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. N Engl J Med 2009 Mar 26;360(13):1320-1328. [doi: 10.1056/NEJMoa0810084] [Medline: 19297566]
10. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. N Engl J Med 2005 May 12;352(19):1977-1984. [doi: 10.1056/NEJMoa043739] [Medline: 15888698]
11. Klotz L. Active surveillance for prostate cancer: for whom? J Clin Oncol 2005 Nov 10;23(32):8165-8169. [doi: 10.1200/JCO.2005.03.3134] [Medline: 16278468]
12. Latini DM, Hart SL, Knight SJ, et al. The relationship between anxiety and time to treatment for patients with prostate cancer on surveillance. J Urol 2007 Sep;178(3 Pt 1):826-831. [doi: 10.1016/j.juro.2007.05.039] [Medline: 17632144]
13. Fujita K, Landis P, McNeil BK, Pavlovich CP. Serial prostate biopsies are associated with an increased risk of erectile dysfunction in men with prostate cancer on active surveillance. J Urol 2009 Dec;182(6):2664-2669. [doi: 10.1016/j.juro.2009.08.044] [Medline: 19836757]
14. Cooperberg MR, Broering JM, Kantoff PW, Carroll PR. Contemporary trends in low risk prostate cancer: risk assessment and treatment. J Urol 2007 Sep;178(3 Pt 2):S14-S19. [doi: 10.1016/j.juro.2007.03.135] [Medline: 17644125]
15. Valerio M, Cerantola Y, Eggener SE, et al. New and established technology in focal ablation of the prostate: a systematic review. Eur Urol 2017 Jan;71(1):17-34. [doi: 10.1016/j.eururo.2016.08.044] [Medline: 27595377]
16. Ramsay CR, Adewuyi TE, Gray J, et al. Ablative therapy for people with localised prostate cancer: a systematic review and economic evaluation. Health Technol Assess 2015 Jul;19(49):1-490. [doi: 10.3310/hta19490] [Medline: 26140518]

XSL•FO
RenderX

17. Azzouzi AR, Vincendeau S, Barret E, et al. Padeliporfin vascular-targeted photodynamic therapy versus active surveillance in men with low-risk prostate cancer (CLIN1001 PCM301): an open-label, phase 3, randomised controlled trial. Lancet Oncol 2017 Feb;18(2):181-191. [doi: 10.1016/S1470-2045(16)30661-1] [Medline: 28007457]
18. Blana A, Murat FJ, Walter B, et al. First analysis of the long-term results with transrectal HIFU in patients with localised prostate cancer. Eur Urol 2008 Jun;53(6):1194-1201. [doi: 10.1016/j.eururo.2007.10.062] [Medline: 17997026]
19. Blana A, Hierl J, Rogenhofer S, et al. Factors predicting for formation of bladder outlet obstruction after high-intensity focused ultrasound in treatment of localized prostate cancer. Urology 2008 May;71(5):863-867. [doi: 10.1016/j.urology.2007.12.076] [Medline: 18313119]
20. Ganzer R, Fritsche HM, Brandtner A, et al. Fourteen-year oncological and functional outcomes of high-intensity focused ultrasound in localized prostate cancer. BJU Int 2013 Aug;112(3):322-329. [doi: 10.1111/j.1464-410X.2012.11715.x] [Medline: 23356910]
21. Rischmann P, Gelet A, Riche B, et al. Focal high intensity focused ultrasound of unilateral localized prostate cancer: a prospective multicentric hemiablation study of 111 patients. Eur Urol 2017 Feb;71(2):267-273. [doi: 10.1016/j.eururo.2016.09.039] [Medline: 27720531]
22. van Velthoven R, Aoun F, Marcelis Q, et al. A prospective clinical trial of HIFU hemiablation for clinically localized prostate cancer. Prostate Cancer Prostatic Dis 2016 Mar;19(1):79-83. [doi: 10.1038/pcan.2015.55] [Medline: 26597660]
23. Feijoo ERC, Sivaraman A, Barret E, et al. Focal high-intensity focused ultrasound targeted hemiablation for unilateral prostate cancer: a prospective evaluation of oncologic and functional outcomes. Eur Urol 2016 Feb;69(2):214-220. [doi: 10.1016/j.eururo.2015.06.018] [Medline: 26164416]
24. Ahmed HU, Freeman A, Kirkham A, et al. Focal therapy for localized prostate cancer: a phase I/II trial. J Urol 2011 Apr;185(4):1246-1254. [doi: 10.1016/j.juro.2010.11.079] [Medline: 21334018]
25. Fütterer JJ, Briganti A, De Visschere P, et al. Can clinically significant prostate cancer be detected with multiparametric magnetic resonance imaging? A systematic review of the literature. Eur Urol 2015 Dec;68(6):1045-1053. [doi: 10.1016/j.eururo.2015.01.013] [Medline: 25656808]
26. Siddiqui MM, Rais-Bahrami S, Turkbey B, et al. Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. JAMA 2015 Jan 27;313(4):390-397. [doi: 10.1001/jama.2014.17942] [Medline: 25626035]
27. Hoeks CMA, Schouten MG, Bomers JGR, et al. Three-Tesla magnetic resonance-guided prostate biopsy in men with increased prostate-specific antigen and repeated, negative, random, systematic, transrectal ultrasound biopsies: detection of clinically significant prostate cancers. Eur Urol 2012 Nov;62(5):902-909. [doi: 10.1016/j.eururo.2012.01.047] [Medline: 22325447]
28. Pinto PA, Chung PH, Rastinehad AR, et al. Magnetic resonance imaging/ultrasound fusion guided prostate biopsy improves cancer detection following transrectal ultrasound biopsy and correlates with multiparametric magnetic resonance imaging. J Urol 2011 Oct;186(4):1281-1285. [doi: 10.1016/j.juro.2011.05.078] [Medline: 21849184]
29. Dickinson L, Ahmed HU, Allen C, et al. Magnetic resonance imaging for the detection, localisation, and characterisation of prostate cancer: recommendations from a European consensus meeting. Eur Urol 2011 Apr;59(4):477-494. [doi: 10.1016/j.eururo.2010.12.009] [Medline: 21195536]
30. Glatzer W, Zapf W. Lebensqualität in Der Bundesrepublik: Objektive Lebensbedingungen Und Subjektives Wohlbefinden [Book in German]: Campus Verlag; 1984.
31. Buddeberg C. Psychosoziale Medizin, 3rd edition: Springer; 2003.
32. The WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med 1995 Nov;41(10):1403-1409. [doi: 10.1016/0277-9536(95)00112-K]
33. Dinkel A, Jahnen M. Patient-reported outcomes—the role of the patient's subjective perspective for research and clinical care. Urologie 2024 Sep;63(9):886-892. [doi: 10.1007/s00120-024-02405-4] [Medline: 39110186]
34. Haraldstad K, Wahl A, Andenæs R, et al. A systematic review of quality of life research in medicine and health sciences. Qual Life Res 2019 Oct;28(10):2641-2650. [doi: 10.1007/s11136-019-02214-9] [Medline: 31187410]
35. van der Heijden TGW, de Ligt KM, Hubel NJ, et al. Exploring the role of health-related quality of life measures in predictive modelling for oncology: a systematic review. Qual Life Res 2025 Feb;34(2):305-323. [doi: 10.1007/s11136-024-03820-y] [Medline: 39652111]
36. Tannock IF, Pe ML, Booth CM, et al. Importance of responder criteria for reporting health-related quality-of-life data in clinical trials for advanced cancer: recommendations of Common Sense Oncology and the European Organisation for Research and Treatment of Cancer. Lancet Oncol 2025 Sep;26(9):e499-e507. [doi: 10.1016/S1470-2045(25)00288-8] [Medline: 40907528]
37. Penson DF, Litwin MS, Aaronson NK. Health related quality of life in men with prostate cancer. J Urol 2003 May;169(5):1653-1661. [doi: 10.1097/01.ju.0000061964.49961.55] [Medline: 12686803]
38. Ganzer R, Hadaschik B, Pahernik S, et al. Prospective multicenter phase II study on focal therapy (hemiablation) of the prostate with high intensity focused ultrasound. J Urol 2018 Apr;199(4):983-989. [doi: 10.1016/j.juro.2017.10.033] [Medline: 29107031]

XSL•FO
RenderX

39. Cornford P, van den Bergh RCN, Briers E, et al. EAU-EANM-ESTRO-ESUR-ISUP-SIOG guidelines on prostate cancer-2024 update. Part I: screening, diagnosis, and local treatment with curative intent. Eur Urol 2024 Aug;86(2):148-163. [doi: 10.1016/j.eururo.2024.03.027] [Medline: 38614820]
40. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 3;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
41. The EORTC QLQ-C30 scoring manual. : European Organisation for Research and Treatment of Cancer; 2001 URL: https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf [accessed 2026-05-27]
42. Schwarz R, Hinz A. Reference data for the quality of life questionnaire EORTC QLQ-C30 in the general German population. Eur J Cancer 2001 Jul;37(11):1345-1351. [doi: 10.1016/s0959-8049(00)00447-0] [Medline: 11435063]
43. Hjermstad MJ, Fayers PM, Bjordal K, Kaasa S. Health-related quality of life in the general Norwegian population assessed by the European Organization for Research and Treatment of Cancer Core Quality-of-Life Questionnaire: the QLQ-C30 (+ 3). J Clin Oncol 1998 Mar;16(3):1188-1196. [doi: 10.1200/JCO.1998.16.3.1188] [Medline: 9508207]
44. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998 Jan;16(1):139-144. [doi: 10.1200/JCO.1998.16.1.139] [Medline: 9440735]
45. Snaith RP. The Hospital Anxiety and Depression Scale. Health Qual Life Outcomes 2003 Aug 1;1:29. [doi: 10.1186/1477-7525-1-29] [Medline: 12914662]
46. Herrmann C, Buss U. Vorstellung und Validierung einer deutschen Version der Hospital Anxiety and Depression Scale (HAD-Skala): ein Fragebogen zur Erfassung des psychischen Befindens bei Patienten mit körperlichen Beschwerden. Diagnostica 1994;40:143-154.
47. Herrmann-Lingen C, Buss U, Snaith RP. Hospital Anxiety and Depression Scale – Deutsche Version [German Version] Deutsche Adaptation Der Hospital Anxiety and Depression Scale (HADS) von R P Snaith Und A S Zigmond, 3rd edition: Verlag Hans Huber; 2011.
48. Hopstaken JS, Bomers JGR, Sedelaar MJP, Valerio M, Fütterer JJ, Rovers MM. An updated systematic review on focal therapy in localized prostate cancer: what has changed over the past 5 years? Eur Urol 2022 Jan;81(1):5-33. [doi: 10.1016/j.eururo.2021.08.005] [Medline: 34489140]
49. Pellegrino A, Cirulli GO, Mazzone E, et al. Focal therapy for prostate cancer: what is really needed to move from investigational to valid therapeutic alternative? A narrative review. Ann Transl Med 2022 Jul;10(13):755. [doi: 10.21037/atm-22-50] [Medline: 35957731]
50. Shoji S, Naruse J, Ohno S, et al. Focal therapy using high-intensity focused ultrasound with intraoperative prostate compression for patients with localized prostate cancer: a multi-center prospective study with 7 year experience. Prostate Cancer Prostatic Dis 2025 Sep;28(3):782-788. [doi: 10.1038/s41391-024-00921-0] [Medline: 39580599]
51. Lodeizen O, de Bruin M, Eggener S, et al. Ablation energies for focal treatment of prostate cancer. World J Urol 2019 Mar;37(3):409-418. [doi: 10.1007/s00345-018-2364-x] [Medline: 29943219]
52. Albisinni S, Mélot C, Aoun F, et al. Focal treatment for unilateral prostate cancer using high-intensity focal ultrasound: a comprehensive study of pooled data. J Endourol 2018 Sep 12;32(9):797-804. [doi: 10.1089/end.2018.0130] [Medline: 29790383]
53. Scheltema MJ, Chang JI, van den Bos W, et al. Impact on genitourinary function and quality of life following focal irreversible electroporation of different prostate segments. Diagn Interv Radiol 2018 Sep;24(5):268-275. [doi: 10.5152/dir.2018.17374] [Medline: 30211680]
54. Flegar L, Baunacke M, Buerk BT, et al. Decision regret and quality of life after focal therapy with vascular-targeted photodynamic therapy (TOOKAD®) for localized prostate cancer. Urol Int 2022;106(9):903-908. [doi: 10.1159/000520084] [Medline: 34814157]

Abbreviations

AS: active surveillance
AUC: area under the curve
CONSORT: Consolidated Standards of Reporting Trials
EORTC: European Organization for Research and Treatment of Cancer
FDR: false discovery rate
HADS: Hospital Anxiety and Depression Scale
HIFU: high-intensity focused ultrasound
HRQoL: health-related quality of life
mpMRI: multiparametric magnetic resonance imaging
PCa: prostate cancer
PIRADS: Prostate Imaging Reporting and Data System
QLQ-C30: Quality of Life Questionnaire—Core 30
ROC: receiver operating characteristic

XSL•FO
RenderX

RP: radical prostatectomy
RT: radiotherapy

Edited by M Balcarras; submitted 26.Jan.2026; peer-reviewed by C Lei, S Shoji; revised version received 21.Apr.2026; accepted 14.May.2026; published 15.Jun.2026.

Please cite as:
Franz T, Hartung C, Ganzer R, Holze S, Stolzenburg JU
Patient-Reported Outcomes on Quality of Life and Psychological Distress After Focal High-Intensity Focused Ultrasound for Prostate Cancer: Prospective Multicenter Study
JMIR Cancer 2026;12:e91984
URL: https://cancer.jmir.org/2026/1/e91984
doi:10.2196/91984

© Toni Franz, Cornelia Hartung, Roman Ganzer, Sigrun Holze, Jens-Uwe Stolzenburg. Originally published in JMIR Cancer (https://cancer.jmir.org), 15.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Reinforcement Learning–Based Digital Therapeutic Intervention for Postprostatectomy Incontinence: Development and Pilot Feasibility Study

Fan Fan[1,2], MSN; Hao Huang[1], PhD; Jingwen Yan[3], MSN; Chao-Yue Xu[1], MSN; Xiuhua Wu[1], BSc; Chunmei Zhou[1], BSc; Dandan Wen[1], BSc; Hai Huang[1], PhD; Ho Cheung Li[2], PhD; Yihong Qiu[3], BSc

[1]Department of Urology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, Guangdong, China

[2]The Nethersole School of Nursing, Faculty of Medicine, Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Department of Nursing, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, Guangdong, China

Corresponding Author:
Yihong Qiu, BSc
Department of Nursing
Sun Yat-sen Memorial Hospital
Sun Yat-sen University
No. 107, Yanjiang West Road, Yuexiu District
Guangzhou, Guangdong, 510120
China
Phone: 86 1 392 519 8885
Email: qiuyihong@126.com

Abstract

Background: Postprostatectomy incontinence (PPI) is a common complication after robot-assisted radical prostatectomy and significantly impairs patients' quality of life. Although behavioral interventions such as pelvic floor muscle training and bladder diaries are evidence-based, their effectiveness is often limited by poor adherence and lack of personalization.

Objective: This study aimed to develop and evaluate a reinforcement learning (RL)–driven clinical behavioral intervention-supporting system (CBISs) for adaptive, personalized rehabilitation in patients with PPI.

Methods: The study comprised 2 sequential stages. First, the CBISs was developed through (1) construction of a medical record database from a prospective cohort of PPI patients using standardized 3-day bladder diaries, (2) design of functional modules and user interfaces based on clinical rehabilitation needs, and (3) development of an RL model using XGBoost (extreme gradient boosting) and Bayesian optimization to generate individualized training plans. Second, a separate cohort of 16 patients participated in a single-arm, pre-post pilot study to evaluate feasibility and preliminary outcome trends over a 3-month intervention period, with assessments based on bladder diary parameters and system usage metrics.

Results: The CBISs successfully implemented an adaptive, closed-loop behavioral rehabilitation framework that dynamically tailored training recommendations according to individual voiding patterns, fluid intake behaviors, and adherence signals. Feasibility outcomes were favorable, with high system engagement observed throughout the intervention (mean usage frequency 5.2, SD 1.1 times per day). In exploratory pre-post analyses (n=16), consistent directional improvements were observed across multiple outcomes. Mean daytime urinary frequency decreased from 5.74 (SD 1.21) episodes per day to 4.69 (SD 1.08) episodes per day, while median nighttime urinary frequency declined from 1.8 (IQR 1.6-2.2) episodes per night to 1.0 (IQR 1.0-1.6) episodes per night. Median incontinence episodes were reduced from 7.0 (IQR 6.0-11.0) episodes per day to 4.0 (IQR 2.0-6.0) episodes per day. Objective urine leakage measured by the 1-hour pad test decreased from a median of 8.5 (IQR 4.0-19.0) g to 3.5 (IQR 2.0-9.0) g. Patient-reported symptom burden, assessed using the International Consultation on Incontinence Questionnaire–Short Form (ICIQ-UI SF), showed a median reduction from 14.0 (IQR 12.0-20.0) points to 9.0 (IQR 6.0-16.0) points. Although several within-participant changes were statistically detectable, effect magnitudes varied across individuals. Given the single-arm design, small sample size, and lack of a control group, findings are presented as exploratory and hypothesis-generating rather than confirmatory of clinical efficacy.

Conclusions: The CBISs represents the first RL-powered digital therapeutic system for PPI, enabling adaptive, evidence-based behavioral optimization. By addressing limitations of static rehabilitation protocols and declining adherence, it offers a scalable approach for personalized PPI management. Future multicenter trials are needed to confirm its clinical effectiveness.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e83375) doi:10.2196/83375

KEYWORDS

behavioral rehabilitation; clinical decision support; digital therapeutics; mobile health; pelvic floor muscle training; prostate cancer; quality of life; reinforcement learning; self-management; telemedicine; urinary incontinence

Introduction

Prostate cancer (PCa) is one of the most prevalent malignant tumors in men and currently represents the second leading cause of cancer-related mortality worldwide [1]. It is estimated that by 2025, approximately 313,780 new cases of PCa will be diagnosed in the United States, accounting for 15.4% of all newly diagnosed cancers [2]. Over the past decade, robot-assisted radical prostatectomy has been increasingly recognized as a primary treatment for localized PCa [3]. Owing to the removal of the prostate, seminal vesicles, and surrounding tissues, as well as the reconstruction of the urinary tract during surgery, patients frequently experience varying degrees of postoperative bladder dysfunction. Among these complications, postprostatectomy incontinence (PPI) is the most common functional sequela, severely affecting patients' daily activities, sexual function, and overall quality of life (QoL) [4-6]. The reported incidence of PPI varies widely, ranging from approximately 4.2% to 87%, largely due to differences in definitions of urinary incontinence, follow-up duration, and assessment methodologies [7]. PPI symptoms typically persist for 3-12 months and, in some cases, may extend for up to 10 years, rendering urinary incontinence a major long-term concern during postoperative recovery [8,9]. Consequently, improving PPI and related QoL outcomes remains a central objective in the comprehensive management of survivors of PCa [10].

Behavioral therapies, including pelvic floor muscle training (PFMT), bladder retraining, urge suppression techniques, and lifestyle modification, constitute the cornerstone of conservative management for urinary incontinence. High-quality evidence from systematic reviews and meta-analyses, including Cochrane reviews, has consistently demonstrated that behavioral therapy–based interventions can reduce incontinence episodes and improve patient-reported QoL across diverse populations [11,12]. These interventions are widely recommended as first-line or adjunctive strategies in clinical practice, reflecting their favorable safety profiles and broad applicability. Nevertheless, the effectiveness of behavioral rehabilitation in real-world settings is frequently compromised by inadequate adherence, limited access to professional supervision, and insufficient patient understanding of correct training techniques.

Advancing from this foundation, digital therapeutics (DTx) have emerged as a promising modality for delivering structured rehabilitation programs remotely, thereby improving accessibility and continuity of care for patients following postprostatectomy. DTx are interventions delivered via digital platforms [13], which have the potential to mitigate the challenges posed by chronic diseases, such as enhancing patient adherence and improving patient self-management [14]. By enabling real-time monitoring, personalized feedback, and flexible scheduling, digital platforms may help address many of the practical limitations inherent in traditional rehabilitation models [15]. However, many existing digital interventions for urinary incontinence remain largely static, relying on predefined rules or uniform training protocols that do not adequately account for individual variability in symptoms, adherence patterns, or recovery trajectories.

In this context, reinforcement learning (RL) represents a promising methodological framework for advancing personalization within digital therapeutic systems. RL is a machine learning approach whereby intelligent agents iteratively optimize decision-making models through cycles of action, feedback, and strategy adjustment [16]. Unlike traditional supervised learning methods, RL is particularly well suited to sequential decision-making problems where the effects of an action may be delayed, outcomes are uncertain, and the optimal strategy depends on the evolving state of the individual. In digital health applications, RL has been used to personalize "just-in-time" interventions—for example, adapting the timing, content, or intensity of behavioral prompts—based on each patient's history, context, and response pattern, thereby improving engagement and clinical outcomes compared with static or rule-based approaches [17]. Moreover, RL models are capable of processing sparse or noisy reward signals arising from heterogeneous patient behaviors and complex intervention-response relationships, making them especially suitable for individualized rehabilitation scenarios [18]. Applying it to DTx may facilitate individualized patient rehabilitation training.

In this study, to address the need for personalized rehabilitation, a Clinical Behavioral Intervention-Supporting System (CBISs) was developed. Using RL, the CBISs provides adaptive, individualized training guidance to patients with urinary incontinence following robot-assisted radical prostatectomy, aiming to accelerate functional recovery and improve QoL through data-driven support.

Therefore, the objectives of this study were to develop and preliminarily evaluate an RL–based CBISs integrated into a digital therapeutic platform to provide individualized pelvic floor and behavioral training for men with PPI after robot-assisted radical prostatectomy. The study was designed to examine whether an RL-driven CBISs could (1) demonstrate feasibility and acceptability, with favorable trends in urinary continence recovery and patient-reported QoL; and (2) enhance adherence to and engagement with the prescribed training over time.

Methods

Overview

The study was structured in two sequential phases, (1) the development of the CBISs, and (2) a feasibility and preliminary efficacy pilot test of the CBISs conducted with a separate cohort of 16 patients. The development of the CBISs consisted of three

XSL•FO
RenderX

main stages, (1) construction of a clinical behavioral database, (2) design of system functionalities and user interfaces based on rehabilitation needs, and (3) development and integration of RL models to support individualized training.

Recruitment of Participants

All patients were recruited during the follow-up visit of the Outpatient Department of Urology at Sun Yat-sen Memorial Hospital, Sun Yat-sen University. Patients were eligible if they were aged 18-75 years, had a confirmed diagnosis of urinary incontinence following radical prostatectomy, and could accurately maintain a voiding diary (recording fluid intake, urination, and leakage). Exclusion criteria comprised serious neurological diseases, concurrent conditions such as urinary tract infection or bladder stones, significant voiding dysfunction, and an inability to comply with the study protocol. During the CBISs construction phase, we recruited a total of 150 patients to participate.

Database Construction

This section details the complete workflow for building the analytical database, including the statistical procedures and tests for data curation and feature engineering, which collectively form the essential foundation for the subsequent development of the prediction model. Bladder diaries were prospectively collected from patients diagnosed with PPI at the urology department of a tertiary care academic hospital. To ensure data robustness, a minimum diary period of 3 days was required based on International Continence Society standards for reliable symptom capture [19].

Sample size determination addressed two dimensions. At the patient level, sample size considerations focused on ensuring sufficient coverage of behavioral variability for model development rather than hypothesis testing. A cohort of 150 patients was deemed adequate to support feature engineering, model training, and internal validation while minimizing the risk of overfitting. At the behavioral event level, more than 10,000 voiding events were targeted to ensure model generalizability, accounting for expected intrapatient variability (mean, 8 voids per day per patient). Sample size considerations therefore accounted not only for the number of patients but also for the volume of behavioral data, with a 3-day behavioral log per patient ensuring minimum model generalizability. Ethical approval for the collection of data required for system construction was obtained from the Medical Ethics Committee of Sun Yat-sen Memorial Hospital, Sun Yat-sen University (approval number: SYSKY -2023-925-01). Data collection was conducted between March 2022 and December 2023.

Determination of Key Features and System Functionalities

The research team conducted a systematic literature review and held 2 rounds of expert meetings to develop a checklist for the rehabilitation needs of patients with PPI. Based on this, researchers and software engineers collaborated to define the functional module division and design the human–computer interaction interface. The CBISs comprises 3 portals: patient portal, a health care team portal, and an administrator portal. The patient portal contains 4 modules: rehabilitation training, bladder diary, incontinence care, and assessment tool. The health care team portal displays the patient's bladder diary details and training completion status.

Reinforcement Learning Model Construction

The research team developed the conceptual plan and proposed the functional requirements of the model; algorithm engineers then constructed and validated the training model, which was subsequently embedded into the CBISs to complete the design and development. Individualized rehabilitation training models for patients with PPI include urination training, fluid intake training, and PFMT. To realize these features, the algorithm engineers used a preconstructed database and performed feature screening and dimensionality reduction through an exhaustive feature engineering step.

Data Preprocessing and Quality Control

Prior to model training, all raw bladder diary records and behavioral logs underwent a structured preprocessing pipeline. Missing values in continuous variables (eg, voided volume and urination interval) were imputed using patient-level mean strategy, while discrete variables (eg, fluid type and urgency score) adopted mode imputation. Abnormal values were detected through IQR filtering and time-series consistency checks. All timestamps were normalized to standard 24-hour format, and temporal alignment of drinking and voiding events was ensured using sliding-window techniques with a 60-minute threshold. Records from patients with <80% diary completion rate or >20% anomalous entries were excluded from analysis to guarantee robust data quality and integrity.

Feature Selection and Model Training Strategy

From the initial 29 behavioral and physiological features, recursive feature elimination and XGBoost's (extreme gradient boosting's) intrinsic gain-based importance ranking were jointly applied to refine input dimensions. Highly collinear features (Pearson r>0.9) were removed to prevent multicollinearity. Model development was performed using a stratified 70/15/15 train/validation/test split, preserving the distribution of leakage severity and adherence levels. Bayesian optimization was used to fine-tune hyperparameters (eg, learning rate, max_depth, and subsample) with the goal of maximizing prediction accuracy and minimizing overfitting. Model performance was evaluated using a composite metric, including area under the receiver operating characteristic curve (AUC), F_1-score, and mean absolute error (MAE). A 5-fold cross-validation strategy was used during the training phase to ensure model generalizability and robustness.

Reinforcement Learning Framework and Deployment

The customized offline RL architecture is adopted in the CBISs to support adaptive behavioral rehabilitation while clinical safety is maintained. The RL component is designed to function as a decision-support layer rather than as an autonomous optimizer of clinical outcomes, reflecting the exploratory nature of this pilot study.

Clinical state representations are constructed from multiday behavioral snapshots encompassing key metrics such as PFMT adherence, voiding intervals, fluid intake ratios, and

XSL•FO
RenderX

patient-reported outcomes. These inputs closely correspond to the information typically evaluated by clinicians when adjusting rehabilitation strategies, thereby ensuring that the RL state definition remains clinically interpretable and grounded in routine practice.

At each decision point t, the system observes a behavioral state s_t and selects an action a_t from a predefined discrete action space. Actions correspond to clinically admissible behavioral interventions; for example, when a patient demonstrates stable adherence to the prescribed training program but reports repeated episodes of urinary leakage, the system may recommend a modest extension of PFMT duration (eg, +15 minutes) while maintaining the current training intensity. Subsequent adjustments to PFMT intensity are then guided by changes in the frequency of reported leakage, allowing the intervention to progress in a gradual and clinically responsive manner. All actions are constrained within predefined safety ranges to avoid abrupt or excessive changes.

Reward Function Formulation

The reward function is designed to reflect clinically meaningful behavioral changes rather than to directly optimize long-term clinical outcomes. At each decision point t, the reward r_t is computed as a weighted composite function:



where A_t represents short-term adherence to prescribed training tasks, C_t reflects task completion and user engagement, Δ_t captures the magnitude of abrupt behavioral changes between consecutive recommendations, and P_t denotes penalty terms applied when predefined safety thresholds are approached (eg, overly frequent voiding prompts or rapid escalation of training schedules). All components are normalized to bounded ranges prior to aggregation to ensure numerical stability and interpretability.

In practical clinical terms, this formulation supports gradual and responsive rehabilitation adjustments. For instance, when a patient reports an increase in urinary leakage episodes, the system may recommend extending PFMT duration by 15 minutes, representing a moderate increase in task demand. If the patient demonstrates high adherence and engagement following this adjustment and a subsequent reduction in leakage frequency is observed, penalty terms decrease and the revised training plan is maintained. Conversely, if task completion or engagement declines and leakage frequency fails to improve, penalty values increase, indicating that the adjustment may be poorly tolerated. In such cases, the system may revise the recommendation by reducing the training extension (eg, to an additional 5 minutes) and continue monitoring adherence, engagement, and leakage trends before further modification. This iterative process mirrors routine clinical reasoning, emphasizing tolerability, patient response, and safety rather than rigid optimization.

Q Value Estimation and Policy Learning

Based on the defined reward, action values are estimated using standard Q-learning principles. The Q value represents the expected cumulative future reward of selecting action a_t under state and is updated offline according to:



where α is the learning rate and γ is the discount factor. Q values are learned entirely from historical behavioral data using offline batch updates, without real-time exploration in patients.

Conceptually, Q values estimate how beneficial a specific behavioral adjustment is expected to be over time. Actions associated with higher Q values are more likely to be recommended, as they have historically led to stable adherence and acceptable behavioral patterns.

To handle variability in patient engagement and avoid overreliance on historical patterns, an ε-greedy action selection strategy is used. In most situations, the system selects the action with the highest estimated Q value; however, with a small probability ε, alternative actions are explored. In simple terms, the model usually follows what has worked best before, but occasionally tests other reasonable options to prevent overly rigid behavior, particularly for patients with inconsistent adherence.

System Usability and Technical Validation

Usability testing was conducted with a pilot group of 30 patients drawn from the database construction cohort, recruited from September to November 2023, who participated in a 4-week trial from December 2023 to January 2024. Task completion rates, interaction time, feedback scores, and daily user feedback during usage were systematically recorded. The mean training plan navigation time was 56 seconds (SD 12 seconds), with a >92% completion rate of daily bladder diaries. The CBISs demonstrated full functional compatibility with both Android (v8.0+) and iOS (v13+) platforms. Backend performance stress testing indicated stable response under 1000 concurrent users with <500 ms latency. The final release version underwent security penetration testing to ensure robustness against common cybersecurity threats, such as injection attacks, data breaches, and unauthorized data access attempts.

Integration of Human–Artificial Intelligence Shared Decision-Making

To reinforce clinical safety and ensure alignment with real-world care standards, a “human–artificial intelligence (AI) shared decision-making” mechanism is integrated into the CBISs. This module operates as a safety net for low-confidence or guideline-sensitive scenarios, complementing the core RL model.

Specifically, when the model generates a recommendation with low predictive certainty or when the patient presents high behavioral variability or complex comorbidities, the CBISs triggers clinician review. Health care professionals evaluate the suggestion in a semiautomated interface, which also displays relevant confidence scores and patient history snapshots.

In addition, a rule-based expert constraint layer is provided, allowing clinicians to embed patient-specific limitations such as contraindications, comorbidity exclusions, or institutional

XSL•FO
RenderX

guidelines (eg, American Urological Association 2024 and International Continence Society). All final behavior plans are generated through an iterative approval flow that synthesizes algorithmic output, patient feedback, and clinician judgment.

This human–AI collaboration mechanism improves accountability, enhances clinical trustworthiness, and ensures DTx remain aligned with evidence-based care pathways.

Pilot Test of Clinical Behavioral Intervention-Supporting System

This pilot study was designed as an exploratory, single-arm, self-controlled evaluation to assess feasibility, adherence, and outcome trends rather than to establish definitive clinical efficacy. A separate cohort of 16 patients with PPI was recruited, distinct from the 150 participants was enrolled in the database construction phase. Participant recruitment and CBISs implementation occurred between January 2024 to March 2024. Participants eligibility criteria were consistent with the formal research protocol. Patient age and time since surgery were recorded. Voiding characteristics, including timing, intervals, and nocturia episodes (both frequency and total volume) were systematically collected. Patient compliance was evaluated based on completion of prescribed CBISs diary entries, with actual usage frequency tracked as the adherence metric. All participants were enrolled to use only the CBISs for postoperative rehabilitation for 3 months.

Both baseline data collection and postintervention outcome assessment were conducted over 5-day periods, with the mean values of these 5-day measurements used as pre- and postintervention results for comparison. The following parameters were recorded: daily voided volume (mL), daily fluid intake (mL), daytime urinary frequency, nighttime urinary frequency, incontinence episodes, urine leakage measured by the 1-hour pad test (g), total score of the ICIQ-UI SF (range 0-21), daily occurrence of urinary urgency, sensation of incomplete bladder emptying, postvoid dribbling, and dysuria (recorded as present or absent).

For each participant, the mean preintervention value was calculated as the arithmetic average of the baseline observations (the first 5 days of recording), yielding the preintervention daily average. The mean postintervention value was derived similarly from the final 5 days of recording. For symptom data recorded as binary outcomes (present=1, absent=0), symptom frequencies were calculated as continuous variables. Preintervention symptom frequency was defined as (number of days the symptom was present during the 5-day baseline period) / 5 × 100%. Postintervention symptom frequency was defined as (number of days the symptom was present during the final 5-day intervention period) / 5 × 100%. For all measured parameters, the change score (Δ) for each participant was computed as: Δ = postintervention value – preintervention value.

All statistical analyses were conducted using R software (version 4.4.2; R Foundation for Statistical Computing). Given the exploratory nature of this pilot study, analyses were performed to describe within-participant change patterns and estimate effect magnitude rather than to test confirmatory hypotheses. A 2-tailed α level of .05 was adopted for descriptive purposes only.

The distribution of individual change scores (Δ) was assessed using the Shapiro-Wilk test. Paired-sample 2-tailed *t* tests or Wilcoxon signed-rank tests were applied as appropriate based on distributional assumptions. Binary symptom variables were converted to symptom frequency percentages and analyzed as continuous outcomes when normality assumptions were met.

Spearman rank correlation coefficient was used to explore associations between changes in selected outcomes. No adjustment for multiple comparisons was performed due to the pilot design and limited sample size, and statistical findings should be interpreted cautiously.

Ethical Considerations

Ethical approval for data collection required for system construction was obtained from the Medical Ethics Committee of Sun Yat-sen Memorial Hospital, Sun Yat-sen University (approval number: SYSKY-2023-925-01). All participants provided electronic informed consent approved by the Medical Ethics Committee (SYSKY-2023-925-01), explicitly detailing (1) bladder diary and behavioral data collection for RL model training to generate personalized recommendations, (2) anonymized data use for system-wide model improvement, (3) right to withdraw data or consent at any time without impact on care, (4) potential implications of AI-driven personalization (eg, adaptive vs static guidance). Consent forms used plain language with examples and were available in Chinese and English.

To mitigate bias from incomplete or anomalous data, preprocessing excluded records with <80% diary completion or >20% anomalies (IQR filtering and time-series checks). Stratified sampling (70/15/15 train/validation/test) preserved leakage severity and adherence distributions. Model fairness was monitored through subgroup performance (age and time since surgery) during 5-fold cross-validation, with clinician override for low-confidence predictions ensuring equity.

For data security, uploaded records were protected using ShangMi 3 (SM3; a cryptographic hash algorithm for integrity verification) and ShangMi 4 (SM4; a symmetric block encryption algorithm for secure data storage and transmission), in accordance with national cryptographic standards. Data handling adhered to China's Personal Information Protection Law and the General Data Protection Regulation–equivalent standards for health data. Health care professionals accessed aggregated views through the CBISs using an authorization key. Additionally, patient-side interfaces integrated real-time alerts for suspicious login attempts or data anomalies, ensuring prompt intervention and maintaining data integrity and patient privacy.

Results

Demographic and Disease-Related Characteristics of Participants

In the first phase of this study, a total of 150 participants were included who provided data for system development, model construction, and demographic characterization. From this

XSL•FO
RenderX

cohort, a subset of 30 additional participants was selected to participate in the focused usability testing phase of CBISs.

The full modeling cohort (N=150) had an age range of 43-82 years, with a mean age of 54.32 (SD 7.51) years. Participants included in the usability testing subsample (n=30) showed a comparable age distribution, ranging from 43-81 (mean 53.70, SD 7.30) years.

Regarding educational attainment, the modeling cohort was predominantly composed of 46 (30.67%) individuals with junior high school education, followed by 35 (23.33%) participants with vocational school education, 28 (18.67%) individuals with undergraduate education, 22 (14.67%) with senior high school education, 12 (8.00%) with primary school education, 6 (4.00%) with postgraduate education, and 1 (0.67%) with no formal education.

In the usability testing subsample, junior high school education was the most prevalent, reported by 13 (43.33%) participants, followed by junior college education with 6 (20%) participants. Senior high school and primary school education were each represented by 4 (13.33%) participants, while vocational school education accounted for 2 (6.67%) participants, and undergraduate education for 1 (3.33%) participants. No participants reported postgraduate education or no formal schooling.

With respect to physical characteristics, the mean BMI of the modeling cohort was 23.44 (SD 3.11) kg/m^2. The usability testing subsample demonstrated a similar BMI distribution, with a mean value of 24.01 (SD 3.18) kg/m^2. Overall, participants included in the usability testing phase were broadly comparable to the full modeling cohort in terms of age and BMI, while exhibiting modest differences in educational composition.

In the pilot test of CBISs, participants used the CBISs an average of 5.2 (SD 1.1) times per day. A total of 11 (68.75%) participants met or exceeded the preset usage target, while 5 (31.25%) participants fell below the target, with a minimum usage of 3 times per day. This indicates generally high engagement with the system during the intervention period. All participants, being older adults, ranged in age from 61-74 years and had a mean age of 67.6 (SD 4.3) years. Among them, 11 (68.75%) participants were within 1-3 months after PCa surgery, and 5 (31.25%) participants were within 3-6 months postoperatively.

The Model of Clinical Behavioral Intervention-Supporting System: Feature Engineering

Based on the patient bladder diary and behavioral monitoring data, a total of 29 features in 6 categories were constructed in this study (Table 1), with the feature construction methods detailed in Table 2. Following data cleaning, the dataset was partitioned via stratified sampling. Model construction was performed based on XGBoost, incorporating Bayesian optimization for model and parameter tuning.

XSL•FO
RenderX

Table 1. Patient data collection form for clinical behavioral intervention-supporting system (CIBSs).

Data category	Instruction
Basic information	
Gender	Discrete variables (male/female)
Age	Continuous variable (years)
BMI	Continuous variable (kg/m^2); associated with symptom typing, behavioral competence, and training intensity appropriateness
Time of onset	Continuous variable; determines symptom chronicity
Urinary flow rate parameter (Q_{max})	Continuous variable; identifies storage, voiding disorders, or bladder hypo-compliance, etc
Urinary flow rate parameter (Q_{ave})	Continuous variable
Residual urine volume test records	Continuous variable
Urinary behavior	
Time of urination	Continuous variable (min); establishment of a 24-hour urinary rhythm
Urinary output	Continuous variable (mL); partially collected automatically by portable urine flow rate instrumentation
Urinary flow rate	Continuous variable (mL/min); automatically collected by portable urinary flow rate instruments
Urination interval	Continuous variable (min); imputed from voiding time difference
Nocturnal urination	Continuous variable (times/mL); includes nocturnal urine frequency and volume
Urinary urgency event	Continuous variable (grade); graded 0-5, subjectively scored by the patient
Symptoms associated with urination	Discrete variables, for example, presence of “urine odor,” “slow urine flow,” “painful urination,” “burning sensation,” “divergent urinary stream,” “feeling of not emptying,” “dribbling after urination,” and so on
Drinking behavior	
Drinking time	Continuous variable (specific time); establish temporal linkage with urination time
Drinking volume	Continuous variable (mL)
Drinking type	Discrete type variable, for example, water, tea, coffee, functional drinks, etc
Daily activity and sleep behavior	
Activity intensity (steps and body movement)	Continuous variable (steps/kcal); from wearable device or mobile health app
Activity period	Discrete variable (daytime/nighttime);
Sleeping time/number of night arousals	Continuous variable (min/times); basis for determining the rhythm between nocturia and nighttime arousal
Specialty scales and subjective scoring data	
Overactive Bladder Syndrome Score	Continuous variable; quantitative grading of symptoms to identify overactive bladder
International Prostate Symptom Score	Continuous variable; determination of symptom severity during storage/voiding phase
International Consultative Committee on Incontinence Questionnaire Short Form	Continuous variable; incontinence symptoms and quality of life implications
Self-assessment Scale for Anxiety and Depression	Continuous variable; assess anxiety and depressive states
Pittsburgh Sleep Quality Index–Sleep Quality Score	Continuous variable; determine the extent to which nocturia affects sleep
System interaction and training feedback data	
Record of training completion	Whether bladder training, water intake program, relaxation training, pelvic floor muscle training, etc, are completed as planned
Patient feedback score	Daily record of perceived efficacy, difficulty, and compliance
Abnormal events	Such as symptom exacerbation or incomplete training prompts

XSL•FO
RenderX

Data category	Instruction
Frequency of use, number of times of punching clock	Judge adherence and compliance level

Table 2. Feature engineering methodology for clinical behavioral intervention-supporting system (CIBSs).

Feature category and name	Source data items	Construction method	Clinical significance
Temporal features[a]			
Daytime mean voiding interval	Urination time, day/night encoding	Mean time difference between voids during 9 AM to 9 PM	Evaluates bladder storage stability
Nocturnal polyuria ratio	Nocturnal urine volume, 24-hour total volume	∑ (nocturnal urine volume) / ∑ (24-hour urine volume) × 100%	Core diagnostic indicator for nocturnal polyuria
Urgency event frequency	Urgency event markers	Voids with urgency (7 days) / total voids (7 days)	Quantifies overactive bladder severity
Ratio features[b]			
Nocturia-arousal association index	Nocturnal voids, sleep interruptions	(Voided volume) / (voided volume + postvoid residual) × 100%	Differentiates pure nocturia from sleep disorders
Effective voiding rate	Voided volume, postvoid residual	Nocturia episodes / sleep interruptions	Reflects bladder emptying efficiency
Irritant beverage intake ratio	Fluid type, intake volume	Coffee + tea intake (mL) / total intake (mL) × 100%	Assesses bladder irritant exposure
Time-period features[c]			
Workday-weekend voiding pattern difference	Voiding time, date type	Mean weekend voiding interval − mean workday voiding interval	Detects stress-related voiding behavior changes
Postexercise urgency marker	Activity period, urgency markers	Binary indicator: urgency within 30 minutes after vigorous exercise	Identifies exercise-induced urgency
Contextual features[d]			
Fluid-void temporal association index	Intake time, voiding time	Voids within 60 minutes post intake / total voids	Evaluates bladder hypersensitivity
PFMT[e] adherence decay rate	Exercise logs, time	Adherence rate (last 3 days) / adherence rate (week 1)	Quantifies decline in behavioral intervention compliance

[a]Based on mean, extreme, and periodic variations within a fixed time window.

[b]Used for reflecting structural relationships.

[c]Used for behavioral pattern recognition.

[d]Describes behavioral markers triggered by specific conditions.

[e]PFMT: pelvic floor muscle training.

Model Training and Validation

The preconstructed database was used to support feature matching and behavioral pattern recognition, which informed the generation of personalized training recommendations. The RL component subsequently updated these recommendations over time based on patient interaction and behavioral feedback, completing the adaptive model construction process (Figure 1).

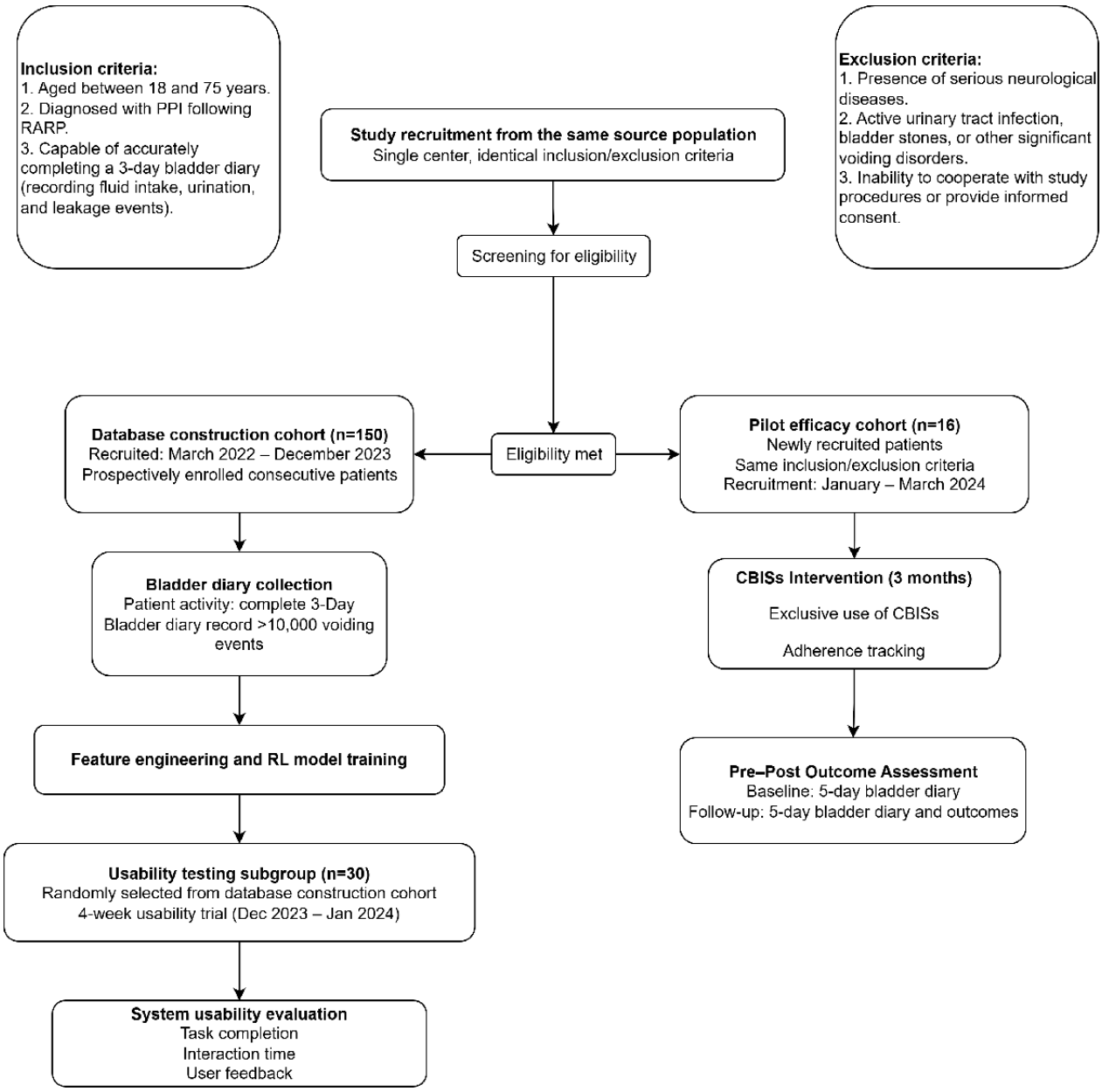
In the current implementation, the RL framework did not incorporate explicit clinical outcome measures (eg, reduction in leakage episodes or symptom score improvement) as direct reward signals. Instead, reward signals were defined using conservative, behavior-centered proxy indicators, including task completion, adherence stability over time, and avoidance of abrupt behavioral changes potentially associated with symptom aggravation. This design choice was intentionally adopted to ensure clinical safety and model stability in this exploratory pilot setting, where outcome responses are delayed, heterogeneous, and not suitable for short-horizon optimization.

Clinical outcome measures were therefore reserved exclusively for post hoc exploratory evaluation rather than real-time reward optimization. Model evaluation metrics (eg, AUC, F_1-score, and MAE) were used internally during development and tuning to support system calibration but were not reported as final performance outcomes, as this study was not designed as a standalone predictive model validation.

XSL•FO
RenderX

Rather than optimizing a single mathematically fixed reward function, the RL component operated within a clinician-supervised decision-support framework, integrating multiple behavior-related feedback signals under predefined safety constraints. Accordingly, the RL module was designed to support adaptive behavioral guidance within a clinically supervised system rather than to function as an autonomous predictive model intended for direct performance benchmarking.

Figure 1. Patient flow diagram for clinical behavioral intervention-supporting system (CBISs) development and pilot feasibility evaluation. PPI: postprostatectomy incontinence; RARP: robot-assisted radical prostatectomy.



The Clinical Behavioral Intervention-Supporting System Mobile App

Overview

Based on the literature review and research team meetings, the main functional modules of the CBISs were identified (Figure 1). The CBISs provide patients with personalized voiding schedules, dietary and fluid intake plans, and pelvic floor muscle contraction training to enhance bladder function. A distinctive feature of the CBISs is its consideration of skin-related symptoms potentially caused by incontinence, offering patients guidance on incontinence-related nursing interventions, that is, the incontinence skin care function. Based on their evaluation of perineal skin condition and severity, the CBISs recommends tasks such as daily cleansing, application of skin-protective dressings, and use of incontinence products. The health care team and administrator portals also receive early warning notifications regarding the occurrence of incontinence dermatitis in patients.

Key Features and Capabilities

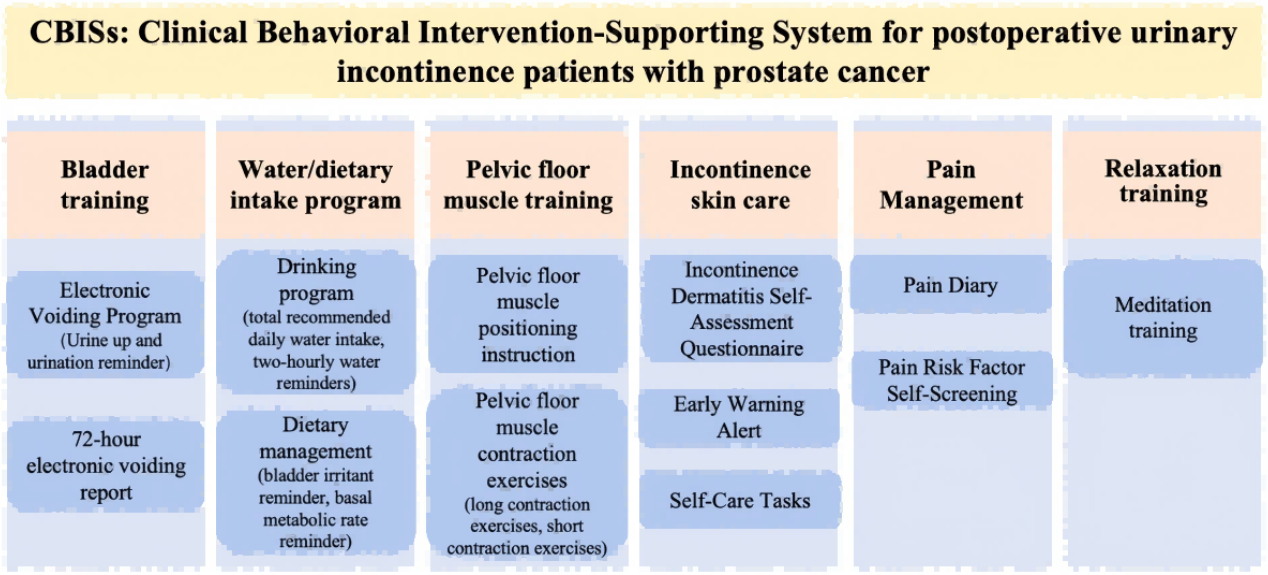
The functional sections of the app and its user interface were identified. The app has 2 interfaces: the home page and the personal center. The home page is divided into 4 sections:

XSL•FO
RenderX

"Training Sections," "Bladder Diary," "Assessment Tools," and "Incontinence Care." The "Personal Center" page provides the user's basic information, contact options for doctors or nurses, and counseling services (Figure 2). Users are required to register an account upon first login and can access the app after using their account number and password. Patients can access the training program after registering their behavioral logs for 3 days.

In the training section, the app provides training programs based on the aforementioned RL models. The initial training program is based on the raw data provided by the patient, and the training content specifically includes voiding interval time, drinking time schedules, rising urine training (bladder capacity training), and pelvic floor muscle contraction and relaxation training. Patient's data are regularly evaluated to assess rehabilitation progress and adjust the training program accordingly. The learning model iteratively adjusts the program, while the health care team may also adjust the program manually. The "Contact Doctor/Nurse" module offers patients support and training guidance throughout the course of their illness to ensure that they maintain good habits after discharge (Figure 3).

Figure 2. Key functional modules of clinical behavioral intervention-supporting system (CBISs).



XSL•FO
RenderX

Figure 3. Clinical behavioral intervention-supporting system (CBISs) user interface. (A) Initial app screen—outpatient; (B) initial app screen—inpatient (patient clinical and demographic characteristics); (C) personal homepage (top to bottom: patient overview—Illness severity, demographics, disease details, and responsible medical personnel; records and support—health and user records, alongside user assistance; rehabilitation log—tracks rehabilitation activities and outcomes; service purchase—handles the acquisition of additional services); (D) bladder ultrasound monitoring; (E) urinary pad test; and (F) questionnaire entry screen.



Safety and Feasibility

Regarding digital confidentiality, patients upload data by completing their own records as well as using wearable devices. Because the data are time series in nature, the app assigns timestamps to each data point and incorporates preprocessing interfaces to filter missing values and outliers. The study follows SM3 and SM4 encryption algorithms for desensitized and encrypted storage of data, alongside hierarchical management of data permissions, facilitating separate review processes for health care professionals and patients. Data privacy complies with the Personal Information Protection Act.

XSL•FO
RenderX

Pilot Evaluation of Clinical Behavioral Intervention-Supporting System Feasibility and Outcome Trends

Participants used the CBISs an average of 5.2 (SD 1.1) times per day. Following implementation of the CBISs-guided bladder function training program, directional changes were observed across multiple bladder diary parameters, incontinence-related outcomes, and patient-reported symptom measures (Tables 3 and 4). Given the exploratory nature of this pilot study and the limited sample size, results are presented to describe change patterns rather than to support confirmatory inference.

Mean voided volume increased from 984.6 (SD 132.7) mL at baseline to 1128.4 (SD 148.2) mL post intervention. Fluid intake also increased over the intervention period, suggesting changes in hydration and voiding behaviors consistent with the individualized training guidance delivered by the system.

Daytime urinary frequency showed a reduction from 5.74 (SD 1.21) episodes per day to 4.69 (SD 1.08) episodes per day. Nighttime urinary frequency similarly declined, with median values decreasing from 1.8 (IQR 1.6-2.2) episodes per night to 1.0 (IQR 1.0-1.6) episodes per night. Total urinary frequency demonstrated a comparable downward trend over the intervention period, reflecting changes in urinary storage and voiding patterns.

Incontinence episodes declined from a median of 7.0 (IQR 6.0-11.0) episodes per day at baseline to 4.0 (IQR 2.0-6.0) episodes per day after the intervention. Pad test measurements indicated lower postintervention urine leakage values compared with baseline, with a shift toward milder leakage severity categories observed among participants. Specifically, the number of patients classified as having severe or greater leakage decreased, while the number of patients classified as having mild leakage increased.

Patient-reported outcomes showed parallel trends. ICIQ-UI SF scores were lower after the intervention compared with baseline, indicating a reduction in perceived symptom burden. Exploratory correlation analysis suggested alignment between changes in patient-reported symptom scores and reductions in objective urine leakage measures; however, these associations should be interpreted cautiously given the small sample size and exploratory design.

The proportion of patients reporting urinary urgency and sensation of incomplete bladder emptying was lower following the intervention period. Dysuria incidence also declined, whereas changes in postvoid dribbling were less pronounced and more variable across individuals.

Effect size estimates were calculated to describe the magnitude of observed changes across outcomes. Several parameters demonstrated moderate to large effect size values; however, these estimates are provided for descriptive and hypothesis-generating purposes only and should not be interpreted as evidence of definitive clinical benefit in the absence of a concurrent control group.

Overall, the pilot results indicate that the CBISs can be feasibly implemented in postprostatectomy patients and is associated with consistent directional changes across multiple objective and subjective outcome domains. No causal conclusions regarding clinical efficacy can be drawn from this exploratory analysis.

Table 3. Exploratory changes in bladder diary parameters before and after the clinical behavioral intervention-supporting system (CBISs)–guided training program. Statistical comparisons are exploratory and intended to describe within-participant change patterns rather than to support confirmatory inference. Effect sizes are reported for descriptive purposes only.

Variable	Preintervention, mean (SD)	Postintervention, mean (SD)	Mean difference, 95% CI	2-tailed *t* test (*df*)	*P* value	Cohen *d*
Voided volume (mL)	984.6 (132.7)	1128.4 (148.2)	143.8 (69.2 to 218.4)	4.32 (15)	<.001	1.08
Fluid intake (mL)	1246.8 (181.4)	1428.2 (198.6)	181.4 (95.6 to 267.2)	5.12 (15)	<.001	1.28
Daytime urinary frequency	5.74 (1.21)	4.69 (1.08)	–1.05 (–1.68 to –0.42)	–3.56 (15)	.003	0.89
Total urinary frequency	7.86 (1.90)	5.96 (1.24)	–1.90 (–2.56 to –1.24)	–6.12 (15)	<.001	1.53
Urinary urgency frequency, %	68.8 (18.2)	31.3 (16.8)	–37.5 (–47.5 to –27.5)	–8.23 (15)	<.001	1.08
Incomplete emptying sensation frequency, %	58.8 (19.5)	28.8 (15.6)	–30.0 (–40.6 to –19.4)	–6.15 (15)	<.001	1.28
Postvoid dribbling frequency, %	25.0 (15.8)	18.8 (12.4)	–6.2 (–14.8 to 2.4)	–1.56 (15)	0.139	0.89

XSL•FO
RenderX

Table 4. Exploratory changes in incontinence-related outcomes and patient-reported symptoms before and after the clinical behavioral intervention-supporting system (CBISs)–guided training program. Statistical analyses were conducted to characterize the direction and magnitude of changes in this pilot study. *P* values and effect sizes should be interpreted cautiously due to the small sample size and lack of a control group. For dysuria, the value represents percentage change.

Variable	Preintervention, median (IQR)	Postintervention, median (IQR)	Median difference	*z* value	*P* value	Effect size[a], *r*
Nighttime urinary frequency, n	1.8 (1.6-2.2)	1.0 (1.0-1.6)	–0.8	–3.82	<.001	0.68
Incontinence episodes, n	7.0 (6.0-11.0)	4.0 (2.0-6.0)	–3.0	–5.56	<.001	0.99
Urine leakage, g	8.5 (4.0-19.0)	3.5 (2.0-9.0)	–5.0	–5.56	<.001	0.99
ICIQ-UI SF[b] score	14.0 (12.0-20.0)	9.0 (6.0-16.0)	–5.0	–5.32	<.001	0.95
Dysuria incidence, %	43.8	25.0	–18.8*	–3.41	.001	0.61

[a]Effect size interpretation: Cohen *d*: small (0.2), medium (0.5), large (0.8); r effect size: small (0.1), medium (0.3), large (0.5).

[b]ICIQ-UI SF: International Consultation on Incontinence Questionnaire–Short Form.

Discussion

Overview

The pilot study provided preliminary evidence of improvements in urinary control and patient-reported symptoms, with the CBISs simulating professional supervision by dynamically adjusting behavioral prescriptions based on adherence patterns, symptom trajectories, and real-time data. Pilot results indicated enhanced urinary control and improved subjective symptom experiences among patients with PPI.

PPI etiology remains incompletely understood, with risk factors including age, BMI, prostate size, oncologic and surgical factors, and shorter membranous urethral length [20-23]. Despite refined surgical techniques and preoperative precautions, PPI persists, severely impacting QoL, finances, and psychology. Conservative treatment—lifestyle modifications, bladder diaries, and PFMT—is first-line for mild PPI, yet a significant implementation gap exists due to poor adherence influenced by fatigue, transportation, time constraints [24-27], inadequate PFMT technique comprehension, and lower cognitive ability [28]. Supervised rehabilitation outperforms unsupervised approaches [29,30].

Traditional in-person care faces geographical, resource, and adherence limitations. Digital health interventions extend support beyond clinics [31]. This study developed CBISs, which collects multidimensional data (diet, urination, activities, adherence, and perceived severity) to generate individualized RL-driven recommendations through XGBoost with rule-engine integration, continuously optimizing through cross-validation while preventing overfitting.

Several digital therapeutic approaches for PPI have been reported, including cognitive-behavioral therapy–based telehealth interventions for urinary incontinence control and QoL [31,32], mobile health apps combining bladder diary logging with PFMT prompts [33], perioperative telehealth programs with remote monitoring calls [28], WeChat-based health education and extended care services [34,35], social media platforms for patient education and community support [36], and proactive digital health interventions aimed at symptom reduction [37]. These platforms primarily focus on delivering educational content, standardized exercise instructions, symptom tracking, and periodic feedback through telehealth communication, messaging apps, or social media engagement. While such approaches improve accessibility and patient awareness, their rehabilitation protocols are generally static, relying on rule-based prompts, fixed schedules, or clinician-mediated adjustments. Consequently, they have limited capacity to adapt training intensity, content, or timing in response to individual patient trajectories, fluctuating adherence, or delayed and heterogeneous treatment responses.

In contrast, the CBISs introduces an RL–driven, closed-loop behavioral intervention framework that emphasizes continuous personalization rather than predefined content delivery, consistent with recent applications of RL in noncommunicable disease management and personalized medicine [38,39]. By iteratively integrating multidimensional patient data—including voiding behaviors, fluid intake patterns, and training adherence—the CBISs dynamically adjusts rehabilitation guidance in response to individual progress and engagement over time. This adaptive design is particularly suited to behavioral rehabilitation contexts, where patient responses evolve gradually and vary substantially across individuals, representing a key limitation of conventional static or rule-based digital platforms.

From a usability and scalability perspective, the CBISs is designed to translate complex behavioral information into actionable, individualized recommendations while reducing reliance on continuous clinician involvement. Clinicians primarily function as supervisors supported by automated monitoring and alert mechanisms, which may help mitigate workforce constraints in real-world clinical settings. Although this pilot study was not designed to establish definitive efficacy, the observed favorable trends in urinary symptom–related outcomes suggest that RL-driven adaptability may help address common challenges of existing digital rehabilitation approaches, including declining adherence and variable response patterns.

Collectively, these features distinguish the CBISs from existing digital therapeutic platforms by shifting the focus from content-centered delivery to behavior-centered, data-driven optimization. This paradigm more closely reflects the dynamic nature of rehabilitation processes and provides a plausible

XSL•FO
RenderX

pathway for improving long-term engagement and personalization in PPI management.

Several limitations warrant consideration. First, transparency and interpretability remain inherent challenges of machine learning–based systems. Although clinician oversight was incorporated into the CBISs workflow, the "black box" nature of RL algorithms necessitates ongoing monitoring to ensure the clinical appropriateness and reliability of generated recommendations. For example, the model may recommend aggressive voiding interval reductions (±15 min) for patients with irregular adherence decay without clearly explaining feature interactions (eg, fluid intake influencing urgency events). In pilot testing, 3 of 16 (18.8%) cases triggered low-confidence alerts routed to clinician override due to unexplained Q value shifts during nocturia spikes, highlighting the need for explainable AI techniques like Shapley additive explanations analysis in future iterations. CBISs relies on patient-reported bladder diaries and potential wearable integration, introducing variability from self-reporting inaccuracies (eg, underreported leakage) or device nonuse (12% dropout in usability testing). Participation may have induced Hawthorne effects, with improvements potentially attributable to heightened attention from research involvement rather than CBISs alone. Additionally, while pilot (n=16) and modeling (n=150) cohorts shared consistent inclusion and exclusion criteria, all participants were motivated outpatients from routine clinical settings, potentially overrepresenting compliant, tech-savvy patients and limiting generalizability to broader PPI populations with lower motivation or compliance challenges. In addition, the system processes large volumes of sensitive patient-level data, underscoring the importance of robust data security, privacy protection, and governance mechanisms for real-world implementation.

The pilot evaluation used a prospective self-controlled pre-post design, which is appropriate for early-stage investigation of interventions requiring high patient engagement and ethical sensitivity. By using patients as their own controls, this design reduced interindividual variability and improved statistical efficiency under small-sample conditions, supporting preliminary exploration of feasibility and potential effect sizes. The concurrent use of objective measures (pad test) and validated patient-reported outcomes (ICIQ-UI SF) further strengthened result interpretation.

Nevertheless, the absence of a concurrent control group limits causal inference, and the observed changes may be partially attributable to nonspecific factors such as time effects, repeated measurement, or expectancy-related bias. The relatively short follow-up period also precludes assessment of long-term effectiveness, sustainability of behavioral change, and delayed adverse effects.

Future studies incorporating appropriate control groups, blinded outcome assessment, and extended follow-up are warranted to establish the definitive clinical value of this intervention. In addition, future work will explore the integration of validated and delayed clinical outcome signals into the RL framework once sufficient longitudinal and controlled trial data become available.

Future studies with complete longitudinal labeling and independent test cohorts will enable formal reporting of model performance metrics such as AUC, F_1-score, and MAE. The lack of a fully formalized mathematical reward function reflects an intentional design choice in this exploratory phase and will be addressed in future work when sufficient validated outcome data become available. Comparative evaluation against simpler baseline models was beyond the scope of this feasibility-focused study and will be addressed in future controlled investigations.

Conclusions

This study describes the development and preliminary evaluation of a behavioral therapy–based CBISs designed to support personalized rehabilitation for PPI. By integrating multidimensional patient data—including voiding behaviors, fluid intake patterns, PFMT adherence, and patient-reported outcomes—the CBISs applies an RL framework to deliver adaptive, individualized behavioral guidance beyond static information delivery.

Rather than replacing clinician decision-making, the CBISs functions as a digital extension of guideline-based care, with the potential to enhance continuity, personalization, and accessibility of rehabilitation support outside traditional clinical settings. The findings from this early-stage study suggest feasibility and favorable trends, supporting further investigation of this approach.

Future multicenter randomized controlled trials are warranted to address several key research questions including whether (1) RL-driven personalization improves continence recovery, QoL, and adherence compared with standard or nonpersonalized digital rehabilitation; (2) observed benefits are sustained over longer follow-up periods; (3) the system performs across diverse patient populations, clinical environments, and levels of baseline motivation; and (4) adaptive, closed-loop interventions confer incremental value over conventional rule-based digital platforms. Addressing these questions will be essential for defining the clinical effectiveness, generalizability, and implementation potential of the CBISs.

As DTx continue to evolve, systems such as the CBISs may offer a scalable pathway toward precision rehabilitation, supporting individualized recovery trajectories for patients navigating functional challenges following prostate cancer surgery.

Acknowledgments

The authors gratefully acknowledge all participants and individuals whose contributions made this study possible.

XSL•FO
RenderX

Funding

This work was supported by the Key Technologies Research and Development Program of the Ministry of Science and Technology of China (grant: 2023YFC3606004) and Yat-Sen Scholarship for Young Nursing Scientists of Sun Yat-sen Memorial Hospital, Sun Yat-sen University (grant: YXHL202102). The funders had no role in the study design, data collection, analysis, interpretation, manuscript preparation, or the decision to publish this work.

Data Availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Authors' Contributions

FF contributed to conceptualization, data curation, formal analysis, methodology, project administration, funding acquisition, and writing of the original draft. Hao H contributed to conceptualization, data curation, formal analysis, methodology, and writing of the original draft. CYX contributed to data curation, formal analysis, methodology, and writing of the original draft. XW contributed to formal analysis and writing—review and editing. CZ contributed to formal analysis, methodology, and writing—review and editing. DW contributed to resources, conceptualization, methodology, and writing—review and editing. JY and YQ contributed to project administration, resources, supervision, conceptualization, methodology, and writing—review and editing. JY contributed to resources, software, methodology, and writing of the original draft. Hai H contributed to resources, supervision, conceptualization, methodology, and writing—review and editing.

Conflicts of Interest

None declared.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263 [FREE Full text] [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Prostate Cancer - Cancer Stat Facts.: National Cancer Institute URL: https://seer.cancer.gov/statfacts/html/prost.html [accessed 2025-12-17]
3. Lu H, Han T, Li F, Yang J, Hou Z. Global trends and hotspots in research of robotic surgery in oncology: a bibliometric and visual analysis from 2002 to 2021. Front Oncol 2022;12:1055118 [FREE Full text] [doi: 10.3389/fonc.2022.1055118] [Medline: 36439475]
4. Ávila M, Patel L, López S, Cortés-Sanabria L, Garin O, Ferrer F. Patient-reported outcomes after treatment for clinically localized prostate cancer: a systematic review and meta-analysis. Cancer Treat Rev 2018;66:23-44. [doi: 10.1016/j.ctrv.2018.03.005] [Medline: 29673922]
5. Clements MB, Gmelich CC, Vertosick EA, Hu JC, Sandhu JS, Scardino PT. Have urinary function outcomes after radical prostatectomy improved over the past decade? Cancer 2022;128(5):1066-1073 [FREE Full text] [doi: 10.1002/cncr.33994] [Medline: 34724196]
6. Vernooij RWM, Cremers RGHM, Jansen H, Somford DM, Kiemeney LA, van Andel G. Urinary incontinence and erectile dysfunction in patients with localized or locally advanced prostate cancer: a nationwide observational study. Urol Oncol 2020;38(9):735.e17-735.e25. [doi: 10.1016/j.urolonc.2020.05.022] [Medline: 32680820]
7. Gacci M, De Nunzio C, Sakalis V, Rieken M, Cornu J, Gravas S. Latest evidence on post-prostatectomy urinary incontinence. J Clin Med 2023;12(3):1190 [FREE Full text] [doi: 10.3390/jcm12031190] [Medline: 36769855]
8. Al Hussein Al Awamlh B, Wallis CJD, Penson DF, Huang L, Zhao Z, Conwill R. Functional outcomes after localized prostate cancer treatment. JAMA 2024;331(4):302-317. [doi: 10.1001/jama.2023.26491] [Medline: 38261043]
9. Wu X, Wong CH, Gandaglia G, Chiu PK, EAU–Young Academic Urology Prostate Cancer Working Party (YAU-PCa WP). Urinary continence in high-risk prostate cancer after robot-assisted radical prostatectomy. Curr Opin Urol 2023;33(6):482-487. [doi: 10.1097/MOU.0000000000001127] [Medline: 37646515]
10. Eastham JA, Auffenberg GB, Barocas DA, Chou R, Crispino T, Davis JW. Clinically localized prostate cancer: AUA/ASTRO guideline, part II: principles of active surveillance, principles of surgery, and follow-up. J Urol 2022;208(1):19-25. [doi: 10.1097/JU.0000000000002758] [Medline: 35536148]
11. Fernandes ACN, Jorge CH, Weatherall M, Ribeiro IV, Wallace SA, Hay-Smith EJC. Pelvic floor muscle training with feedback or biofeedback for urinary incontinence in women. Cochrane Database Syst Rev 2025;3(3):CD009252. [doi: 10.1002/14651858.CD009252.pub2] [Medline: 40066950]
12. Vecchio M, Chiaramonte R, DI Benedetto P. Management of bladder dysfunction in multiple sclerosis: a systematic review and meta-analysis of studies regarding bladder rehabilitation. Eur J Phys Rehabil Med 2022;58(3):387-396 [FREE Full text] [doi: 10.23736/S1973-9087.22.07217-3] [Medline: 35102733]

XSL•FO
RenderX

13. Carrera A, Manetti S, Lettieri E. Rewiring care delivery through Digital Therapeutics (DTx): a machine learning-enhanced assessment and development (M-LEAD) framework. BMC Health Serv Res 2024;24(1):237 [FREE Full text] [doi: 10.1186/s12913-024-10702-z] [Medline: 38395905]
14. Agarwal P, Mukerji G, Desveaux L, Ivers NM, Bhattacharyya O, Hensel JM. Mobile app for improved self-management of type 2 diabetes: multicenter pragmatic randomized controlled trial. JMIR Mhealth Uhealth 2019;7(1):e10321 [FREE Full text] [doi: 10.2196/10321] [Medline: 30632972]
15. Ahmad Z. Integration of wearable digital urinary sensors and patient-reported outcomes to monitor post-prostatectomy incontinence: a prospective pilot study. Urology 2025;206:25-30. [doi: 10.1016/j.urology.2025.07.038] [Medline: 40714011]
16. Di S, Petch J, Gerstein HC, Zhu R, Sherifali D. Optimizing health coaching for patients with type 2 diabetes using machine learning: model development and validation study. JMIR Form Res 2022;6(9):e37838. [doi: 10.2196/37838] [Medline: 36099006]
17. Ghasemi P, Greenberg M, Southern DA, Li B, White JA, Lee J. Personalized decision making for coronary artery disease treatment using offline reinforcement learning. NPJ Digit Med 2025;8(1):99. [doi: 10.1038/s41746-025-01498-1]
18. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med 2018;24(11):1716-1720 [FREE Full text] [doi: 10.1038/s41591-018-0213-5] [Medline: 30349085]
19. Ito H, Abrams P, Lewis AL, Young GJ, Blair PS, Cotterill N. Use of the International Consultation on Incontinence Questionnaires bladder diary in men seeking therapy for lower urinary tract symptoms. Eur Urol Focus 2022;8(1):66-74. [doi: 10.1016/j.euf.2021.12.009] [Medline: 35027329]
20. Heesakkers J, Farag F, Bauer RM, Sandhu J, De Ridder D, Stenzl A. Pathophysiology and contributing factors in postprostatectomy incontinence: a review. Eur Urol 2017;71(6):936-944. [doi: 10.1016/j.eururo.2016.09.031] [Medline: 27720536]
21. Kim M, Park M, Shim M, Choi S, Lee SM, Lee E. Effect of preoperative urodynamic detrusor overactivity on post-prostatectomy incontinence: a systematic review and meta-analysis. Int Urol Nephrol 2016;48(1):53-63. [doi: 10.1007/s11255-015-1141-7] [Medline: 26507516]
22. Lardas M, Grivas N, Debray TPA, Zattoni F, Berridge C, Cumberbatch M. Patient- and tumour-related prognostic factors for urinary incontinence after radical prostatectomy for nonmetastatic prostate cancer: a systematic review and meta-analysis. Eur Urol Focus 2022;8(3):674-689 [FREE Full text] [doi: 10.1016/j.euf.2021.04.020] [Medline: 33967010]
23. Pastore AL, Palleschi G, Illiano E, Zucchi A, Carbone A, Costantini E. The role of detrusor overactivity in urinary incontinence after radical prostatectomy: a systematic review. Minerva Urol Nefrol 2017;69(3):234-241 [FREE Full text] [doi: 10.23736/S0393-2249.16.02790-9] [Medline: 27808489]
24. Johnson EE, Mamoulakis C, Stoniute A, Omar MI, Sinha S. Conservative interventions for managing urinary incontinence after prostate surgery. Cochrane Database Syst Rev 2023;4(4):CD014799 [FREE Full text] [doi: 10.1002/14651858.CD014799.pub2] [Medline: 37070660]
25. Luo L, Chen X, Xie H, Zhou J, Li L. Development and evaluation of a rehabilitation training compliance scale for patients with urinary incontinence. BMC Nurs 2023;22(1):147 [FREE Full text] [doi: 10.1186/s12912-023-01326-5] [Medline: 37138310]
26. Santa Mina D, Hilton WJ, Matthew AG, Awasthi R, Bousquet-Dion G, Alibhai SMH. Prehabilitation for radical prostatectomy: a multicentre randomized controlled trial. Surg Oncol 2018;27(2):289-298. [doi: 10.1016/j.suronc.2018.05.010] [Medline: 29937184]
27. Zhang AY, Fu AZ, Moore S, Zhu H, Strauss G, Kresevic D. Is a behavioral treatment for urinary incontinence beneficial to prostate cancer survivors as a follow-up care? J Cancer Surviv 2017;11(1):24-31 [FREE Full text] [doi: 10.1007/s11764-016-0557-0] [Medline: 27341843]
28. Goode PS, Johnson TM, Newman DK, Vaughan CP, Echt KV, Markland AD. Perioperative mobile telehealth program for post-prostatectomy incontinence: a randomized clinical trial. J Urol 2022;208(2):379-387. [doi: 10.1097/JU.0000000000002697] [Medline: 35389239]
29. Baumann FT, Reimer N, Gockeln T, Reike A, Hallek M, Ricci C. Supervised pelvic floor muscle exercise is more effective than unsupervised pelvic floor muscle exercise at improving urinary incontinence in prostate cancer patients following radical prostatectomy - a systematic review and meta-analysis. Disabil Rehabil 2022;44(19):5374-5385. [doi: 10.1080/09638288.2021.1937717] [Medline: 34550846]
30. Junwen S, Rongjiang W. The efficacy of the WeChat app combined with pelvic floor muscle exercise for the urinary incontinence after radical prostatectomy. Biomed Res Int 2020;2020:6947839 [FREE Full text] [doi: 10.1155/2020/6947839] [Medline: 32258138]
31. Izidoro LCDR, Mata LRFD, Azevedo C, Paula AAPD, Pereira MG, Santos JEMD. Cognitive-behavioral program to control lower urinary tract symptoms after radical prostatectomy: a randomized clinical trial. Rev Bras Enferm 2022;75(5):e20210818 [FREE Full text] [doi: 10.1590/0034-7167-2021-0818] [Medline: 35920515]
32. de Resende Izidoro LC, Azevedo C, Pereira MG, Chianca TCM, Borges CJ, de Almeida Cavalcante Oliveira LM. Effect of cognitive-behavioral program on quality of life in men with post-prostatectomy incontinence: a randomized trial. Rev

XSL•FO
RenderX

Esc Enferm USP 2024;58:e20240187 [FREE Full text] [doi: 10.1590/1980-220X-REEUSP-2024-0187en] [Medline: 39475393]
33. Estevam FEB, Machado AF, Azevedo C, Izidoro LCDR, Anjos FMDSD, Oliveira HMD. Mobile health application for the treatment of urinary incontinence after radical prostatectomy: development and quality analysis. Rev Esc Enferm USP 2024;58:e20240119 [FREE Full text] [doi: 10.1590/1980-220X-REEUSP-2024-0119en] [Medline: 39443287]
34. Gu T, Wang H, Shen Y, Jiang H, Gu Y, Shi J. The application of health education with the combination of hospital WeChat platform and teach-back method in improving the quality of life of patients with prostate cancer after operation. Chin J Pract Nurs 2024;40(18):1403-1410 [FREE Full text] [doi: 10.3760/cma.j.cn211501-20230813-00266]
35. Hou L, Ma W, Jia F, Zhang B. Evaluation of the efficacy of WeChat-based extended care service for patients with urinary incontinence after radical prostatectomy for prostate cancer. Int J Nurs 2024;43(11):2099-2102. [doi: 10.3760/cma.j.cn221370-20230424-00493]
36. Struck JP, Siegel F, Kramer MW, Tsaur I, Heidenreich A, Haferkamp A. Substantial utilization of Facebook, Twitter, YouTube, and Instagram in the prostate cancer community. World J Urol 2018;36(8):1241-1246. [doi: 10.1007/s00345-018-2254-2] [Medline: 29523948]
37. Hou S, Qiao W, Li Y, He H, Wu B, Dai Y. Effectiveness of proactive health interventions in reducing symptoms and enhancing self-efficacy and self-management in prostate cancer survivors: a randomized controlled trial. J Cancer Surviv 2025;19(5):1761-1773. [doi: 10.1007/s11764-024-01706-z] [Medline: 39542992]
38. Zhao Y, Chaw JK, Liu L, Chaw SH, Ang MC, Ting TT. Systematic literature review on reinforcement learning in non-communicable disease interventions. Artif Intell Med 2024;154:102901. [doi: 10.1016/j.artmed.2024.102901] [Medline: 38838400]
39. Venkatesan L, Benjamin LS, Satchi NS. Reinforcement learning in personalized medicine: a comprehensive review of treatment optimization strategies. Cureus 2025;17(4):e82756. [doi: 10.7759/cureus.82756] [Medline: 40406780]

Abbreviations

AI: artificial intelligence
AUC: area under the receiver operating characteristic curve
CBIS: clinical behavioral intervention-supporting system
DTx: digital therapeutics
ICIQ-UI SF: International Consultation on Incontinence Questionnaire–Short Form
MAE: mean absolute error
PCa: prostate cancer
PFMT: pelvic floor muscle training
PPI: postprostatectomy incontinence
QoL: quality of life
RL: reinforcement learning
SM3: ShangMi 3
SM4: ShangMi 4
XGBoost: extreme gradient boosting

Edited by M Balcarras; submitted 01.Sep.2025; peer-reviewed by W Xia, U Anceschi, R Chiaramonte; comments to author 17.Nov.2025; revised version received 04.Jan.2026; accepted 05.Jan.2026; published 06.Feb.2026.

Please cite as:
Fan F, Huang H, Yan J, Xu CY, Wu X, Zhou C, Wen D, Huang H, Li HC, Qiu Y
Reinforcement Learning–Based Digital Therapeutic Intervention for Postprostatectomy Incontinence: Development and Pilot Feasibility Study
JMIR Cancer 2026;12:e83375
URL: https://cancer.jmir.org/2026/1/e83375
doi:10.2196/83375
PMID:

©Fan Fan, Hao Huang, Jingwen Yan, Chao-Yue Xu, Xiuhua Wu, Chunmei Zhou, Dandan Wen, Hai Huang, Ho Cheung Li, Yihong Qiu. Originally published in JMIR Cancer (https://cancer.jmir.org), 06.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is

XSL•FO
RenderX

properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Hashimoto Thyroiditis and Progression of Papillary Thyroid Cancer: 10-Year Retrospective Cohort Study

Xin Li[1*], MD; Xiangyun Yao[2*], MD; Rui Shan[3], MM; Fang Mei[4,5], PhD; Shibing Song[1], MD; Bangkai Sun[6], MS; Chunhui Yuan[1], MD; Zheng Liu[3], PhD

[1]Department of General Surgery, Peking University Third Hospital, Beijing, China

[2]Department of Ultrasound, Peking University Third Hospital, Beijing, China

[3]Department of Maternal and Child Health, School of Public Health, Peking University, 38 Xueyuan Road , Haidian District, Beijing, China

[4]Department of Pathology, Peking University Third Hospital, Beijing, China

[5]School of Basic Medical Sciences, Peking University Health Science Center, Beijing, China

[6]Information Management and Big Data Center, Peking University Third Hospital, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Zheng Liu, PhD
Department of Maternal and Child Health, School of Public Health, Peking University, 38 Xueyuan Road , Haidian District, Beijing, China

Abstract

Background: In recent years, the global incidence of thyroid cancer has been increasing.

Objective: This study aimed to examine the association between Hashimoto thyroiditis (HT) and papillary thyroid cancer (PTC) progression under active surveillance (AS).

Methods: Our retrospective study was conducted at Peking University Third Hospital and included 203 patients with PTC who underwent AS for ≥6 months before surgery. Patients were first categorized into 2 groups: the HT group (n=90) and the non-HT group (n=113). Cox proportional hazards models were then used to evaluate the association between HT and PTC progression during AS, adjusting for age, sex, baseline tumor size, BMI, pregnancy status, number of tumor foci, and thyroid-stimulating hormone level. Subgroup analyses stratified by the 6 covariates mentioned above were also applied to explore the potential effect modification.

Results: No significant difference was observed between the HT and non-HT groups in PTC progression-free survival (hazard ratio [HR] 1.11, 95% CI 0.61 - 1.99; *P*=.74), tumor enlargement-free survival (HR 1.02, 95% CI 0.56 - 1.86; *P*=.95), or lymph node metastasis-free survival (HR 1.76, 95% CI 0.31 - 10.12; *P*=.52). Subgroup analyses revealed a potential interaction between HT and BMI. Among patients who were overweight or obese (BMI >24 kg/m^2), HT was significantly associated with an increased risk of disease progression (HR 6.32, 95% CI 1.84 - 21.69; *P*=.003), while among patients with BMI ≤24 kg/m^2, no association between HT and progression risk was observed (*P*=.01).

Conclusions: We found no evidence of association between HT and PTC progression during AS. However, the relationship between HT and PTC progression may be modified by overweight or obesity status.

(JMIR Cancer 2026;12:e80535) doi:10.2196/80535

KEYWORDS

papillary thyroid cancer; Hashimoto thyroiditis; progression; retrospective cohort.; progression-free survival; lymph node

Introduction

Thyroid cancer is not only the most common malignant tumor of the head and neck but also the most prevalent endocrine malignancy. In recent years, the global incidence of thyroid cancer has been steadily increasing. Between 1990 and 2020, the number of newly diagnosed cases worldwide rose from 95,000 to 586,000 [1-3]. In China, the age-standardized incidence rate increased from 3.21 per 100,000 in 2005 to 24.64 per 100,000 in 2022 [4,5]. Clinically, papillary thyroid cancer (PTC) is the most frequently encountered, representing approximately 80% of all thyroid cancer cases worldwide [6].

As understanding of PTC has evolved, active surveillance (AS) has been increasingly adopted by Japan and Korea as part of a comprehensive management approach for low-risk PTC [7,8]. The US and Chinese guidelines also consider AS an acceptable treatment approach [9,10]. AS is an emerging management strategy that offers a more conservative alternative to immediate

XSL•FO
RenderX

surgery [9]. It involves active monitoring of the patient's condition without initiating surgical treatment unless there is evidence of tumor progression [11]. Active monitoring primarily refers to performing neck ultrasound every 6 months for the first 1 to 2 years, followed by annual examinations. Despite growing interest in AS for low-risk PTC, there is still no clear consensus on which patients should undergo immediate surgery and which are appropriate candidates for AS.

Hashimoto thyroiditis (HT) is an autoimmune disorder characterized by immune-mediated destruction of thyroid cells through both cellular and humoral mechanisms [12]. The incidence of HT has been increasing steadily over the years, and its coexistence with PTC is relatively common [13-15]. Some studies have suggested that HT may be a risk factor for the development of PTC, while others have indicated that HT could be a protective factor against postoperative recurrence of PTC [13,14,16].

Nevertheless, the vast majority of existing research has focused solely on the postoperative recurrence, overlooking the increasingly prevalent management strategy of AS in recent years. Currently, there is limited evidence on the association between HT and PTC progression during AS. Importantly, the risk factors for postoperative recurrence and those for disease progression under AS may differ substantially [9,17]. Identifying the risk factors for tumor progression during AS is crucial for selecting appropriate candidates, helping to avoid both undertreatment and overtreatment [9]. Moreover, patients with HT tend to be younger and more likely to be female compared to those without HT [14]; both sex and age have been independently associated with disease progression [18,19]. Therefore, it is essential to investigate the role of HT in PTC progression while carefully accounting for potential confounders.

Two reviews have investigated the role of HT in PTC progression [3,13]; however, they have several notable limitations:

1. Lack of stringent inclusion and exclusion criteria: most of the included studies focused on patients who underwent surgery, providing limited insight into the role of HT during AS.
2. Failure to account for confounding factors: key confounders such as sex and age were not adequately considered. For example, the observed association between HT and favorable prognosis in some studies may be attributed to the higher proportion of female patients in the HT group.
3. Unclear definition of outcome measures: these reviews did not clearly distinguish between the role of HT in tumor progression during AS and its effect on recurrence after surgery. This distinction is important, as the extent of surgery may differ for patients with HT and may itself act as a confounding factor in recurrence outcomes.
4. Lack of quantitative synthesis: one of the reviews did not include a quantitative meta-analysis, limiting the ability to draw robust conclusions [3].

Therefore, we investigated the relationship between HT and progression-free survival of PTC under AS—including tumor enlargement and lymph node metastasis (LNM)—while controlling for multiple potential confounding factors.

Methods

Ethical Considerations

The study was approved by the Medical Ethics Committee of Peking University Third Hospital (ethical project number: IRB00006761-M2022721). The requirement for informed consent was waived by the Ethics Committee due to the retrospective nature of the study. All patient data were deidentified and anonymized prior to analysis to ensure the privacy of individuals.

Study Population

The Electronic Medical Record system of Peking University Third Hospital was searched to retrieve cases of thyroid surgery with a subsequent diagnosis of PTC, covering the period from January 2012 to September 2022.

The AS strategy was routinely implemented at our institution beginning in 2015. Since then, all patients diagnosed with low-risk PTC have been introduced to the AS option during outpatient consultations. Following comprehensive counseling by clinicians regarding the risks associated with thyroid surgery, the potential need for thyroid hormone replacement therapy, and the possibility of disease progression during AS, patients make their decision based on individual circumstances—such as pregnancy plans and overall health status—on whether to pursue AS. Patients independently decided whether to undergo AS at our center.

Based on previous studies [20-23] and domain knowledge, the inclusion criteria for the study population were as follows: (1) patients with a pathological diagnosis of PTC in surgical paraffin-embedded specimens; this criterion was feasible because the specimen acquisition method is consistently documented in all pathology reports; (2) patients who underwent at least two thyroid ultrasound examinations prior to surgery at our center; and (3) patients who underwent ≥6 months of preoperative surveillance; (4) patients without surgical contraindications. Patients who presented with LNM or extrathyroidal extension (ETE) at baseline were excluded, as AS is generally not recommended for individuals with evidence of metastasis. For patients with multifocal PTC, only the lesion with the largest mean tumor diameter was considered for analysis [24]. Ultimately, based on the presence or absence of HT, the study population was divided into the HT group and the non-HT group.

Following relevant guidelines and previous studies [14,25], the criteria for diagnosing HT were defined as meeting at least one of the following conditions: (1) pathological evidence of HT in the peritumoral thyroid tissue on the postoperative pathology report or (2) thyroid function test results within 30 days before surgery showing thyroglobulin antibody ≥60 U/mL or ≥4.5 IU/mL, or thyroid peroxidase antibody ≥60 U/mL or ≥34 IU/mL.

Progression During Surveillance

The primary outcome of this study was progression-free survival, defined as the time from baseline to the progression

XSL•FO
RenderX

of PTC. Tumor progression was determined based on either of the following criteria [26] : (1) tumor enlargement, defined as an increase of ≥3 mm in any tumor diameter or (2) newly detected LNM during surveillance. In cases of suspected progression, original pathological, ultrasonographic, and surgical records were independently reviewed by senior clinicians and investigators to ensure accurate determination of disease progression.

Data Extraction

All authors involved in data entry for this study received standardized and targeted training to ensure consistency and comparability of the collected data. The demographic characteristics, imaging examinations, laboratory results, and pathological data extracted from the Electronic Medical Record system were entered and consolidated using Epidata software. To minimize subjective bias during data extraction, patient information was anonymized. Quality control was implemented at multiple stages. Initial and midterm audits were independently conducted by investigators: for each data entry personnel, 20 records were randomly selected and cross-checked against the original source documents to verify accuracy. After data entry was completed, a final spot-check review was performed by a senior physician to ensure the overall accuracy and reliability of the dataset. Data filtering was implemented in R (R Core Team) language according to the predetermined enrollment criteria.

Statistical Analyses

First, we compared baseline characteristics between the HT group and the non-HT group. For categorical variables, frequencies and percentages were reported, and differences between groups were assessed using the Pearson chi-square test or the Fisher exact test when the expected cell count was ≤5. For continuous variables, normality was tested using the Shapiro-Wilk test. Since the data did not follow a normal distribution, group differences were analyzed using the Mann-Whitney *U* test.

Next, to examine the association between HT and progression-free survival, we first tested the proportional hazards assumption using Schoenfeld residuals. We then constructed a multivariable Cox proportional hazards regression model, adjusting for potential confounders, including age, sex, baseline maximal tumor diameter, BMI, pregnancy status, number of tumor foci, and thyroid-stimulating hormone (TSH) level [27,28]. To visually illustrate the survival differences between the HT and non-HT groups, adjusted Kaplan-Meier PTC progression-free survival curves were plotted.

Further subgroup analyses were conducted based on median age (≥36 or <36 y), median TSH level (>1.72 or ≤1.72 µIU/mL), sex (male or female), baseline tumor size (the maximum diameter >1 or ≤1 cm), number of tumor foci (single or multiple), BMI (>24.0 or ≤24.0 kg/m^2), and pregnancy status (pregnant or not). We evaluated the modifying effects of these variables on the association between HT and progression-free survival using interaction terms (subgroup variable × HT/non-HT group) within the Cox regression model. In the subgroup analyses, covariates other than the stratifying variable were adjusted for accordingly. For example, when stratifying by sex, adjustments were made for age, TSH level, baseline tumor size, number of tumor foci, BMI, and pregnancy status. Similarly, when stratifying by age, adjustments were made for sex, TSH level, baseline tumor size, number of tumor foci, BMI, and pregnancy status.

All statistical analyses were performed using R software (version 4.4.3; R Foundation for Statistical Computing). A 2-sided *P* value of <.05 was considered statistically significant.

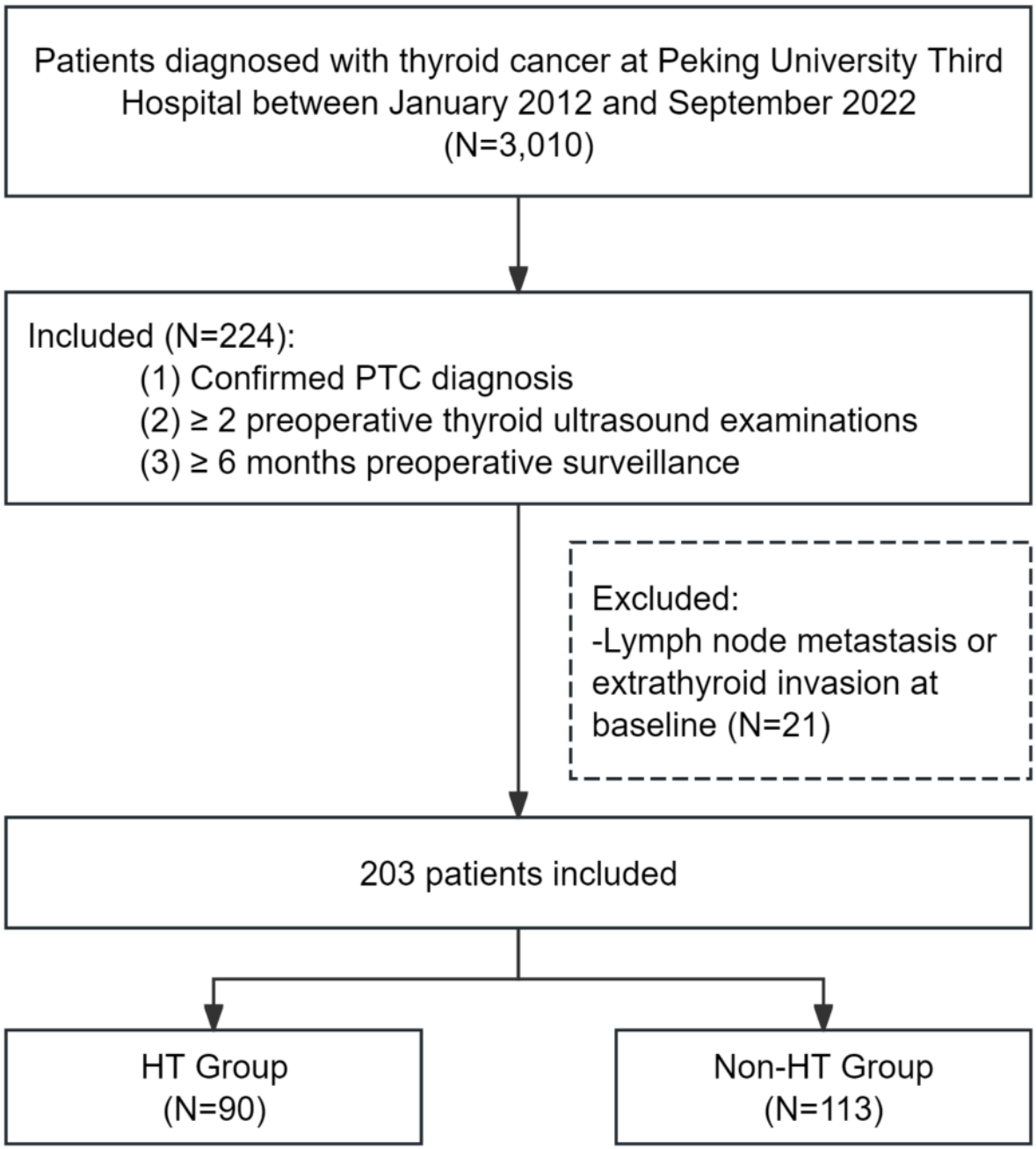
Results

Characteristics of the Study Population

A total of 203 patients with PTC under AS were included in the study. The median age was 36 (IQR 30-42) years, and 187 (92%) patients were female. The median preoperative surveillance duration was 1.45 (IQR 0.89-2.60) years. Patients were categorized into the HT group (n=90) and the non-HT group (n=113). The study flow diagram is presented in Figure 1, and the baseline characteristics of both groups are summarized in Table 1. Significant differences in age and sex were observed between the 2 groups (both *P*<.001). Table 2 shows the progression status of the study population.

XSL•FO
RenderX

Figure 1. Study flow diagram of the inclusion. HT: Hashimoto thyroiditis; PTC: papillary thyroid carcinoma.



XSL•FO
RenderX

Table . Baseline characteristics of the study population.

Characteristics	HT[a] group (n=90)	Non-HT group (n=113)	*P* value
Age (y), median (IQR)	34.5 (30.0-39.8)	39.0 (31.0-45.0)	<.001[b]
Baseline maximal tumor diameter (cm), median (IQR)	0.7 (0.6-1.1)	0.8 (0.6-1.2)	.41[b]
Surveillance duration (y), median (IQR)	1.53 (0.83-2.74)	1.44 (0.93-2.49)	.86[b]
Sex, n (%)			<.001[c]
Female	89 (98.9)	98 (86.7)	
BMI (kg/m^2), median (IQR)	23.4 (21.1-25.4)	23.4 (21.5-26.9)	.42[b]
Pregnancy status, n (%)			.47[d]
Yes	8 (8.9)	6 (5.3)	
Number of tumor foci, n (%)			.06[d]
Single	53 (58.9)	82 (72.6)	
Multiple	37 (41.1)	31 (27.4)	
Thyroid-stimulating hormone (TSH) level (μIU/mL), median (IQR)[e]	1.83 (1.12-2.56)	1.64 (1.26-2.18)	.29[b]

[a]HT: Hashimoto thyroiditis.

[b]Mann-Whitney *U* test.

[c]Fisher exact test.

[d]Pearson χ^2 test with Yates' continuity correction.

[e]TSH levels were missing in 93 patients during the observation period.

Table . Progression status of the study population.

Progression type	HT[a] group (n=90)	Non-HT group (n=113)	*P* value
PTC[b] progression, n (%)			
Yes	30 (33.3)	32 (28.3)	.54[c]
Tumor enlargement, n (%)			
Yes	28 (31.1)	31 (27.4)	.68[c]
Lymph node metastasis, n (%)			
Yes	6 (6.7)	3 (2.7)	.19[d]

[a]HT: Hashimoto thyroiditis.

[b]PTC: papillary thyroid carcinoma.

[c]Pearson χ^2 test with Yates' continuity correction.

[d]Fisher exact test.

Association Between HT and Progression-Free Survival

The Schoenfeld residuals test indicated that our data met the proportional hazards assumption (*P*>.05). During the surveillance period, 63 (31.0%) patients experienced progression, with 30 (33.3%) patients in the HT group and 33 (29.2%) patients in the non-HT group. Among the 63 patients who progressed, 6 experienced both tumor enlargement and LNM, 54 had tumor enlargement only, and 3 had LNM only.

In the multivariable Cox regression analysis, after adjusting for age, sex, and baseline mean tumor diameter, no significant difference in PTC progression-free survival was observed between the HT and non-HT groups (hazard ratio [HR] 1.11, 95% CI 0.61 - 1.99; *P*=.74). Similarly, no significant differences were found in tumor enlargement-free survival (HR 1.02, 95% CI 0.56 - 1.86; *P*=.95) or LNM-free survival (HR 1.76, 95% CI 0.31 - 10.12; *P*=.52; Table 3).

The Kaplan-Meier curves for PTC progression-free survival revealed that the survival curves of the HT and non-HT groups were nearly identical, indicating no significant difference between the 2 groups (Figure 2).

XSL•FO
RenderX

Table . Association between HT[a] and progression-free survival.

Outcomes	HR[b, c] (95% CI)	*P* value
PTC[d] progression	1.11 (0.61-1.99)	.74
Tumor enlargement	1.02 (0.56-1.86)	.95
Lymph node metastasis	1.76 (0.31-10.12)	.52

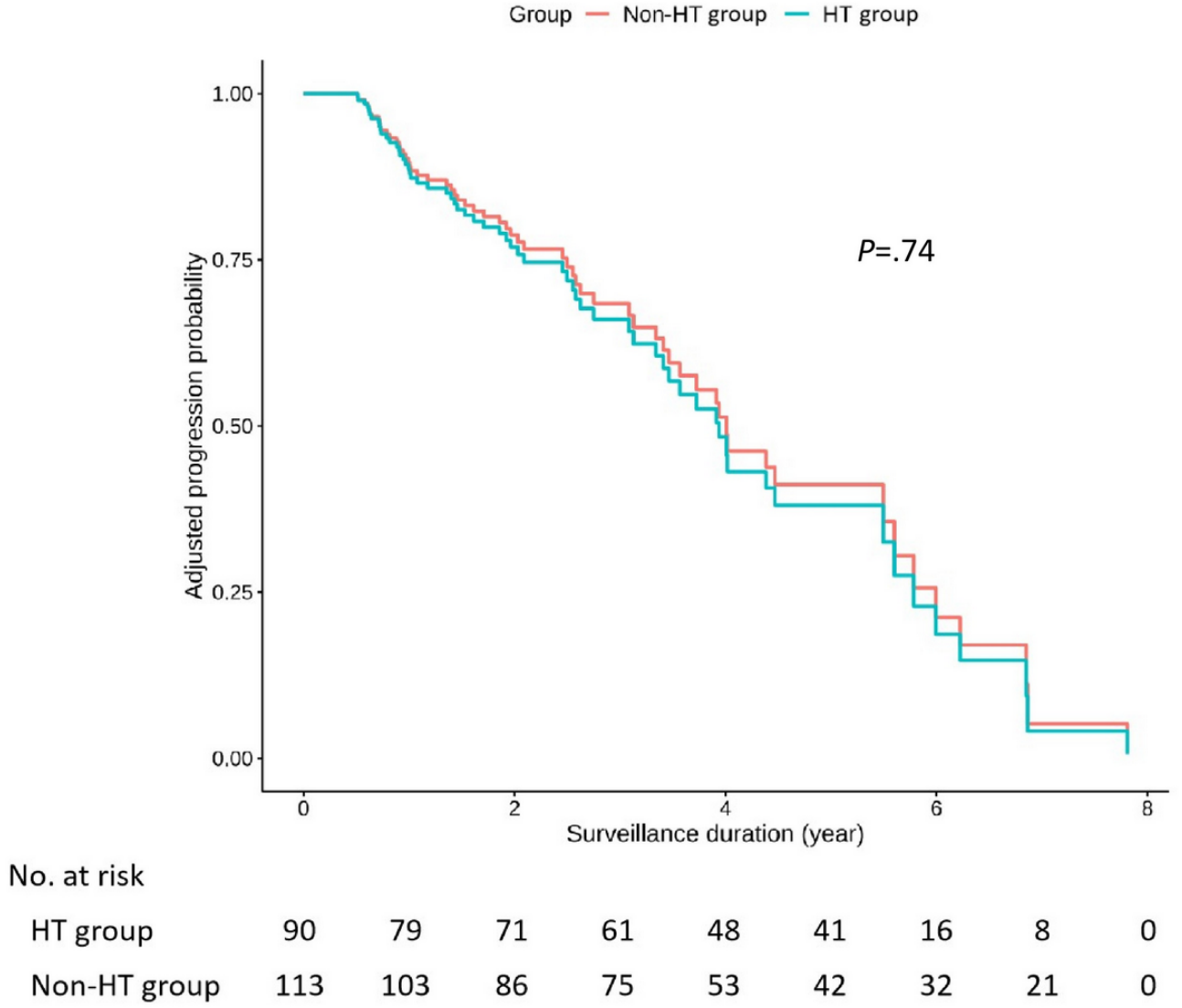
[a]HT: Hashimoto thyroiditis.

[b]HR: hazard ratio.

[c]Adjust for age, sex, baseline maximal tumor diameter, BMI, pregnancy status, number of tumor foci, and thyroid-stimulating hormone level.

[d]PTC: papillary thyroid carcinoma.

Figure 2. Kaplan-Meier papillary thyroid carcinoma progression-free survival curves. Adjust for age, sex, baseline maximal tumor diameter, BMI, pregnancy status, number of tumor foci, and thyroid-stimulating hormone level. HT: Hashimoto thyroiditis.



Sensitivity Analyses

In the sensitivity analyses, no significant difference in PTC progression-free survival was observed between the HT and non-HT groups (HR 1.23, 95% CI 0.60 - 2.53; *P*=.58). Similarly, no significant differences were found in tumor enlargement–free survival (HR 1.08, 95% CI 0.51 - 2.28; *P*=.84) or LNM-free survival (HR 2.44, 95% CI 0.23 - 25.23; *P*=.46). The findings from the sensitivity analyses were consistent with those of the primary analyses, suggesting the robustness of the results.

Subgroup Analyses

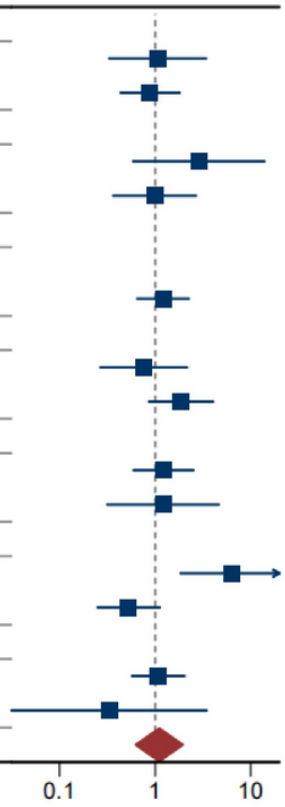
Figure 3 shows the forest plot of subgroup analyses. There was no significant difference in the association between HT and progression-free survival across subgroups classified by median age (36 y), median TSH level (1.72 µIU/mL), sex, baseline tumor size (the maximum diameter >1 or ≤1 cm), number of

XSL•FO
RenderX

tumor foci, or pregnancy status ($P_{interaction}$>.05). Notably, we found a potential interaction between HT and BMI ($P_{interaction}$<.01). Among patients who were overweight or obese (BMI >24 kg/m^2), HT was associated with a significantly higher risk of progression (HR 6.32, 95% CI 1.84 - 21.69; *P*=.003). In contrast, no association between HT and progression risk was observed in patients with BMI ≤24 kg/m^2 (*P*=.10).

Figure 3. Subgroup analysis of the association between Hashimoto thyroiditis (HT) and progression-free survival. TSH levels were missing in 93 patients during the observation period. In the male subgroup, only 1 patient had HT, and no progression was observed in this case; therefore, an estimate could not be provided. TSH: thyroid-stimulating hormone.

Subgroup	Number of cases	Hazard ratio (95% CI)	*P* value	*P* for interaction
Age				**.92**
≥36 years old	100	1.06 (0.33-3.42)	.92	
<36 years old	103	0.89 (0.43-1.81)	.74	
TSH				**.37**
>1.72 uIU/mL	55	2.87 (0.593-13.99)	.19	
≤ 1.72 uIU/mL	55	0.99 (0.36-2.68)	.97	
Sex				**≥.99**
Male	16	-		
Female	187	1.21 (0.65-2.25)	.55	
Tumor size at baseline				**.45**
>1cm	55	0.76 (0.27-2.16)	.61	
≤ 1cm	148	1.87 (0.86-4.06)	.11	
Number of tumor foci				**.67**
Single	135	1.23 (0.60-2.53)	.58	
Multiple	68	1.21 (0.31-4.66)	.78	
BMI				**.001**
>24.0 kg/m²	85	6.32 (1.84-21.69)	.003	
≤ 24.0 kg/m²	118	0.53 (0.25-1.12)	.10	
Pregnant				**.95**
Yes	14	1.08 (0.57-2.04)	.82	
No	189	0.33 (0.03-3.47)	.36	
Overall	**203**	**1.11 (0.61-1.99)**	**.74**	



Discussion

Summary of the Findings

Over a median surveillance period of 1.45 years, our analysis demonstrated no significant association between HT and PTC progression under AS. This finding was consistent for both tumor enlargement and LNM. The results remained consistent in the sensitivity analysis. In the subgroup analyses, no significant association between HT and progression-free survival was found across strata defined by age, TSH level, sex, baseline tumor size, number of tumor foci, or pregnancy status. However, a potential interaction between HT and BMI was identified. Among patients who were overweight or obese (BMI >24 kg/m²), HT was significantly associated with an increased risk of progression.

These results suggest that HT is not independently associated with PTC progression during AS in the general population. However, our subgroup analysis revealed a potential interaction between HT and BMI, suggesting that the association between HT and PTC progression may be modified by overweight or obesity status. Overweight is an increasingly recognized clinical condition. The association between obesity and ETE, tumor multifocality, larger tumor size, as well as LNM remains controversial. However, it has been established that a positive correlation exists between elevated BMI and the presence of the BRAFV600E mutation [29]. Although HT has been proven to be a factor for a better prognosis, this study confirms that, compared to other clinical indicators, an association between HT and PTC progression can be observed specifically within the overweight subgroup in this study. Although the current understanding of the mechanisms underlying the interaction between obesity and HT in PTC progression remains unclear, we speculate that their association may involve the following aspects: (1) HT is associated with a lower BRAF V600E mutation rate [14], while obesity is associated with a higher BRAF V600E mutation rate [30,31]. When both conditions coexist, their opposing effects on gene mutation might offset each other. (2) HT is characterized by lymphocytic infiltration in thyroid tissue. Patients with HT and PTC have been observed to have reduced Tregs and increased interleukin 10 secretion [32]. Conversely, obesity induces Treg expansion and elevated interleukin 10 [33], which may indirectly promote tumor immune escape by suppressing the cytotoxic functions of CD8$^+$ T cells and natural killer cells, thereby counteracting the influence of HT on PTC. (3) Regarding metabolism, in the state of obesity, increased levels of free fatty acids provide more energy for tumor cells.

XSL•FO
RenderX

Further studies with larger cohorts and mechanistic investigations were warranted to validate and elucidate this interaction.

Comparison With Previous Studies

Previous studies have provided substantial evidence supporting a potential association between HT and PTC prognosis. For example, Marotta et al [34] conducted a multicenter retrospective cohort study in Italy involving 301 patients with PTC, of whom 42.5% had coexisting HT and reported significantly longer recurrence-free survival in those with HT. Similarly, Xu et al [14] performed a single-center retrospective cohort study in China involving 9210 patients with PTC, with a 19% prevalence of coexistent HT. Xu et al also suggested that HT was associated with better PTC prognosis.

Nevertheless, the association between HT and postoperative recurrence may differ substantially from its relationship with disease progression under AS. Yet, this distinction has been largely overlooked, as few studies have specifically investigated the impact of HT on PTC progression during AS. In a Korean cohort of 699 patients with PTC managed with AS, Lee et al [17] reported that tumor progression was associated with diffuse thyroid disease (DTD) as detected by ultrasound, with a 2.3-fold increased risk of progression in patients with HT compared to those without. Interestingly, although our study did not find a statistically significant association between HT and PTC progression under AS in the general population, we observed a 6.32-fold increased risk of progression associated with HT in patients who were overweight or obese. It is important to highlight that our study design differs from that of Lee et al [17]. First, HT represents only one type of DTD, which also includes other conditions such as Graves' disease and simple goiter. Second, according to Lee et al [17], DTD was diagnosed based on preoperative ultrasound examinations, while in our study, the diagnosis of HT was primarily based on postoperative pathological confirmation and preoperative serum antibody levels measured within 30 days before surgery.

Limitations and Strengths

Our study has several strengths. First, we applied strict eligibility criteria, including only patients who had undergone AS for more than 6 months preoperatively, to ensure that sufficient follow-up time was available for detecting clinically meaningful changes. The relatively low enrollment proportion is attributable to 2 main factors: (1) AS has only been widely adopted in recent years, and many preoperative cases did not undergo this protocol. (2) Some nonlocal patients or those seeking examinations at other hospitals underwent only a single ultrasound at our institution prior to surgery, making them largely ineligible. We also excluded patients with baseline LNM or ETE to reduce confounding by indication. Second, we employed Cox proportional hazards models to fully utilize time-to-event data, providing more robust estimates. Third, all outcome events were individually reviewed by experienced clinicians to ensure the accuracy of progression assessment. Fourth, beyond the overall analysis, we examined the association between HT and specific types of progression (including tumor enlargement and LNM) and conducted sensitivity analyses to assess the robustness of our findings. Furthermore, we investigated potential interactions and associations across various patient subgroups.

Nevertheless, our results should be interpreted with caution. First, this was a retrospective study based on data extracted from electronic medical records. We plan to conduct a well-designed, prospectively followed cohort study with standardized data collection to validate our findings. Second, in clinical practice, the decision to undergo AS is influenced not only by the biological risk of tumor progression but also by patient preferences. For instance, patients with significant anxiety may opt for surgery even if their clinical condition qualifies for observation. Such preference-based decisions could introduce selection bias in our study population. Third, as a single-center study, the generalizability of our findings may be limited. Fourth, the limited sample size restricted our adjustment to only a constrained set of confounders. Future studies with larger cohorts are needed to control for a broader range of potential confounding factors. Fifth, the relatively short follow-up period may have limited our ability to detect long-term disease progression dynamics. In several instances, surgical intervention occurred before tumor progression could be observed, which could introduce observational time bias and potentially affect the accuracy of outcome assessments.

These findings should be regarded as hypothesis-generating and may help inform the design of future prospective studies with longer surveillance and more comprehensive data collection.

Clinical Implications

Given the intensive resources required to conduct prospective cohort studies on AS, there are currently very few studies addressing this issue globally. Our study provides a foundation for future research and has important implications for clinical practice. With a methodologically sound design, we contribute relatively high-quality evidence suggesting that, in general, patients with PTC with coexisting HT who meet clinical criteria for AS do not face an increased risk of disease progression. However, among patients with HT who are overweight or obese (BMI >24 kg/m^2), particular attention should be paid to a potentially elevated risk of PTC progression.

Conclusion

The findings of our retrospective cohort study implied that for patients with PTC, the presence of HT alone might not be a major factor in determining eligibility for AS versus immediate surgery, except in cases where the patients were overweight or obese. We advocate for future prospective cohort studies on AS to validate these findings before they are applied to clinical decision-making.

XSL•FO
RenderX

Acknowledgments

The generative artificial intelligence (GAI) tool used was DeepSeek-V3.1-Terminus. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

ZL was supported by the National Natural Science Foundation of China (82373694), Beijing Municipal Natural Science Foundation (7262165), Young Elite Scientists Sponsorship Program by CAST (2023QNRC001), Beijing Education Sciences Planning Program during the 14th Five-Year Plan (BECA23111), and the Fundamental Research Funds for the Central Universities (BMU2021YJ030). Other authors received no financial support for this research.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request via email.

Authors' Contributions

Conceptualization: ZL (lead)
Data curation: BS (supporting), CY (supporting), FM (equal), SS (supporting)
Formal analysis: RS (equal), XL (equal)
Review and editing: BS (supporting), CY (supporting), FM (supporting), SS (supporting)
Writing – original draft: XL (equal), XY (equal)
Writing – review & editing: RS (supporting), XL (equal), XY (equal), ZL (lead)

Conflicts of Interest

None declared.

References

1. Deng Y, Li H, Wang M, et al. Global burden of thyroid cancer from 1990 to 2017. JAMA Netw Open 2020 Jun 1;3(6):e208759. [doi: 10.1001/jamanetworkopen.2020.8759] [Medline: 32589231]
2. Pizzato M, Li M, Vignat J, et al. The epidemiological landscape of thyroid cancer worldwide: GLOBOCAN estimates for incidence and mortality rates in 2020. Lancet Diabetes Endocrinol 2022 Apr;10(4):264-272. [doi: 10.1016/S2213-8587(22)00035-3] [Medline: 35271818]
3. Osborne D, Choudhary R, Vyas A, et al. Hashimoto's thyroiditis effects on papillary thyroid carcinoma outcomes: a systematic review. Cureus 2022 Aug;14(8):e28054. [doi: 10.7759/cureus.28054] [Medline: 36120263]
4. Wang J, Yu F, Shang Y, Ping Z, Liu L. Thyroid cancer: incidence and mortality trends in China, 2005-2015. Endocrine 2020 Apr;68(1):163-173. [doi: 10.1007/s12020-020-02207-6] [Medline: 32002755]
5. Han B, Zheng R, Zeng H, et al. Cancer incidence and mortality in China, 2022. J Natl Cancer Cent 2024 Mar;4(1):47-53. [doi: 10.1016/j.jncc.2024.01.006] [Medline: 39036382]
6. Lin RX, Yang SL, Jia Y, Wu JC, Xu Z, Zhang H. Epigenetic regulation of papillary thyroid carcinoma by long non-coding RNAs. Semin Cancer Biol 2022 Aug;83:253-260. [doi: 10.1016/j.semcancer.2021.03.027] [Medline: 33785446]
7. Tuttle RM, Fagin JA, Minkowitz G, et al. Natural history and tumor volume kinetics of papillary thyroid cancers during active surveillance. JAMA Otolaryngol Head Neck Surg 2017 Oct 1;143(10):1015-1020. [doi: 10.1001/jamaoto.2017.1442] [Medline: 28859191]
8. Kim HI, Jang HW, Ahn HS, et al. High serum TSH level is associated with progression of papillary thyroid microcarcinoma during active surveillance. J Clin Endocrinol Metab 2018 Feb 1;103(2):446-451. [doi: 10.1210/jc.2017-01775] [Medline: 29211863]
9. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association Guidelines Task Force on thyroid nodules and differentiated thyroid cancer. Thyroid 2016 Jan;26(1):1-133. [doi: 10.1089/thy.2015.0020] [Medline: 26462967]
10. Endocrinology Branch of the Chinese Medical Association, Thyroid and Metabolic Surgery Group, Surgery Branch of Chinese Medical Association, Head and Neck Tumor Professional Committee of the Chinese Anti-Cancer Association, Nuclear Medicine Branch of the Chinese Medical Association, Thyroid Cancer Professional Committee of the Chinese Anti-Cancer Association, Thyroid Surgeon Committee of the Surgeon Branch of the Chinese Medical Doctor Association, Ultrasound Medicine Branch of the Chinese Medical Association. Guidelines for the diagnosis and management of thyroid nodules and differentiated thyroid cancer (Second edition). Chin J Endocrinol Metab 2023;39(3):181-226. [doi: 10.3760/cma.j.cn311282-20221023-00589]
11. NCCN clinical practice guidelines in oncology-thyroid carcinoma (version 1). National Comprehensive Cancer Network. 2025. URL: https://www.nccn.org/guidelines/category_1 [accessed 2025-12-15]

XSL•FO
RenderX

12. Kaur J, Jialal I. Hashimoto thyroiditis. In: StatPearls: StatPearls Publishing; 2024. URL: https://www.ncbi.nlm.nih.gov/books/NBK459262/ [accessed 2025-12-18] [Medline: 29083758]
13. Lee JH, Kim Y, Choi JW, Kim YS. The association between papillary thyroid carcinoma and histologically proven Hashimoto's thyroiditis: a meta-analysis. Eur J Endocrinol 2013 Mar;168(3):343-349. [doi: 10.1530/EJE-12-0903] [Medline: 23211578]
14. Xu S, Huang H, Qian J, et al. Prevalence of Hashimoto thyroiditis in adults with papillary thyroid cancer and its association with cancer recurrence and outcomes. JAMA Netw Open 2021 Jul 1;4(7):e2118526. [doi: 10.1001/jamanetworkopen.2021.18526] [Medline: 34313737]
15. Caturegli P, De Remigis A, Chuang K, Dembele M, Iwama A, Iwama S. Hashimoto's thyroiditis: celebrating the centennial through the lens of the Johns Hopkins hospital surgical pathology records. Thyroid 2013 Feb;23(2):142-150. [doi: 10.1089/thy.2012.0554] [Medline: 23151083]
16. Xu J, Ding K, Mu L, et al. Hashimoto's thyroiditis: a "double-edged sword" in thyroid carcinoma. Front Endocrinol (Lausanne) 2022;13:801925. [doi: 10.3389/fendo.2022.801925] [Medline: 35282434]
17. Lee JY, Kim JH, Kim YK, et al. US predictors of papillary thyroid microcarcinoma progression at active surveillance. Radiology 2023 Oct;309(1):e230006. [doi: 10.1148/radiol.230006] [Medline: 37906009]
18. Koshkina A, Fazelzad R, Sugitani I, et al. Association of patient age with progression of low-risk papillary thyroid carcinoma under active surveillance: a systematic review and meta-analysis. JAMA Otolaryngol Head Neck Surg 2020 Jun 1;146(6):552-560. [doi: 10.1001/jamaoto.2020.0368] [Medline: 32297926]
19. Lee EK, Moon JH, Hwangbo Y, et al. Progression of low-risk papillary thyroid microcarcinoma during active surveillance: interim analysis of a multicenter prospective cohort study of active surveillance on papillary thyroid microcarcinoma in Korea. Thyroid 2022 Nov;32(11):1328-1336. [doi: 10.1089/thy.2021.0614] [Medline: 36205563]
20. Xiao WC, Li X, Shan R, et al. Pregnancy and progression of differentiated thyroid cancer: a propensity score-matched retrospective cohort study. J Clin Endocrinol Metab 2024 Feb 20;109(3):837-843. [doi: 10.1210/clinem/dgad557] [Medline: 37738427]
21. Oda H, Miyauchi A, Ito Y, et al. Incidences of unfavorable events in the management of low-risk papillary microcarcinoma of the thyroid by active surveillance versus immediate surgery. Thyroid 2016 Jan;26(1):150-155. [doi: 10.1089/thy.2015.0313] [Medline: 26426735]
22. Li X, Xiao WC, Mei F, et al. The association of pregnancy with disease progression in patients previously treated for differentiated thyroid cancer: a propensity score-matched retrospective cohort study. J Womens Health (Larchmt) 2023 Nov;32(11):1174-1181. [doi: 10.1089/jwh.2023.0172] [Medline: 37638829]
23. Ringel MD, Sosa JA, Baloch Z, et al. 2025 American Thyroid Association management guidelines for adult patients with differentiated thyroid cancer. Thyroid 2025 Aug;35(8):841-985. [doi: 10.1177/10507256251363120] [Medline: 40844370]
24. Hwang H, Choi JY, Yu HW, et al. Surgical outcomes in patients with low-risk papillary thyroid microcarcinoma from MAeSTro study: immediate operation versus delayed operation after active surveillance: a multicenter prospective cohort study. Ann Surg 2023 Nov 1;278(5):e1087-e1095. [doi: 10.1097/SLA.0000000000005841] [Medline: 36912439]
25. Ralli M, Angeletti D, Fiore M, et al. Hashimoto's thyroiditis: an update on pathogenic mechanisms, diagnostic protocols, therapeutic strategies, and potential malignant transformation. Autoimmun Rev 2020 Oct;19(10):102649. [doi: 10.1016/j.autrev.2020.102649] [Medline: 32805423]
26. Ito Y, Miyauchi A. A therapeutic strategy for incidentally detected papillary microcarcinoma of the thyroid. Nat Clin Pract Endocrinol Metab 2007 Mar;3(3):240-248. [doi: 10.1038/ncpendmet0428] [Medline: 17315032]
27. Sugitani I, Ito Y, Takeuchi D, et al. Indications and strategy for active surveillance of adult low-risk papillary thyroid microcarcinoma: consensus statements from the Japan Association of Endocrine Surgery Task Force on management for papillary thyroid microcarcinoma. Thyroid 2021 Feb;31(2):183-192. [doi: 10.1089/thy.2020.0330] [Medline: 33023426]
28. Ito Y, Miyauchi A, Fujishima M, et al. Thyroid-stimulating hormone, age, and tumor size are risk factors for progression during active surveillance of low-risk papillary thyroid microcarcinoma in adults. World J Surg 2023 Feb;47(2):392-401. [doi: 10.1007/s00268-022-06770-z] [Medline: 36182976]
29. Lee J, Lee CR, Ku CR, et al. Association between obesity and BRAFV600E mutation status in patients with papillary thyroid cancer. Ann Surg Oncol 2015 Dec;22(S3):683-690. [doi: 10.1245/s10434-015-4765-z]
30. Kim WW, Ha TK, Bae SK. Clinical implications of the BRAF mutation in papillary thyroid carcinoma and chronic lymphocytic thyroiditis. J Otolaryngol Head Neck Surg 2018 Jan 9;47(1):4. [doi: 10.1186/s40463-017-0247-6] [Medline: 29316976]
31. Lee I, Kim HK, Soh EY, Lee J. The association between chronic lymphocytic thyroiditis and the progress of papillary thyroid cancer. World J Surg 2020 May;44(5):1506-1513. [doi: 10.1007/s00268-019-05337-9] [Medline: 31915977]
32. Zhao N, Liu X, Wu C, Liu Y, Zhao X, He X. Changes in Treg numbers and activity in papillary thyroid carcinoma with and without Hashimoto's thyroiditis. J Int Med Res 2020 Apr;48(4):300060520919222. [doi: 10.1177/0300060520919222] [Medline: 32314639]
33. Prakash S, Priyatma, Aasarey R, et al. Role of prokineticins and T-reg cells in obesity-associated metabolic oxidative dysregulation in NAFLD. Sci Rep 2025 Aug 12;15(1):29470. [doi: 10.1038/s41598-025-97969-2] [Medline: 40797035]

XSL•FO
RenderX

34. Marotta V, Sciammarella C, Chiofalo MG, et al. Hashimoto's thyroiditis predicts outcome in intrathyroidal papillary thyroid cancer. Endocr Relat Cancer 2017 Sep;24(9):485-493. [doi: 10.1530/ERC-17-0085] [Medline: 28696209]

Abbreviations

AS: active surveillance
DTD: diffuse thyroid disease
ETE: extrathyroidal extension
HR: hazard ratio
HT: Hashimoto thyroiditis
LNM: lymph node metastasis
PTC: papillary thyroid cancer
TSH: thyroid-stimulating hormone

Edited by N Cahill; submitted 11.Jul.2025; peer-reviewed by TS Sevimli, X Zhu; accepted 01.Dec.2025; published 06.Jan.2026.

Please cite as:
Li X, Yao X, Shan R, Mei F, Song S, Sun B, Yuan C, Liu Z
Hashimoto Thyroiditis and Progression of Papillary Thyroid Cancer: 10-Year Retrospective Cohort Study
JMIR Cancer 2026;12:e80535
URL: https://cancer.jmir.org/2026/1/e80535
doi:10.2196/80535

© Xin Li, Xiangyun Yao, Rui Shan, Fang Mei, Shibing Song, Bangkai Sun, Chunhui Yuan, Zheng Liu. Originally published in JMIR Cancer (https://cancer.jmir.org), 6.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Experiences and Perspectives of Patients With Early Breast Cancer, Caregivers, and Health Care Professionals: Italian Social Media Listening Study

Matteo Basilio Suter[1*], MD; Rosanna D'Antona[2*], PG; Manuelita Mazza[3*], MD; Enrica Francavilla[1*], MiMS; Alice Messi[1*], MA; Abhinav Maurya[4*], MBA; Shiva Padhi[4*], MBA; Diletta Valsecchi[1*], MS

[1]Novartis Farma S.p.A., Viale Luigi Sturzo, 43 - 20154, Milan, Italy

[2]Europa Donna Italia, Milan, Italy

[3]Cancer Institute Fatima bint Mubarak Center - Cleveland Clinic Abu Dhabi, Abu Dhabi, United Arab Emirates

[4]Novartis Healthcare Pvt. Ltd., Hyderabad, India

[*]all authors contributed equally

Corresponding Author:
Diletta Valsecchi, MS
Novartis Farma S.p.A., Viale Luigi Sturzo, 43 - 20154, Milan, Italy

Abstract

Background: Published evidence on patient experiences, perceptions, and challenges related to early breast cancer (eBC) in Italy is limited. Understanding these aspects is critical for improving diagnosis, treatment outcomes, and quality of life (QoL).

Objective: This study used social media listening (SML) to explore the patient journey, treatment perceptions, QoL, and unmet needs of patients with eBC, caregivers, and health care professionals (HCPs) in Italy.

Methods: This retrospective noninterventional SML study analyzed publicly available posts from December 2021 to November 2023 using breast cancer–related keywords in English and Italian through Sprinklr, a web-based aggregator tool. Posts sourced from social media platforms, such as X (formerly known as Twitter), blogs, forums, Facebook, Instagram, and YouTube, were filtered by geographic location to include only users in Italy. Posts were filtered using natural language processing (NLP) for relevance and duplicates, followed by manual review and stakeholder identification (patients, caregivers, and HCPs). Key themes of discussion were identified through thematic analysis of posts across the stages of the patient journey (symptoms, diagnosis, treatment, etc). Ethical guidelines were followed by using anonymized, publicly available data. Descriptive statistics were used to analyze the data, and posts with missing data were excluded. Consequently, denominators varied across analyses and were adjusted based on data availability for specific variables.

Results: Of the 20,008 posts initially extracted, 1580 posts were retained following NLP filtering, and 530 posts were included after manual screening. The majority (493/518, 95%) of the posts were sharing information about diagnosis and treatment journeys, emotional challenges, QoL concerns, and symptoms (eg, lumps, breast pain), while 27% (141/518) of the posts sought information on diagnostic dilemmas, treatment options, and second opinions. Patients contributed 60% (318/530) of the posts, and caregivers contributed 21% (111/530) of the posts, with over half (57/107, 53%) discussing their mothers' diagnosis and treatment struggles. HCPs contributed 16% (85/530) of the posts, primarily sharing clinical trial updates, drug approvals, and disease awareness efforts. A total of 88 posts included discussions on QoL, and eBC significantly impacted patients' emotional, physical, functional, and social well-being. Discussions revealed key unmet needs, including limited awareness of adjuvant therapy options, lack of peer support groups, suboptimal patient-HCP communication, and insufficient access to specialty care facilities.

Conclusions: This study highlights gaps in eBC management related to patient education, HCP communication, and access to specialty care and describes an associated worsening of QoL for patients as reflected in social media posts. Within the limitations of an observational SML design, increasing patient and caregiver awareness of available adjuvant therapies to improve adherence and reduce recurrence risk, alongside expanding access to regional breast cancer centers, may help optimize patient experiences and outcomes. Further research using complementary data sources is needed to confirm and extend these findings.

(JMIR Cancer 2026;12:e73371) doi:10.2196/73371

KEYWORDS

early breast cancer experiences; social media listening; social media analysis; patient journey mapping; quality of life in cancer; diagnosis and treatment challenges; health care professional; patient-HCP communication; emotional and physical well-being; unmet needs in cancer care; breast cancer recurrence

XSL•FO
RenderX

Introduction

In the year 2022, female breast cancer (BC) was projected to be the most prevalent form of cancer globally, accounting for 2.3 million new cases, which constitutes 11.6% of all cancer diagnoses [1]. It is also the leading cause of death among women, accounting for 670,000 deaths globally, and ranked fourth in terms of cancer-related mortality worldwide [1]. Similarly, BC is the most commonly diagnosed cancer among women in Italy. It was estimated that there would be a total of 55,900 new cases in 2023, with over 90% of them being detected in the early stages [2,3].

Despite the availability of various treatment options for early BC (eBC), a significant number of patients still experience recurrence. For luminal BC, there is a long-term cumulative risk, with Italy reporting up to 10.5% distant recurrence within the first 5 years [4].

Patients with BC encounter numerous challenges throughout their journey, from identifying symptoms to receiving a diagnosis, undergoing treatment, and managing the disease [5-7]. Early diagnosis and effective management of BC are crucial in improving patient outcomes and reducing health care expenses [8,9]. To understand the diverse challenges of patients with eBC, conventional methods, such as questionnaires, interviews, and patient-reported outcome measures, are often used to gather qualitative data. However, these studies frequently involve a limited number of participants or require extended durations to be conducted, resulting in a limited understanding of the unique obstacles faced by individual patients [10-13]. Recently, social media listening (SML) has gained prominence as an approach to gain deeper insights into these challenges [14-16].

High internet penetration and presence of various social media platforms have greatly influenced health care [17,18]. Social media platforms have become valuable sources of information, allowing patients to share their beliefs about diseases, treatment experiences, satisfaction with outcomes, and other factors that impact their lives [19].

SML is a new approach for gathering information from social media platforms and can be useful in generating insights from user experiences. SML has been used to monitor discussions on health-related topics in various cancers [16,19-25]. However, there is limited published literature on the use of SML to investigate the needs and experiences of patients with eBC [15]. This study explores SML as a research tool to provide insights into the journey of patients with eBC.

This study aimed to understand the patient journey, treatment perceptions, quality of life (QoL), unmet needs, and concerns of patients with eBC by analyzing the social media posts from patients, caregivers, and health care professionals (HCPs).

Methods

Study Design and Data Source

This study was a noninterventional retrospective analysis of social media data available in the public domain. A comprehensive search was conducted on various social media platforms, including X (formerly known as Twitter), blogs, forums, Facebook community groups, Instagram, and YouTube, from December 2021 to November 2023 using keywords related to BC in both Italian and English languages. The initial keyword list was developed by a multidisciplinary study team comprising researchers with expertise in oncology, social media analytics, and qualitative research. The team first compiled candidate terms based on clinical terminology (eg, “early breast cancer,” “chemotherapy”) and commonly used lay terms (eg, “chemo,” “cancer journey,” “tumor,” “lump”), informed by prior SML and patient-experience literature [19,26]. Preliminary scoping searches were then conducted in the Sprinklr tool to test and refine this list; identify additional relevant terms and hashtags used by patients, caregivers, and HCPs; and optimize precision and recall. Boolean operators (eg, AND, OR) were used to structure the search strings, and all terms were translated into both Italian and English to reflect the bilingual nature of online discourse in Italy. This iterative, expert-informed process aligns with best practices in SML research [19,27,28] and was used to finalize the search strategy prior to full data extraction. The data have been reported in accordance with the SRQR (Standards for Reporting Qualitative Research) checklist (Checklist 1). The specific search terms and list of forums and blogs are reported in Multimedia Appendix 1.

Each social media platform was selected based on its potential to capture unique dimensions of the patient, caregiver, and HCP experience. The inclusion of multiple platforms allowed this study to triangulate insights from different types of user engagement:

- X: Real-time updates, public sentiment, and awareness campaigns
- Facebook groups: Peer support, community discussions, and shared experiences
- Forums: Anonymous, in-depth discussions among patients and caregivers
- Blogs: Detailed personal narratives and reflections on the cancer journey
- Instagram and YouTube: Visual storytelling and video-based sharing of experiences

While blogs are not inherently discussion-based, they offer rich, first-person accounts of the lived experience of illness. Prior SML studies have consistently included blogs as valuable sources of qualitative data [19,26,27].

Forums and blogs were identified through Sprinklr’s indexing of publicly accessible websites and were selected if they (1) hosted user-generated content related to BC, (2) were accessible without registration or password protection, and (3) had contributors whose profiles or content indicated an Italian context (eg, language, self-reported location).

Eligibility Criteria

This study included posts that (1) were publicly available (ie, not private, password-protected, or restricted content); (2) originated from users located in Italy (based on profile or geolocation information); (3) contained 1 or more keywords related to BC; and (4) appeared to reflect the experiences,

XSL•FO
RenderX

perceptions, questions, or opinions of patients with eBC, their caregivers, or HCPs involved in BC care. This study excluded (1) posts that were clearly unrelated to BC, despite containing 1 or more keywords (eg, spam, advertisements, financial content); (2) posts from news portals, media outlets, and scientific journals; (3) posts with commercial or promotional content; and (4) duplicate post links. These eligibility criteria were applied consistently across all platforms and are reported in detail in Multimedia Appendix 1.

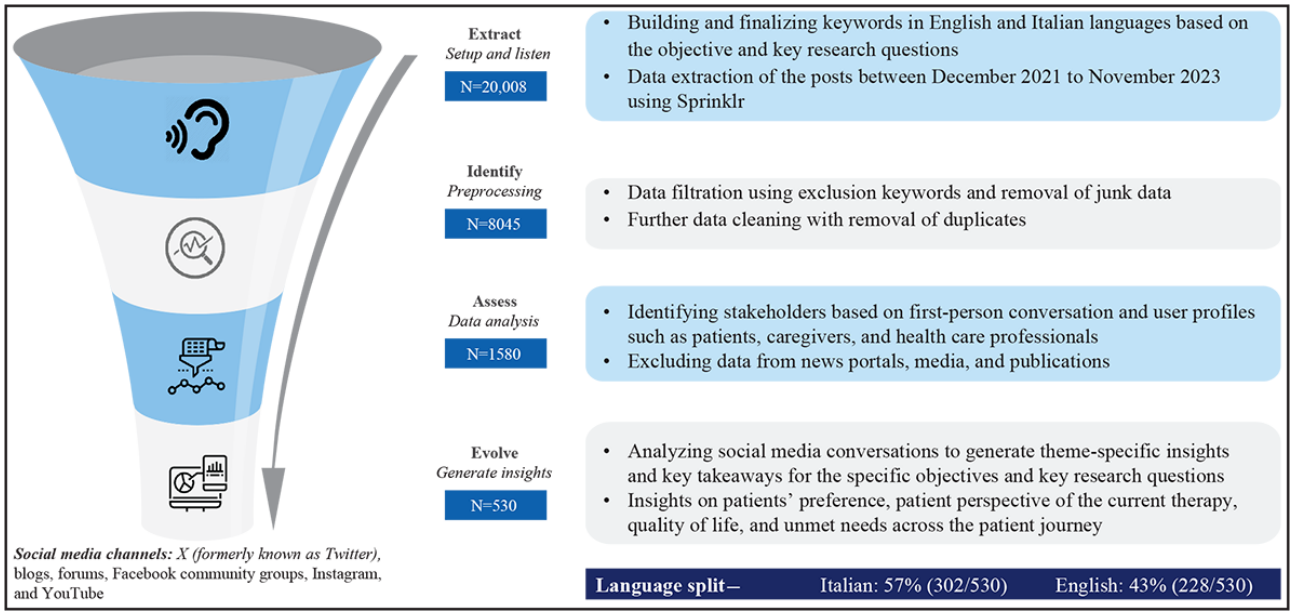
Data Collection, Selection of Posts, and Data Cleaning

Data from public posts were collected using Sprinklr [29], a web-based aggregator tool. This tool allows downloading of posts along with the date stamp and geographic location information of the users. Based on the geographic location of the user profile, only posts from Italy were downloaded. Each post was assigned a unique ID by the tool. Data filtration techniques were applied using exclusion keywords to remove junk data to maintain data quality using natural language processing (NLP). Furthermore, NLP was used to clean the data and eliminate duplicates for reliable analysis.

Data Analysis

Manual screening was conducted to analyze the data. Each post was indexed to identify the stakeholders based on first-person conversations and user profiles, including patients, caregivers, and HCPs. Posts related to news portals, media, and publications were excluded to focus solely on social media posts from patients, caregivers, and HCPs. The analysis included examining the posts based on the social media channel (X, blogs, forums, Facebook community groups, Instagram, and YouTube), the tone of the posts (positive, neutral, or negative), and the key themes of discussion (Figure 1). The themes were identified based on the type of information relevant to the study objective discussed in the posts. A single post could be categorized under more than 1 theme based on the information mentioned in the post.

Figure 1. Social media listening (SML) methodology: from extraction of posts to insights generation. "N" denotes number of posts.



The thematic analysis in this study followed a predominantly deductive approach. The themes were derived based on a conceptual framework aligned with the typical patient journey for individuals diagnosed with eBC. This framework included key stages such as symptom onset, diagnosis, treatment, disease management, and quality of life. Additionally, this study incorporated themes related to online patient behavior, such as sharing disease awareness and discussing clinical trial or research updates, which are commonly observed in social media discourse among patients, caregivers, and HCPs.

The deductive nature of the thematic analysis was guided by predefined categories informed by existing literature on the eBC patient journey and prior SML studies [19,28]. These categories were used to structure the coding process and ensure alignment with the study objectives. While the primary approach was deductive, the analysis remained open to emergent subthemes within each category during manual screening, allowing for flexibility in capturing nuanced insights. Each post was coded and analyzed by a qualified researcher with expertise in oncology, social media analytics, and qualitative research.

The theoretical underpinning of this approach draws from the patient-centered care model and the concept of the illness trajectory, which emphasizes understanding the lived experiences of patients across different stages of their condition. This framework supports the identification of themes that reflect both clinical milestones and psychosocial dimensions of the patient experience. The review and coding were led by a primary analyst (A Maurya), who has expertise in oncology, social media analytics, and qualitative research, using structured spreadsheets and a predefined codebook. To enhance the reliability of the reviewing and coding process, we applied techniques, such as iterative codebook refinement and regular team discussions, to review preliminary themes and ensure consistency with the study objectives. Data analysis and interpretation were led by A Maurya in collaboration with SP, with input from the other

XSL•FO
RenderX

coauthors, who contributed to the refinement of themes and final interpretation of the findings.

For analysis and reporting, the terms "post" and "mention" were used. The term "post" referred to a single post. The term "mention" referred to the number of times a symptom, treatment, diagnostic test, or other parameter was mentioned in each post. It is important to note that the number of mentions was independent of the number of posts. The unit of analysis is posts and mentions, not unique individuals, and due to data anonymization, users could not be tracked across different platforms. Descriptive statistics were used to analyze the data, which were presented in the form of post counts (N), mention counts (n), and percentages (%). The percentage calculation was performed by dividing the number of posts or mentions by the total number of posts or mentions providing specific information, respectively. Missing data were neither imputed nor considered in the analysis. Consequently, denominators varied across analysis and were adjusted based on the availability of data for specific variables.

Ethical Considerations

All data used in this SML study were collected exclusively from publicly accessible sources, without accessing any password-protected or nonpublic information. This study involved social media analysis and did not report any data with identifiable participants. All web-based content was anonymized and is in compliance with the Health Insurance Portability and Accountability Act (HIPAA). Approval was obtained from the Novartis safety registry 1P1R—the governing body that holds oversight on the use of social media by Novartis (DE014963-V1). All relevant local and global laws affecting and relating to the use of social media, aligned with and as reflected in Novartis processes, were followed in the conduct of this study. The authors of this study consent to the publication of the submitted manuscript and declare that no individual patient data requiring consent have been presented.

Results

Overview

A total of 20,008 posts were initially extracted from social media platforms using BC-related keywords during the study period. After filtration using exclusion keywords, removal of junk data and further data cleaning with the removal of duplicates using NLP, it resulted in retaining 1580 posts. These 1580 posts were then screened manually to check for relevance and identify stakeholders based on first-person conversation and user profiles, such as patients, caregivers, and HCPs. After excluding nonrelevant data, such as data from news portals, media, and publications, a total of 530 posts were included in the final analysis for insights generation (Figure 1).

Most posts were from patients (318/530, 60%), compared with 21% (111/530) from caregivers; 16% (85/530) from HCPs; and 3% (16/530) from unidentified persons. Of the 448 posts mentioning the cancer type, 95% (427/448) mentioned de novo BC and 5% (21/448) mentioned recurrent BC. Of the 107 posts mentioning the caregiver relationship with the patient, 53% (57/107) were posts from children discussing their mothers' diagnosis and treatment challenges. Patients' current age was reported in 89 posts, with most (56/89, 63%) posts reporting the patient age range of 21 - 50 years. Time since diagnosis was reported in 90 posts, with 31% (28/90) of the posts mentioning patients being diagnosed within 6 months to 1 year. Demographic and clinical characteristic details are reported in Multimedia Appendix 2.

Social Media Landscape and Online Behavior

Social Media Platform Distribution and Interaction Trends

Forums and X were the most commonly used social media platforms, accounting for 47% (249/530) and 31% (164/530) of the posts respectively, followed by Facebook (69/530, 13%), YouTube (21/530, 4%), Instagram (16/530, 3%), and blogs (11/530, 2%). Posts from 10 forums and 5 blogs were included in the analysis, and a list of forums and blogs is available in Multimedia Appendix 1. Among the forums, Medicitalia+, Aimac, Triste Mondo, and Carenity were the forums with the most posts. Most posts concentrated on the treatment and management of the disease, followed by discussions on diagnosis (Table 1). Among the 518 posts categorized as either sharing or seeking information, approximately 95% (493/518) involved sharing information, while only 27% (141/518) sought information.

Table . Discussion themes from the posts.

Topic of conversation	Mentions of the specific topic of conversation from the total mentions (n=1037), n (%)
Treatment and management	425 (41)
Diagnosis	249 (24)
Quality of life	135 (13)
Symptoms	124 (12)
Disease awareness	73 (7)
Clinical trial or research update	31 (3)

Sharing Information

Patients shared various aspects of their experiences, providing insights into their diagnostic journey, emotional support needs, and treatment courses. In discussing their diagnostic journey, patients shared their experiences of undergoing multiple diagnostic tests and the associated waiting periods, aiming to

XSL•FO
RenderX

reassure their peers who were facing similar challenges. Time since diagnosis was reported in 90 posts. Regarding emotional support, 17% (15/90) of the posts mentioning a recent diagnosis frequently expressed anxiety and concern about the progression of their treatments, whereas 12% (11/90) of the posts mentioning a diagnosis 4 to 6 years earlier were particularly active in addressing queries related to available treatment options and medication choices. Concerns about HCP interactions were also prevalent, with some patients expressing dissatisfaction due to multiple visits and the perceived lack of cordiality from certain HCPs. Furthermore, patients detailed the impact of their condition on their overall QoL and described pain, discomfort, and the effects on personal relationships and overall well-being.

Seeking Information

In the context of information-seeking behavior, patients expressed various concerns, including the interpretation of ambiguous symptoms, wherein they sought clarification from peers regarding concerning lumps or masses, as online resources often exacerbated confusion and anxiety. Additionally, there were diagnostic dilemmas, with individuals inquiring about preparation for upcoming breast examinations and biopsy results. Regarding treatment options, patients asked about improved disease management strategies to overcome the adverse effects of medications. Furthermore, patients expressed apprehension about medical consultations, especially regarding the frequency of medical consultations in the absence of clear guidance, and frequently sought second opinions from HCPs. This was largely driven by anxiety and worry induced by online testimonials read by patients.

A summary of insights from individual social media channels is provided in Multimedia Appendix 3.

Insights From Patients' and Caregivers' Discussions

Online Engagement and Support Seeking Pattern

Insights from patients' and caregivers' discussions highlighted that women aged 31 - 40 years, as reported in 26% (23/89) of the posts, were more active in sharing their patient journey and seeking peer experiences and community inclusiveness. Online platforms served as their preferred method for seeking clarification and addressing concerns, with approximately 30% (28/90) of the posts mentioning active online engagement within the first year of their diagnosis.

Symptoms

Key symptoms reported by patients included lumps, cysts, and breast pain (Table 2). Of the 128 mentions on symptoms, approximately 33% (42/128) of the mentions revealed that women identified a lump or mass as the first symptom through self-examination after a diagnostic checkup. Some mentions revealed that patients discovered their BC only through routine mammograms, as they were asymptomatic. Individuals educated themselves on self-examination at home by reading patient testimonials or watching videos or blogs from peers or medical websites. However, the uncertainty surrounding initial symptoms left patients unsure about whether to wait or visit a doctor, as the information available online was inconclusive. The physical impact of the symptoms, such as changes in breast appearance and severity of nipple discharge, made patients feel self-conscious and uncomfortable.

Table . Key symptoms reported in posts.

Key symptom	Mentions (%) of the specific symptom from the total mentions (n=128), n (%)
Lumps or mass	42 (33)
Cyst or calcification	20 (16)
Breast pain	17 (13)
Skin irritation	15 (12)
Nodules	14 (11)
Swelling of breasts	12 (9)
Nipple discharge	5 (4)
Others	5 (4)

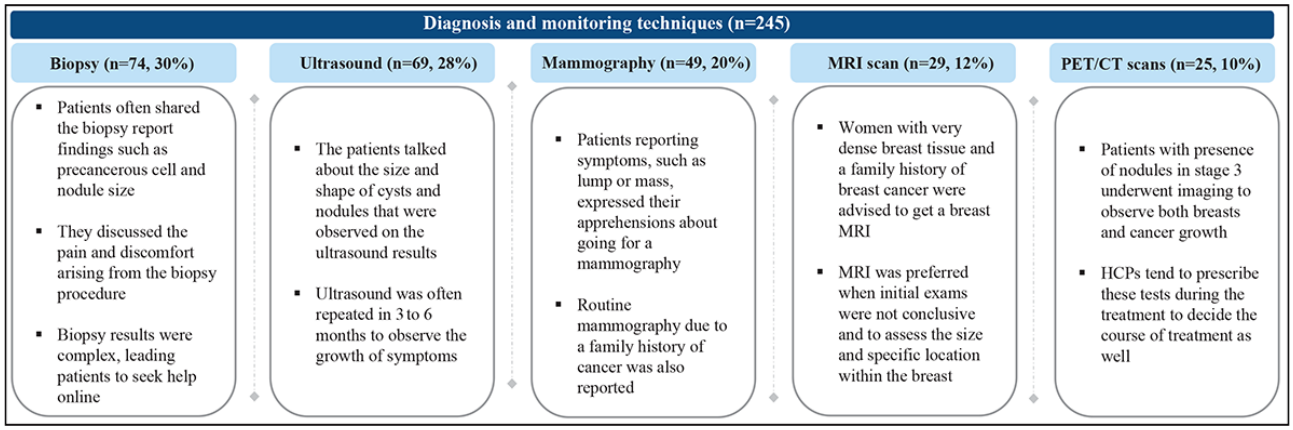
Diagnosis

Diagnosis and Monitoring: Procedures, Preferences, and Concerns

Diagnostic procedures, such as biopsy, ultrasonography, and mammography, played a vital role in confirming the diagnosis, accounting for 77% (190/245) of the mentions. These procedures were often repeated to establish a definitive diagnosis by radiologists and oncologists (Figure 2).

XSL•FO
RenderX

Figure 2. Diagnosis and monitoring techniques reported in posts. "n" denotes the number of mentions. CT: computed tomography; HCP: health care professionals; MRI: magnetic resonance imaging; PET: positron emission tomography.



The analysis of diagnostic and monitoring techniques from 245 mentions highlighted distinct preferences and concerns among patients. Biopsy, representing 30% (74/245) of the mentions, was frequently discussed, with patients sharing results that revealed conditions, such as precancerous cells and nodule sizes. Of the 73 mentions in which patients underwent biopsy, 22% (16/73) reported experiencing pain and discomfort after the procedure and often sought additional online resources due to the complexity of the biopsy results. Ultrasound, accounting for 28% (69/245) of the mentions, frequently involved discussions about the size and shape of cysts and nodules observed in the results, with follow-up ultrasounds typically scheduled every 3 to 6 months to monitor symptom progression. Mammography, mentioned in 20% (49/245) of the mentions, commonly induced patient apprehensions, especially in those reporting symptoms, such as lumps or masses. It was also routinely used by individuals with a family history of cancer. Magnetic resonance imaging (MRI) scans, comprising 12% (29/245) of the mentions, were particularly useful for patients with dense breast tissue and a significant family history of BC. MRIs were preferred when initial exams lacked conclusive results, offering a comprehensive assessment of the size and specific locations of the tumors within the breast. Positron emission tomography (PET) or computed tomography (CT) scans, constituting 10% (25/245) of mentions, were predominantly used to evaluate nodule presence in advanced cancer stages and track cancer growth. These scans were crucial during treatment phases to assist HCPs in determining appropriate treatment courses. Overall, these findings underscored the diverse diagnostic approaches and patient experiences in breast health monitoring.

Key Drivers for Diagnosis

Drivers for diagnosis included routine mammography and ultrasound screening, particularly among younger adults who were more informed. A family history of cancer made patients more vigilant in noticing symptoms and undergoing regular screenings. The self-examination for BC symptoms, proactive communication from gynecologists or breast specialists, and clear guidance from oncologists were also key factors.

Treatment

Surgery to Adjuvant Therapies: Experiences and Expectations

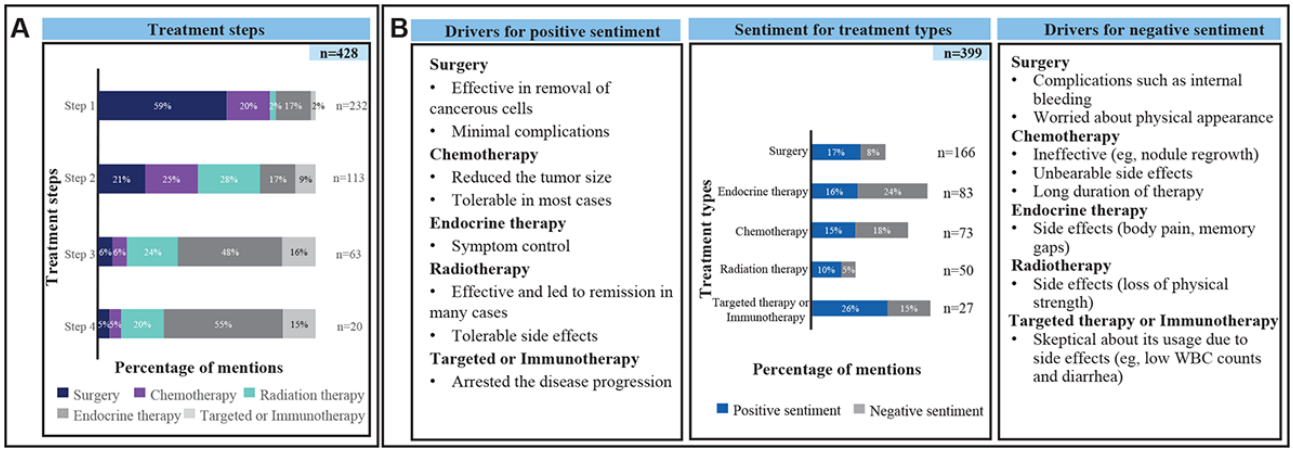
Before initiating treatment, patients often sought validation online for their prescribed treatments. Mastectomy was the most mentioned surgical method (137/232 mentions, 59%) for managing local cancerous cells at the first stage. After surgery, oncologists advised patients to undergo histological analysis to determine the further treatment course. Specific drug names were rarely discussed, but patients sought medications that had fewer side effects and higher efficacy. Some common chemotherapy and hormonal drugs mentioned for the first to third steps of treatment included paclitaxel, letrozole, and tamoxifen among others. Patients expressed anxiety about starting adjuvant therapies due to concerns about severe side effects. Adjuvant therapy was used intermittently to provide better relief to patients at high risk of disease progression. From the patients' and caregivers' perspective, HCPs rarely had transparent conversations with patients about the risks associated with hormonal or adjuvant therapies.

Treatment Sentiment

Patients experienced relief from surgery due to effective cancer cell removal and minimal complications. However, long-term treatments were challenging, negatively affecting their QoL. Chemotherapy reduced tumor size but had severe side effects and long durations. Endocrine therapy helped with symptom control but caused issues, such as body pain and memory gaps. Radiation therapy was effective, leading to remission but also resulted in the loss of physical strength. Targeted therapy or immunotherapy arrested disease progression but was met with skepticism due to side effects, such as low white blood cell counts and diarrhea (Figure 3).

XSL•FO
RenderX

Figure 3. (A) Treatment steps and (B) treatment sentiment. "n" denotes the number of mentions. WBC: white blood cell.



Unmet Treatment Needs

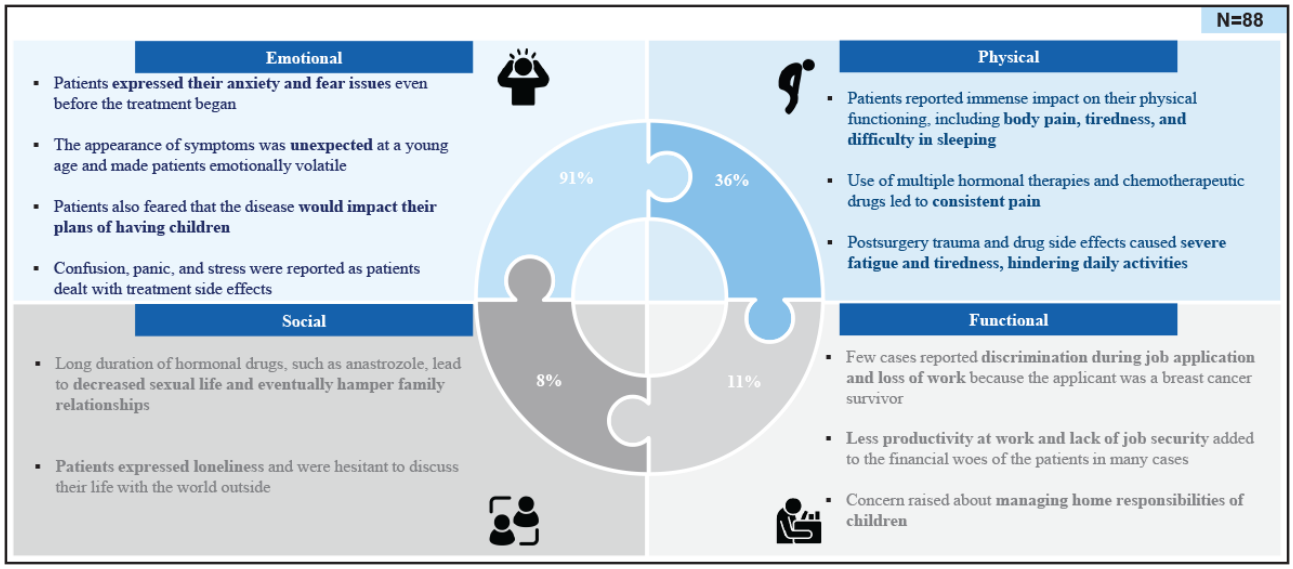
The treatment journey for patients with eBC is often further complicated by a lack of therapy knowledge and poor patient-HCP interactions. One of the key unmet needs identified is the perception of care.

- Limited treatment knowledge: Patients frequently lack awareness of available adjuvant treatment options that are critical for effective eBC management.
- Lack of trusted online peer groups: During treatment initiation, patients often seek support from trusted online peer groups, authenticated Google searches, or medical blogs. However, the absence of a standardized care system leaves them without proper guidance during this challenging period.
- Limited access to specialty breast care units: There is a pressing need for more tertiary cancer care centers across various regions to enhance patient access to specialized care.
- Poor patient-HCP communication: Treatment steps are often not openly discussed by HCPs during consultations, which places the burden of understanding and managing the treatment course on the patients themselves.
- Need for advanced drugs: Patients with recurrence highlighted the need for more advanced drugs to prevent recurrence.

Quality of Life

A total of 88 posts included discussions regarding QoL. eBC substantially impacted patients' QoL, affecting emotional well-being (80/88, 91%) on domains, such as anxiety and fear; physical functioning (32/88, 36%) in body pain, tiredness, and difficulty in sleeping; functional well-being (10/88, 11%), characterized by less productivity at work and lack of job security; and social well-being (7/88, 8%), including decreased sexual life, eventually hampering family relationships and resulting in loneliness (Figure 4).

Figure 4. Impact of disease on quality of life of patients with early breast cancer. "N" denotes the number of posts.



Recurrence and Associated Challenges

A total of 21 posts reported information on recurrence. Patients and caregivers expressed anger and disbelief upon recognizing symptoms of recurrence, which were often detected only after the cancer had metastasized to other organs. Symptoms, such as nodules and swollen breasts, were reported in 24% (5/21) of the posts on relapse. This led to repeated diagnostic tests,

XSL•FO
RenderX

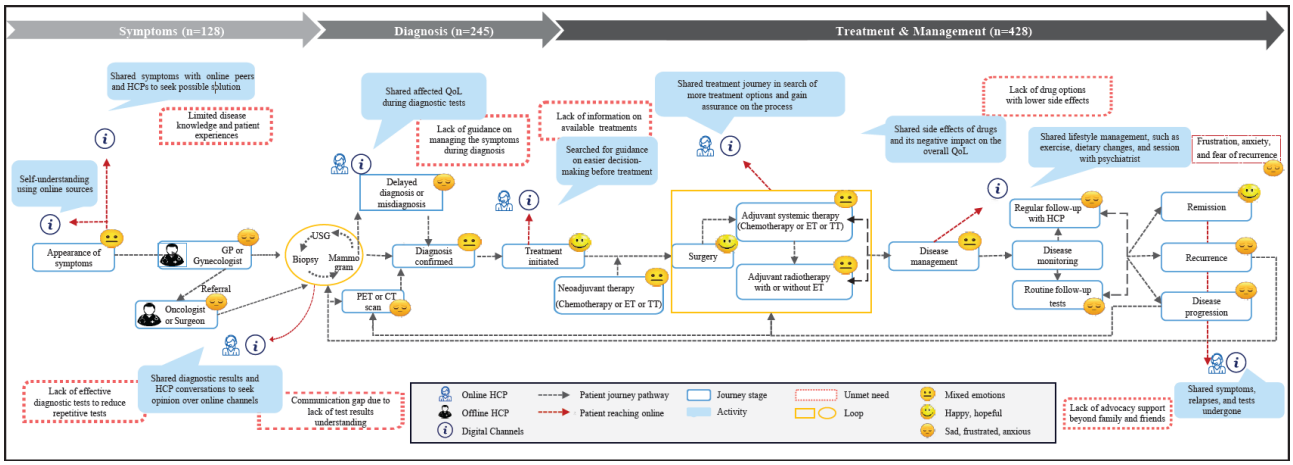
including CT, MRI, and PET scans as described in 33% (7/21) of the posts. The increase in disease activity on imaging tests was associated with feelings of devastation, as described in 33% (7/21) of the posts.

Recurrence had a profoundly negative impact on patients' QoL, causing frustration with the repetitive diagnostic tests and the treatment process. Overall, 19% (4/21) of the posts highlighted delays in disease monitoring, often due to false positive or negative biomarker results. Instances of negligence by radiologists were mentioned, with 10% (2/21) of the posts reporting that symptoms were initially ignored, leading to a later diagnosis of recurrence after seeking second opinions. Additionally, the long waiting periods for treatment in some areas of Italy prompted some patients to consider relocating for faster access to care, as reported in 10% (2/21) of the posts.

Considering 12 posts reporting data on remission, 59% (7/12) of the posts reported that patients were in remission for more than 6 years, 33% (4/12) for 3 - 6 years, and 8% (1/12) for less than 3 years.

The patient journey highlighted the reliance on digital channels for information, support, and communication through different stages, while it highlighted both online and offline interactions. Throughout their journey, patients experienced various emotions, including anxiety, sadness, frustration, hope, and happiness. Positive emotions were typically associated with successful treatments and management, while negative emotions often resulted from misdiagnoses, treatment side effects, and disease recurrence. Figure 5 illustrates the patient journey map, focusing on digital touchpoints—online platforms (Facebook, X, Instagram, YouTube, forums, and blogs) where patients seek information, support, and communicate with HCPs and peer networks—at various stages: symptoms, diagnosis, and treatment and management.

Figure 5. Digital touchpoints across the patient journey map. "n" denotes the number of mentions of the specific attribute. CT: computed tomography; ET: endocrine therapy; GP: general practitioner; HCP: health care professional; MRI: magnetic resonance imaging; PET: positron emission tomography; QoL: quality of life; TT: targeted therapy; USG: ultrasonography.



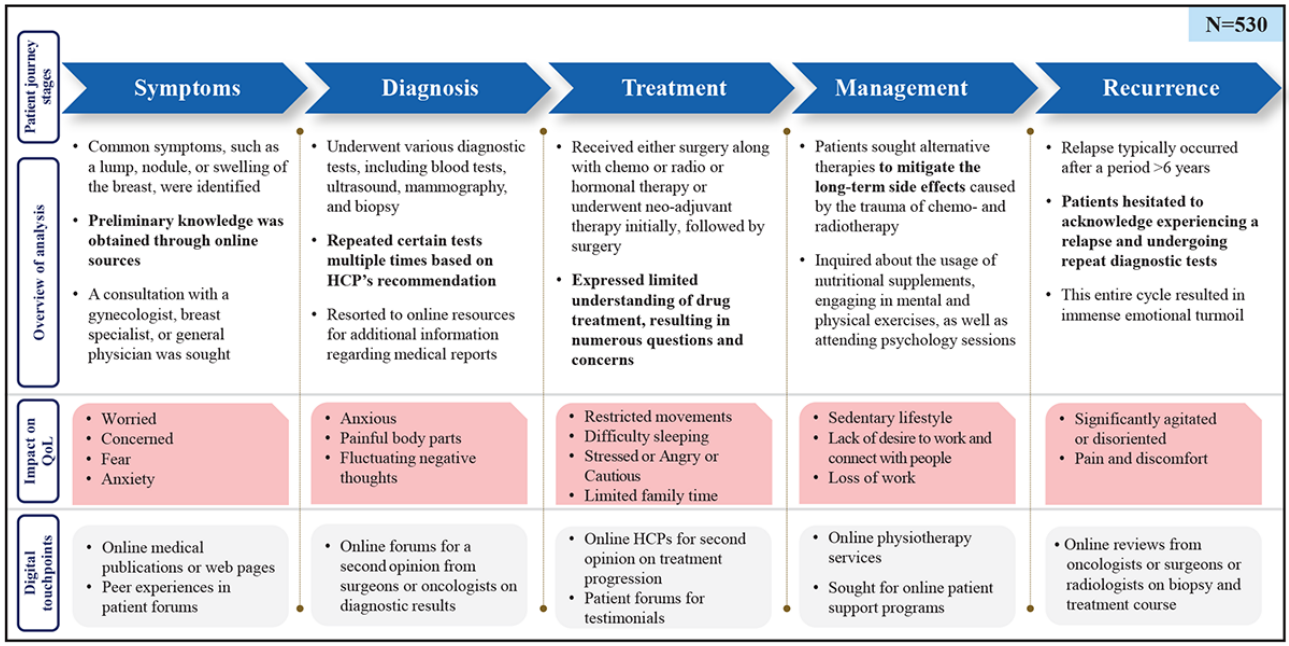
Insights From HCPs

This study highlighted optimism among HCPs regarding advancements in eBC care. They shared research through book publications focusing on diagnostic techniques and provided guidance on radiation therapy. Clinical advancements were discussed, including the role of artificial intelligence in BC management, phase 3 clinical trial results, and insights on new drugs for BC. Furthermore, a lot of emphasis was placed on raising disease awareness among patients. This included discussions on proactive measures for fertility preservation, sharing advancements in screening techniques, and underlining the importance of regular check-ups. HCPs also emphasized the harmful impacts of alcohol consumption on BC risk. They actively engaged patients through various channels to promote better disease management and prevention.

A summary of insights from this study at various stages of the journey of patients with eBC is shown in Figure 6.

XSL•FO
RenderX

Figure 6. Summary of insights at various stages of the journey of patients with early breast cancer (eBC). "N" denotes the number of posts. HCP: health care professional; QoL: quality of life.



Discussion

Overview of Findings

Key Themes Identified Across the Journey of Patients With eBC in Italy

This SML study provides insights into how patients with eBC, caregivers, and HCPs in Italy describe their experiences, perceptions, and unmet needs on social media. Through the analysis of posts from various social media platforms over a 2-year period, we identified recurring themes related to initial symptoms and diagnosis, treatment experiences and concerns (including adjuvant therapies), impact on QoL, perceived recurrence and ongoing monitoring, communication gaps with HCPs, and access to support and specialized care. These findings underscore the significance of social media as a supplementary resource for understanding patient experiences in real-world settings and highlight critical areas that necessitate attention to improve patient care and support systems.

Social Media Landscape and Online Behavior

This study reveals that most social media posts were initiated by patients themselves, highlighting their willingness to share experiences and seek peer support. This aligns with previous research that indicated that patients are increasingly turning to online platforms for emotional and informational support [16,30]. The prevalent use of forums and X emphasizes the role of these platforms in disseminating information and fostering community engagement [14,17]. However, as the sample is limited to users who are active online, these insights may reflect the perspectives of more digitally engaged patients rather than all individuals with eBC.

Initial Symptoms and Diagnosis

The most common symptoms reported in this study include lumps, cysts, and breast pain. These early indicators drive patients to seek diagnostic interventions, such as mammography, biopsy, and ultrasound, with patients having varied anxiety levels toward these methods.

Patients frequently discussed the diagnostic journey, expressing concerns about the multiple tests and the associated waiting periods. These findings were similar to those of a previous study that found that prolonged diagnostic processes contribute to substantial patient anxiety and stress [31]. Patients' inability to describe their disease stage during discussions indicated a generally low level of awareness about the disease and its clinical characteristics (eg, staging and interpretation of test reports). Consistent with previous research [32,33], the findings of this study indicated that self-examination for BC symptoms, proactive communication from gynecologists or breast specialists, and clear guidance from oncologists were key drivers for diagnosis, supporting the early identification of BC and the potential for improved survival outcomes. Given the nature of SML data, these observations are hypothesis-generating and should be interpreted as qualitative insights rather than definitive epidemiologic estimates.

Treatment Experiences and Sentiments

Patients also expressed great concern regarding the long durations and side effects of chemotherapy and endocrine therapy, leading them to seek alternative treatment strategies. This aspect of patient experience is well documented in the literature, where adverse effects of treatment often result in diminished QoL [34]. The challenges in understanding the complexities of adjuvant therapies indicate a need for more transparent communication from HCPs [4]. Furthermore, negative perceptions of adjuvant therapies significantly impact treatment adherence. Prolonged low adherence or nonadherence increases the risk of recurrence, ultimately leading to poorer patient outcomes [35-37]. In our dataset, patients often described doubts about the benefits of ongoing treatment and concerns about side effects. Although actual adherence behavior could

XSL•FO
RenderX

not be measured, these narratives signal areas where enhanced patient-HCP dialogue and shared decision-making may be beneficial.

The discussions reflected a lack of open communication from HCPs regarding treatment options. Similar findings were observed in a social media study conducted on patients with metastatic BC from Europe [19]. Collectively, these findings indicate that enhancing the clarity, frequency, and empathy of patient-HCP communication may help address some of the concerns expressed in online discussions.

Impact on QoL

The impact of eBC on patients' QoL was another critical theme identified in this study. Consistent with the literature, the findings of this study show that eBC substantially affects emotional well-being, physical functioning, functional well-being, and social well-being [38]. Patients reported anxiety, fear, pain, and fatigue, which also influenced their productivity at work and personal relationships. The prevalence of emotional distress among patients with eBC underscores the importance of comprehensive supportive care that addresses both the physical and psychological needs of patients [34,39-42]. Because the data were derived from spontaneous posts, individuals experiencing higher levels of distress may have been more likely to share their experiences online, potentially amplifying the prominence of negative QoL themes.

Recurrence, Ongoing Screening, and Continuous Therapy

The findings of this study indicated that recurrence had a substantial impact on QoL, necessitating regular imaging tests and repeated cycles of chemotherapy, which became an integral part of the treatment journey. This study highlights a critical unmet need for timely and accurate monitoring to prevent recurrence and mitigate long-term impacts. Although we could not independently verify clinical details, such as recurrence status or specific testing protocols, repeated patient discussions regarding continuous imaging, biomarker testing, and changes in therapy underscore the burden of long-term surveillance from the patient perspective.

Unmet Needs and Communication Gaps

Despite the availability of various treatment options, a notable number of patients indicated unmet needs in treatment knowledge and patient-HCP communication. The lack of trusted online peer groups and limited access to specialty breast care units were also prominent concerns. These findings are supported by previous studies [43-45] that emphasized the importance of accessible, timely, and accurate information for patients to make informed decisions about their care, as well as the role of patient advocacy groups [46].

Poor patient-HCP communication was frequently mentioned, with patients expressing dissatisfaction with the clarity and frequency of consultations. Patients complained about multiple tests and ineffective communication with HCPs about test results, prompting patients to seek help online. Effective communication is critical in oncology care, yet often remains suboptimal, contributing to patient distress and a sense of being unsupported [11,19]. This suggests that targeted interventions to improve patient-HCP interactions could enhance patient satisfaction and perceived outcomes.

Role of HCPs and Social Media

HCPs also engaged actively on social media, sharing advancements in eBC care and offering guidance on disease management. This active participation aligns with the findings from previous studies that underscore the role of social media in bridging knowledge gaps and fostering public health education [47,48]. The discussions by HCPs on clinical trials, diagnostic techniques, and preventive measures, such as regular screenings and lifestyle modifications, reflect an ongoing effort to raise disease awareness and improve patient care. Nevertheless, our analysis suggests that HCP-driven content may not always directly address the specific questions and concerns raised by patients and caregivers, indicating an opportunity to better align online communication with patient or caregiver needs.

New Insights From the Findings of This Study

New insights from the findings of this study are as follows:

1. This SML study revealed gaps that are not typically captured by traditional research methods, including the generational shift in how younger patients (aged 31 - 40 years) are using online platforms as their primary source of both emotional and informational support. This highlights an urgent need for formal integration of digital tools into cancer care frameworks, for example, through moderated online communities, validated educational content, and digital navigation tools.
2. The study provides new evidence regarding the reported geographic inequities in access to care. Patients in underserved areas reported delays in treatment and follow-up, highlighting a systemic issue that could be addressed with regional policy innovations, such as expanding specialized breast care units or creating mobile health care services. These insights should be interpreted cautiously as they are based on self-reported locations and experiences, but they point to priority areas for further health services research.
3. Recurring diagnostic inaccuracies, particularly in biomarker testing during recurrence monitoring, emerged as a unique and actionable finding. While specific clinical details could not be confirmed, addressing these perceived inaccuracies through standardized protocols and advanced imaging techniques could significantly alleviate patient distress and potentially improve outcomes. Future studies combining SML with clinical data or registry information may help validate and quantify the extent of these issues.

Strengths and Limitations

This study has several strengths and limitations. One key strength is its relatively novel approach, utilizing SML to gather real-world insights into patient experiences and perceptions without intruding on their privacy. This method captures spontaneous and candid discussions among patients, caregivers, and HCPs that are difficult or time-consuming to elicit through traditional research methods, such as surveys or interviews. A

XSL•FO
RenderX

second strength is the breadth and diversity of data sources. Posts were collected from multiple social media platforms (X, Facebook, forums, blogs, Instagram, and YouTube), which provides diverse perspectives and a broader understanding of the journey of patients with eBC in Italy. Third, the study benefits from a systematic methodological framework. The search strategy was developed by a multidisciplinary team and iteratively refined with scoping searches and clearly defined eligibility criteria. Data processing combined NLP-based filtering and manual review to exclude junk, spam, and irrelevant content and focus on posts reflecting experiences of patients, caregivers, and HCPs. Thematic analysis was guided by a patient-journey–based conceptual framework. Together, these elements enhance the methodological rigor, transparency, and reproducibility of the study. Finally, the study generated a substantial volume of qualitative material for a focused national context. From 20,008 initially extracted posts, 530 relevant posts were retained for detailed analysis, enabling an in-depth exploration of specific themes, such as diagnostic pathways, perceptions of adjuvant therapies, QoL impacts, recurrence-related challenges, communication gaps, and access barriers.

However, the study has certain limitations. First, the social media users included in the analysis may not be representative of the broader population of patients with eBC, resulting in a potentially nonrepresentative sample. We attempted to mitigate this by including multiple platforms and both Italian- and English-language posts from users in Italy. However, selection bias toward more digitally engaged, younger, or health-literate individuals is likely. This may overemphasize certain themes (eg, online information-seeking) and underrepresent less connected or older populations. Future studies could integrate SML with survey or registry data to improve generalizability. Second, as the study relies on self-reported information, the accuracy and completeness of the data may vary, and it is not possible to verify the authenticity of individual posts. Third, the retrospective, cross-sectional design captures content from a specific period, from December 2021 to November 2023, which may not fully reflect changes over time or emerging trends. As such, temporal shifts in treatment paradigms, public awareness, or health policy may not be fully captured. Longitudinal SML analyses or repeated cross-sectional assessments could provide a more dynamic picture of evolving patient needs. In addition, the anonymous nature of social media data from publicly available posts prevents tracking individual users across platforms and precludes the verification of key demographics or clinical characteristics (eg, age, disease stage). This limits the ability to draw robust conclusions about specific subgroups. Our subgroup insights (eg, younger patients or those in underserved regions) should therefore be interpreted as indicative patterns rather than statistically confirmed differences. Future work that links deidentified social media data with structured demographic or clinical information, where ethically and legally permissible, could refine these subgroup analyses. Finally, the quality, depth, and specificity of shared information differ considerably between the posts, which may introduce gaps and uncertainty in data analysis and interpretation.

These strengths and limitations provide a comprehensive understanding of the study and emphasize the need for additional research and ongoing methodological improvements in the use of SML to inform patient-centered cancer care.

Conclusions

This study highlights the utility of SML in understanding the nuanced experiences of patients with eBC, caregivers, and HCPs in Italy, as expressed through publicly available social media posts. The identified themes reveal critical unmet needs in patient education, support systems, access to specialty care units, and patient-HCP communication. While the findings are exploratory and reflect the perspectives of social media users rather than all patients with eBC, they provide valuable, real-world insights that can complement traditional research approaches. Addressing these gaps through targeted interventions and leveraging social media's potential for tailored information, peer support, and enhanced patient-HCP communication can enhance patient experiences and perceived outcomes. Further research is encouraged to develop and implement effective strategies for improving cancer care in the digital age, including the integration of digital health tools, evaluation of patient-HCP communication interventions, and linkage of SML data with clinical and health services datasets where feasible and ethical.

Acknowledgments

The authors acknowledge Pradeep Basa (Novartis Healthcare Pvt. Ltd., India) for editorial and writing support, in accordance with the Good Publication Practice (GPP3) guidelines of the International Society for Medical Publication Professionals.

Funding

The study and article processing charges were funded by Novartis Farma S.p.A. (Milan, Italy).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

MBS, RD, MM, EF, A Messi, and DV conceptualized the study design. A Maurya and SP contributed to the conduct of study and data collection. A Maurya, SP, MBS, and DV analyzed and interpreted the data, along with other authors. All authors reviewed the manuscript critically and provided final approval for submission. All named authors meet the International Committee of

XSL•FO
RenderX

Medical Journal Editors (ICMJE) criteria for authorship for this paper and take responsibility for the integrity of the work as a whole.

Conflicts of Interest

MBS, EF, A Messi, A Maurya, SP, and DV are full-time employees at the Novartis group of companies and declare that they have no competing interests. MM has provided consultancy/advice or received reimbursement for travel expenses from Roche, AstraZeneca, Daiichi Sankyo, Pfizer, Novartis, Lilly, ORGANON, and Gentili and declares no other competing interests. RD declares no competing interests related to this work.

Multimedia Appendix 1
Search terms, inclusion and exclusion criteria, and list of forums and blogs.
[XLSX File, 25 KB - cancer_v12i1e73371_app1.xlsx]

Multimedia Appendix 2
Demographic and clinical characteristics.
[DOCX File, 17 KB - cancer_v12i1e73371_app2.docx]

Multimedia Appendix 3
Insights from various social media channels.
[PDF File, 693 KB - cancer_v12i1e73371_app3.pdf]

Checklist 1
SRQR checklist.
[DOCX File, 24 KB - cancer_v12i1e73371_app4.docx]

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
2. I numeri del cancro in Italia 2023 [Report in Italian]. : Associazione Italiana di Oncologia Medica (AIOM); 2023 URL: https://www.aiom.it/wp-content/uploads/2023/12/2023_AIOM_NDC-web.pdf [accessed 2026-02-24]
3. Gori S, Zambelli A, Angiolini C, et al. Follow-up of breast cancer: why is it necessary to start a Consensus in 2024? AboutOpen 2024 Feb;11(1):10-17. [doi: 10.33393/ao.2024.2634]
4. Perrone F, De Laurentiis M, De Placido S, et al. Adjuvant zoledronic acid and letrozole plus ovarian function suppression in premenopausal breast cancer: HOBOE phase 3 randomised trial. Eur J Cancer 2019 Sep;118:178-186. [doi: 10.1016/j.ejca.2019.05.004] [Medline: 31164265]
5. Davis LE, Bubis LD, Mahar AL, et al. Patient-reported symptoms after breast cancer diagnosis and treatment: a retrospective cohort study. Eur J Cancer 2018 Sep;101:1-11. [doi: 10.1016/j.ejca.2018.06.006] [Medline: 30014970]
6. Dinapoli L, Colloca G, Di Capua B, Valentini V. Psychological aspects to consider in breast cancer diagnosis and treatment. Curr Oncol Rep 2021 Mar 11;23(3):38. [doi: 10.1007/s11912-021-01049-3] [Medline: 33709235]
7. Grusdat NP, Stäuber A, Tolkmitt M, et al. Cancer treatment regimens and their impact on the patient-reported outcome measures health-related quality of life and perceived cognitive function. J Patient Rep Outcomes 2022 Feb 21;6(1):16. [doi: 10.1186/s41687-022-00422-5] [Medline: 35190938]
8. Blumen H, Fitch K, Polkus V. Comparison of treatment costs for breast cancer, by tumor stage and type of service. Am Health Drug Benefits 2016 Feb;9(1):23-32. [Medline: 27066193]
9. Subramanian S, Trogdon J, Ekwueme DU, Gardner JG, Whitmire JT, Rao C. Cost of breast cancer treatment in Medicaid: implications for state programs providing coverage for low-income women. Med Care 2011 Jan;49(1):89-95. [doi: 10.1097/MLR.0b013e3181f81c32] [Medline: 21079524]
10. Friedel AL, Siegel S, Kirstein CF, et al. Measuring patient experience and patient satisfaction—how are we doing it and why does it matter? A comparison of European and US American approaches. Healthcare (Basel) 2023 Mar 8;11(6):797. [doi: 10.3390/healthcare11060797] [Medline: 36981454]
11. Lavela SL, Gallan AS. Evaluation and measurement of patient experience. Patient Exp J 2014;1(1):28-36 [FREE Full text]
12. Moser A, Korstjens I. Series: practical guidance to qualitative research. Part 3: sampling, data collection and analysis. Eur J Gen Pract 2018 Dec;24(1):9-18. [doi: 10.1080/13814788.2017.1375091] [Medline: 29199486]
13. Salmani H, Nasiri S, Ahmadi M. The advantages, disadvantages, threats, and opportunities of electronic patient-reported outcome systems in cancer: a systematic review. Digit Health 2024;10:20552076241257146. [doi: 10.1177/20552076241257146] [Medline: 38812853]

XSL•FO
RenderX

14. Kusudo M, Terada M, Kureyama N, et al. Characterizing user demographics in posts related to breast, lung and colon cancer on Japanese Twitter (X). Sci Rep 2024 Mar 18;14(1):6485. [doi: 10.1038/s41598-024-56679-x] [Medline: 38499598]
15. Sreenivasan S, Fang C, Flood EM, Markuzon N, Sze JYY. Insights into the patient experience of hormone therapy for early breast cancer treatment using patient forum discussions and natural language processing. JCO Clin Cancer Inform 2024 Aug;8:e2400038. [doi: 10.1200/CCI.24.00038] [Medline: 39102642]
16. Ure C, Cooper-Ryan AM, Condie J, Galpin A. Exploring strategies for using social media to self-manage health care when living with and beyond breast cancer: in-depth qualitative study. J Med Internet Res 2020 May 25;22(5):e16902. [doi: 10.2196/16902] [Medline: 32364510]
17. Cimiano P, Collins B, De Vuono MC, et al. Patient listening on social media for patient-focused drug development: a synthesis of considerations from patients, industry and regulators. Front Med (Lausanne) 2024;11:1274688. [doi: 10.3389/fmed.2024.1274688] [Medline: 38515987]
18. For industry: using social media. US Food and Drug Administration (FDA). 2024. URL: https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/industry-using-social-media [accessed 2024-08-21]
19. Mazza M, Piperis M, Aasaithambi S, Chauhan J, Sagkriotis A, Vieira C. Social media listening to understand the lived experience of individuals in Europe with metastatic breast cancer: a systematic search and content analysis study. Front Oncol 2022;12:863641. [doi: 10.3389/fonc.2022.863641] [Medline: 35719996]
20. Efstathiou E, Merseburger A, Liew A, et al. Perception of cure in prostate cancer: human-led and artificial intelligence-assisted landscape review and linguistic analysis of literature, social media and policy documents. ESMO Open 2024 May;9(5):103007. [doi: 10.1016/j.esmoop.2024.103007] [Medline: 38744101]
21. Kim D, Lim JH, Kim IH, et al. 1858P Discrepancies between psychological stress and emotional care in cancer patients revealed in a social listening analysis. Ann Oncol 2023 Oct;34:S1003. [doi: 10.1016/j.annonc.2023.09.2807]
22. Chen L, Wang X, Peng TQ. Nature and diffusion of gynecologic cancer–related misinformation on social media: analysis of tweets. J Med Internet Res 2018 Oct 16;20(10):e11515. [doi: 10.2196/11515] [Medline: 30327289]
23. Schück S, Loussikian P, Mebarki A, et al. PCR24 Patient and caregiver perceptions of systemic anticancer treatments in advanced bladder cancer (aBC): results of a social media listening (SML) study conducted in 5 European countries (Eu5). Value Health 2023 Dec;26(12):S453. [doi: 10.1016/j.jval.2023.09.2463]
24. Yekrou D, Guettouche MG, Magdy M. Awareness about genitourinary cancer in North Africa—social listening analytics. Eur Urol Open Sci 2023 Nov;57:S370-S371. [doi: 10.1016/S2666-1683(23)01564-1]
25. Montrone M, Chauhan J, Sagkriotis A, Aasaithambi S, Rodrigues ACF. 1343P Understanding patient and caregiver perceptions of quality of life (QoL) impacts in lung cancer through social listening. Ann Oncol 2021 Sep;32:S1024. [doi: 10.1016/j.annonc.2021.08.1944]
26. Wolffsohn JS, Leteneux-Pantais C, Chiva-Razavi S, et al. Social media listening to understand the lived experience of presbyopia: systematic search and content analysis study. J Med Internet Res 2020 Sep 21;22(9):e18306. [doi: 10.2196/18306] [Medline: 32955443]
27. Freeman TCB, Rodriguez-Esteban R, Gottowik J, Yang X, Erpenbeck VJ, Leddin M. A neural network approach for understanding patient experiences of chronic obstructive pulmonary disease (COPD): retrospective, cross-sectional study of social media content. JMIR Med Inform 2021 Nov 11;9(11):e26272. [doi: 10.2196/26272] [Medline: 34762056]
28. Rodrigues A, Chauhan J, Sagkriotis A, Aasaithambi S, Montrone M. Understanding the lived experience of lung cancer: a European social media listening study. BMC Cancer 2022 Apr 30;22(1):475. [doi: 10.1186/s12885-022-09505-4] [Medline: 35490223]
29. Sprinklr. URL: https://www.sprinklr.com/products/consumer-intelligence/social-listening [accessed 2023-12-01]
30. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res 2016 Aug 26;16(1):442. [doi: 10.1186/s12913-016-1691-0] [Medline: 27562728]
31. Fefferman ML, Stump TK, Thompson D, Simovic S, Medenwald RJ, Yao K. Patient-reported observations on medical procedure timeliness (PROMPT) in breast cancer: a qualitative study. Breast Cancer Res Treat 2024 Nov;208(1):123-132. [doi: 10.1007/s10549-024-07406-7] [Medline: 38965153]
32. Ginsburg O, Yip CH, Brooks A, et al. Breast cancer early detection: a phased approach to implementation. Cancer 2020 May 15;126(Suppl 10):2379-2393. [doi: 10.1002/cncr.32887] [Medline: 32348566]
33. Kuroishi T, Tominaga S, Ota J, et al. The effect of breast self-examination on early detection and survival. Jpn J Cancer Res 1992 Apr;83(4):344-350. [doi: 10.1111/j.1349-7006.1992.tb00113.x] [Medline: 1506268]
34. Prieto-Callejero B, Rivera F, Fagundo-Rivera J, et al. Relationship between chemotherapy-induced adverse reactions and health-related quality of life in patients with breast cancer. Medicine (Baltimore) 2020 Aug 14;99(33):e21695. [doi: 10.1097/MD.0000000000021695] [Medline: 32872042]
35. Todd A, Waldron C, McGeagh L, et al. Identifying determinants of adherence to adjuvant endocrine therapy following breast cancer: a systematic review of reviews. Cancer Med 2024 Feb;13(3):e6937. [doi: 10.1002/cam4.6937] [Medline: 38240343]

XSL•FO
RenderX

36. AlOmeir O, Patel N, Donyai P. Adherence to adjuvant endocrine therapy among breast cancer survivors: a systematic review and meta-synthesis of the qualitative literature using grounded theory. Support Care Cancer 2020 Nov;28(11):5075-5084. [doi: 10.1007/s00520-020-05585-9] [Medline: 32601852]
37. Makubate B, Donnan PT, Dewar JA, Thompson AM, McCowan C. Cohort study of adherence to adjuvant endocrine therapy, breast cancer recurrence and mortality. Br J Cancer 2013 Apr 16;108(7):1515-1524. [doi: 10.1038/bjc.2013.116] [Medline: 23519057]
38. Criscitiello C, Spurden D, Piercy J, et al. Health-related quality of life among patients with HR+/HER2–early breast cancer. Clin Ther 2021 Jul;43(7):1228-1244.e4. [doi: 10.1016/j.clinthera.2021.04.020] [Medline: 34256965]
39. Di Meglio A, Catanuto G, Zambon M, et al. 273P Experiences and preferences about information on treatment (tx)-related side effects (SE) among patients (pts) with early breast cancer (EBC). ESMO Open 2024 May;9:103331. [doi: 10.1016/j.esmoop.2024.103331]
40. Korani OM, Samy SA, Eid R, Helal AM, Shash E. 1675P Impact of adjuvant hormonal therapy on sexual and global quality of life in young Egyptian women with breast cancer. Ann Oncol 2024 Sep;35:S1006. [doi: 10.1016/j.annonc.2024.08.1759]
41. Pelzer F, Tröger W, Reif M, et al. Fatigue and quality of life during neoadjuvant chemotherapy of early breast cancer: a prospective multicenter cohort study. Breast Cancer (Auckl) 2024 Jan;31(1):124-134. [doi: 10.1007/s12282-023-01520-y] [Medline: 37966690]
42. Potter S, Mactier M, Fairhurst K, Stobart H, Gath J, McIntosh S. Understanding patient experiences of early breast cancer treatment: informing future studies to reduce treatment burden. Eur J Cancer 2024 Mar;200:113668. [doi: 10.1016/j.ejca.2024.113668]
43. Ormel I, Magalhaes M, Josephson D, Tracey L, Law S. How to know what to know: information challenges for women in the diagnostic phase of breast cancer. Patient Educ Couns 2021 Jan;104(1):179-185. [doi: 10.1016/j.pec.2020.06.030] [Medline: 32646774]
44. Rattanakanlaya K, Anusasananun B, Chaisurin P, et al. Information needs among Thai women breast cancer patients during the initial treatment phase: a cross-sectional study. J Clin Nurs 2025 Sep;34(9):3826-3836. [doi: 10.1111/jocn.17598] [Medline: 39632510]
45. Lu H, Xie J, Gerido LH, Cheng Y, Chen Y, Sun L. Information needs of breast cancer patients: theory-generating meta-synthesis. J Med Internet Res 2020 Jul 28;22(7):e17907. [doi: 10.2196/17907] [Medline: 32720899]
46. Dvaladze A, Kizub DA, Cabanes A, et al. Breast cancer patient advocacy: a qualitative study of the challenges and opportunities for civil society organizations in low-income and middle-income countries. Cancer 2020 May 15;126:2439-2447. [doi: 10.1002/cncr.32852] [Medline: 32348570]
47. Hazzam J, Lahrech A. Health care professionals' social media behavior and the underlying factors of social media adoption and use: quantitative study. J Med Internet Res 2018 Nov 7;20(11):e12035. [doi: 10.2196/12035] [Medline: 30404773]
48. Khan MN, Faraz A, Jamal AB, et al. A study to see the effect of social media usage among healthcare providers. Cureus 2021 Jul;13(7):e16350. [doi: 10.7759/cureus.16350] [Medline: 34395130]

Abbreviations

BC: breast cancer
CT: computed tomography
eBC: early breast cancer
HCP: health care professional
HIPAA: Health Insurance Portability and Accountability Act
MRI: magnetic resonance imaging
NLP: natural language processing
PET: positron emission tomography
QoL: quality of life
SML: social media listening
SRQR: Standards for Reporting Qualitative Research

Edited by M Balcarras, N Cahill; submitted 17.Mar.2025; peer-reviewed by CS Lam, J Braverman; accepted 08.Jan.2026; published 24.Mar.2026.

Please cite as:
Suter MB, D'Antona R, Mazza M, Francavilla E, Messi A, Maurya A, Padhi S, Valsecchi D
Exploring the Experiences and Perspectives of Patients With Early Breast Cancer, Caregivers, and Health Care Professionals: Italian Social Media Listening Study
JMIR Cancer 2026;12:e73371
URL: https://cancer.jmir.org/2026/1/e73371
doi:10.2196/73371

XSL•FO
RenderX

© Matteo Basilio Suter, Rosanna D'Antona, Manuelita Mazza, Enrica Francavilla, Alice Messi, Abhinav Maurya, Shiva Padhi, Diletta Valsecchi. Originally published in JMIR Cancer (https://cancer.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Acceptability of Sharing Internet Browsing History for Cancer Research: Think-Aloud and Interview Study

Nicola Cara Gradwell, BDS; Mel Ramasawmy, PhD; Sanjula Arora, MSc; Manoj Mistry, MBA; Reshma Punjabi; Christina Derksen, PhD; Suzanne E Scott, PhD

[1]Wolfson Institute of Population Health, Queen Mary University of London, Charterhouse Square, London, United Kingdom

Corresponding Author:
Mel Ramasawmy, PhD
Wolfson Institute of Population Health, Queen Mary University of London, Charterhouse Square, London, United Kingdom

Abstract

Background: Growing interest surrounds how internet search behaviors might provide digital signals of disease prior to diagnosis, for example, when people search symptoms online. Internet browsing data offer novel opportunities for understanding response to symptoms, public health surveillance, and early intervention in conditions such as cancer. However, the acceptability of using such sensitive data in medical research remains unclear, particularly among individuals at higher risk of health and digital exclusion, such as older adults and those from minority ethnic groups or with a lower socioeconomic status.

Objective: This study aims to explore the feasibility and acceptability of using internet browsing history data for health research.

Methods: Participants were purposively sampled to ensure representation from groups at risk of digital and health inequalities via community organizations and charities. We conducted semistructured and think-aloud interviews allowing participants to reflect on hypothetical research involving sharing their internet browsing data. The adapted theoretical framework of acceptability guided the interview structure and coding. The interviews were transcribed, coded in NVivo, and thematically analyzed. Patient and public involvement informed the study approach, participant-facing documents, and the interpretation of the findings.

Results: Twenty participants (10 with a history of cancer and 10 without) were included in the study representing a range of age, gender, and ethnic and socioeconomic groups. Key themes focused on factors necessary for acceptability, including trust, transparency, and control and on perceived feasibility and individual willingness. Trust and transparency were fundamental to participants' willingness to share data. Trust in researchers would have to be earned through clear communication, ethical data handling, and familiarity with a named research team. Privacy concerns were prominent, with participants wanting control over what was shared, particularly regarding nonhealth-related information (such as details related to banking) or activity related to others (such as their children). Potential use or misuse of data beyond the original research purpose caused more concern than the nature of the shared data itself. Digital literacy varied; many expressed concerns over the technical aspects of sharing data. Participants also doubted the value of their individual internet browsing history, for example, as they chose not to search for health information due to the prevalence of misinformation. However, they described wider benefits arising from internet browsing history research, such as potential advancements in early detection and opportunities to promote credible online sources.

Conclusions: Participant recommendations balanced privacy concerns against the potential of internet history data for early diagnosis and health research. The study highlights ethical and inclusive approaches to health research using internet browsing history. Future researchers should consider defining the scope of health-specific data filters, providing user-friendly information and guidance for study participants, and ensuring that participants are able to contact research team members to build trust and facilitate data sharing.

(JMIR Cancer 2026;12:e82009) doi:10.2196/82009

KEYWORDS

cancer; digital inequalities; early diagnosis; health inequalities; health-seeking behavior; internet search; internet search data; smart data; web search

Introduction

For many medical conditions, good prognosis and effective treatment rely on timely diagnosis [1,2]. There is also a large economic burden of delayed diagnosis when disease becomes advanced or requires more invasive treatment, longer hospital stays, and greater impact on quality of life [3,4]. In particular, early detection of cancer is a national priority in England, with the National Health Service (NHS) Long-Term Plan aiming to diagnose 75% of cancers at an early stage by 2028 [5-7]. A patient's pathway to diagnosis of cancer may begin weeks or months before they seek help from a health care provider, when

XSL•FO
RenderX

they first notice changes in their body and make decisions about the next steps, including doing nothing, monitoring, self-managing (eg, through medication or rest), seeking external information, or consulting a health care provider [8]. Misinterpreting symptoms as not needing health care appears to be the most important reason for diagnostic delays, constituting approximately 60% of the overall time from first symptoms to diagnosis in cancer [9].

There is growing interest in exploring how the internet is used in response to symptoms and the decision to seek help. Accessing health information online prior to appointments can empower patients and enable more informed conversations with health care providers [10]. Individuals typically search for specific symptoms rather than diagnoses before visiting a health care professional [11,12], although over 40% of individuals opt against seeing a doctor after researching their symptoms [10,13,14]. Recent research has suggested that internet search activity history may provide valuable early digital signals of disease at an individual level. For example, there has been pilot work into the predictive utility of internet searches in relation to gynecological malignancies [15]. The researchers found differences in online search data between malignant and benign gynecological conditions, up to a year before general practitioner referral. Similar findings have been reported in lung cancer [16] and pancreatic adenocarcinoma [17]. However, studies have found variable opinions on sharing internet browsing history data [18-20], and there is a lack of diversity in studied populations. This has the potential to compound existing inequalities [21]. It is known that older adults, those from minority ethnic groups, and those with the lowest levels of socioeconomic status have differences in linguistics, disease rates, digital literacy, and help-seeking behaviors [21-23]; are less likely to use the internet for health information [24-26]; face barriers to internet access [22,27]; and are more likely to experience inequalities in cancer care and outcomes [28]. Disadvantaged groups are also less likely to be willing to share health information more generally [29,30].

Crucial to successful research into internet browsing history is ensuring this research is acceptable to research participants and achieves representative participation. There is currently a lack of knowledge about public and patients' perceptions of acceptability of using individual level internet search data for medical research or the criteria that would enhance engagement. This is important as low or biased engagement could limit the potential of these studies. In this study, we define acceptability as a multifaceted construct, which reflects how potential research participants consider the proposed research to be appropriate [31]. We apply the adapted theoretical framework of acceptability [32], which highlights 8 constructs relevant to establishing acceptability, to explore the feasibility and acceptability of sharing internet browsing history data for health research among individuals with and without a history of cancer, including those from disadvantaged groups.

Methods

Participants and Recruitment

Recruitment took place through in-person engagement and recruitment posters (with a study email and phone number) at community organizations, local networks, and charities using a purposive sampling strategy to ensure representation of groups at higher risk of poorer health and digital exclusion, such as individuals experiencing higher levels of deprivation, older adults, individuals with lower educational attainment, and ethnic minority groups. Participants were adults of any gender, had access to a portable web-browsing–enabled device; were able to read and speak in English; and self-reported either no history of cancer or a history of cancer as an adult (excluding nonmelanoma skin cancer) to explore whether previous experience of cancer influenced willingness to share (see Multimedia Appendix 1 and Checklist 1 for full recruitment details). Individuals interested in taking part were screened in person or by phone for eligibility and sampling by MR, provided with study materials, and a date for the interview booked in pending consent. Interviews were conducted by MR or SA and took place between September 2024 and January 2025 at the university, community organization (London), or online (any location, England), whichever was preferred by the participant. Recruitment was ceased when we had captured a broad range of experiences and attitudes across a diverse group and achieved sufficient information power [33].

Study Design

This research was led by an experienced health services research team, supported by a research assistant and a medical student. The team has expertise in psychology, anthropology, digital health, inequalities, clinical practice, and policy and was supported by patient and public involvement (PPI) contributors from diverse backgrounds (see the *Patient and Public Involvement* section), allowing us to bring together different perspectives during the study design, data collection, and analysis.

Taking a critical realist position, a mixed "reactive" and "reflective" "think-aloud" interview approach was used [34,35], in which participants were invited to speak their thoughts aloud in response to material prompted by the interviewers. After a "warm-up" think-aloud activity, participants were asked to think aloud while reading sections of a participant information sheet for a hypothetical future study, which asked people to share internet browsing data for health research (either prospective or retrospective data sharing). A second think-aloud activity involved participants viewing information on "Google Takeout" (a Google service that allows users to export data from other Google products, which has been used in previous studies to allow participants to share their internet history). To visualize this, participants were also shown an example image of internet search history and encouraged to view their own internet search history via their personal devices, if available. Participants were then asked broad questions on their attitudes to data sharing and the feasibility of taking part in internet sharing research projects. Interviews lasted approximately 60 minutes and were conducted by MR or SA.

XSL•FO
RenderX

Data Analysis

Interviews were audio-recorded and professionally transcribed verbatim, and notes were made by the interviewer as a memory aid during the think-aloud portion of the interview. A framework analysis approach was employed, guided by the adapted theoretical framework of acceptability [32]. Familiarization and initial inductive and deductive coding were conducted by NCG using NVivo (version 1.6; Lumivero). To ensure consistency in the coding and thematic framework development, MR reviewed a subset of transcripts. Code definitions, grouping, and theme development were led by NCG and refined through discussion with MR, SS, and CD via a series of consensus meetings and validated with patient and public contributors. These discussions were informed by field notes kept by SA and MR, and analysis notes were kept by NG. Microsoft Excel was used to chart codes, develop a framework matrix, and compare responses between participants with and without a history of cancer.

Ethical Considerations

Ethical approval was granted from Queen Mary University of London (QMUL) Research Ethics Committee (reference QME24.0548). The study protocol is available on OSF [36], and there were no deviations from this protocol. Informed consent was obtained from participants through a written consent form that was explained by the researcher following a detailed participant information sheet. Participation was voluntary, and participants could withdraw at any point of the study. Participants' identifying information has not been included in this study. Participants were reimbursed £25 (US $33.50) for their time.

Patient and Public Involvement

Three patient and public involvement contributors were recruited from the Centre for Cancer Screening, Prevention and Early Diagnosis (Queen Mary University of London) PPI pool. This is a working database of over 150 people who have expressed an interest in contributing to research. We requested engagement particularly from people aged 55 years or older and from a Black, Asian, or other minority ethnic background. PPI contributors discussed the overall study approach; reviewed the study documents, such as participant information sheets and interview guides; and validated the interpretation of the results and the study recommendations. Recommended changes that were implemented included improved wording on the study recruitment posters and clarification of the study purpose in the participant information sheets. PPI reporting is included in Checklist 2.

Results

Participant Characteristics

In total, 22 interviews were conducted. Two transcripts were excluded from the final analysis following team discussion. In both interviews, there was little engagement, the participants repeated information included within the question and were unable to provide details about their own experiences. As a result of this, the relevance and reliability of the data were insufficient for meaningful analysis. Half (n=10, 50%) of the participants had no history of cancer and half (n=10, 50%) had a history of cancer, representing a range of age, gender, and ethnic groups. Half of those with no history of cancer (n=5) and one third of those with a history of cancer (n=3) lived in areas described as high deprivation (ie, local postcode scores 1 or 2 according to the English Indices of Deprivation), where 1 represents the highest level of deprivation and 10 represents the lowest level of deprivation [37]. The full details of participant demographics are included in Table 1.

XSL•FO
RenderX

Table . Demographics of interview participants.

Demographic characteristic	Those with history of cancer (n=10), n	Those without history of cancer (n=10), n
Gender		
Male	2	5
Female	8	5
Ethnicity		
White/White British/White—Other	7	5
Black British/Black Caribbean	1	2
Asian British or Asian Other	2	3
Age range (y)		
19 - 34	0	1
35 - 44	1	2
45 - 54	1	2
55 - 64	6	2
65 - 74	0	1
75 - 84	2	2
Indices of deprivation		
1 - 2 (highest deprivation)	3	5
3 - 4	2	4
5 - 6	3	1
7 - 8	1	0
9 - 10 (lowest deprivation)	1	0
Highest level of education		
Primary	0	0
Secondary	3	5
Tertiary (including professional qualifications)	6	3
No qualifications	1	1
Not provided	0	1

Interview Results

The key themes focused on factors necessary for acceptability, including trust, transparency, personal control, and setting digital boundaries and on the perceived feasibility and individual willingness of participants to share internet browsing history for health research. Finally, participants offered recommendations on how future studies using this approach could facilitate trust and encourage participation.

Trust and Transparency as Foundations for Acceptability

Trust and transparency emerged as central factors influencing participants' willingness to share their browsing history data. Trust was not automatic but needed to be actively earned through clear communication, familiarity with the research team, transparency about who would access the data, and integrity in data handling. Direct interaction with researchers, a clear understanding of the study's purpose, and reassurance about data access enhanced participants' sense of security and credibility.

> *I'd feel comfortable because I've met you guys and you've explained clearly what your research is about, if it was a third party, then that's kind of weird because I've never met that person.* [P02, 35-39 years, female, no history of cancer]

Participants consistently articulated that their trust in the research team was closely tied to the perceived transparency of the team's actions and intentions. Providing clear information about the researchers' identities, data use, and opportunities to ask questions would foster reassurance and informed participation. Transparency also entailed setting clear boundaries around what data would be accessed and outlining the safeguards in place to protect privacy. Participants were generally willing to share information provided researchers strictly adhered to stated objectives, maintained ethical standards, and implemented robust data handling protocols, such as anonymization and secure deletion.

> *It is always reassuring to see that you are following all the right things and that you delete information*

XSL•FO
RenderX

> *once you've collected it for your analysis.* [P43, 75-79 years, female, history of cancer]
>
> *Because of the sensitivity of that nature, I need to be 100% sure that this data are being used in the right way, are stored, accessed by a limited number of people for the purpose that is being clearly described and given consent for.* [P54, 35-39 years, male, history of cancer]

Importantly, participants noted that it was not the nature of their data that caused concern but rather the apprehension that would arise if researchers deviated from the agreed scope of use. Such actions would raise doubts about the researchers' intentions and significantly undermine trust. Thus, trust and transparency were seen as mutually reinforcing, whereby transparency was essential to the development and maintenance of trust throughout the research process.

> *Providing that the researchers are open about what they're doing and only do what they say they're going to do, then that's fine. If they start going beyond that then it would concern me … because it makes you question why they're wanting to look for other things.* [P47, 60-64 years, female, history of cancer]
>
> *As long as it's for research … If they are properly trained researchers, they'll be ethical and respectful and will observe people's confidentiality and privacy.* [P43, 75-79 years, female, history of cancer]

Several participants highlighted the importance of having choice and autonomy, including the ability to opt out or withdraw consent at any time. These features reassured participants that their involvement remained voluntary and under their control.

> *It's quite good that it's got an opt-out clause … you don't have to take part if you don't want to and that it won't affect you in any way.* [P26, 55-59 years, female, history of cancer, responding to the PIS]

Trust in the organizations managing search data was a fundamental factor in acceptability, with participants clearly differentiating between trusted public bodies and commercial companies. Institutional affiliations with reputable entities, such as the NHS (United Kingdom) and universities, were seen as markers of legitimacy, with logos, branding, and clear consent processes being mentioned as ways of reinforcing this trust. Commercial companies were often not trusted due to their profit-driven motives and concerns about data misuse, such as selling personal information to health insurance companies leading to potential discrimination.

> *I'm not sure I saw a university or NHS trust logo … Just to give me that added confidence.* [P43, 75-79 years, female, history of cancer, responding to the PIS]
>
> *I think with the NHS it's okay, but if it's a private corporation or some commercial that engages with marketing stuff, then I'd be quite apprehensive about that.* [P02, 35-39 years, female, no history of cancer]
>
> *If it is for other reasons like health insurance, then they can target you and make you pay a higher premium, then I don't think that is fair. And then I'm not happy with it.* [P17, 60-64 years, female, no history of cancer]

Personal Control, Privacy Concerns, and Setting Digital Boundaries

Privacy and data security emerged as critical concerns for participants, reflecting broader mistrust in digital systems. The foundations of trust in data management were heavily dependent on participants' perceptions of how their personal information would be treated. While some participants were open to sharing their data for medical research, many expressed reservations about the potential misuse of their personal information or information of other individuals using their devices, as well as data breaches. They highlighted the importance of maintaining control over their personal data.

> *The only thing I'd worry is if certain things have been hacked.* [P09, 20-24 years, male, no history of cancer]
>
> *It's that right to say yes or no, I have the right to keep my private things private.* [P49, 60-64 years, female, history of cancer]

The concept of privacy was also viewed as inherently linked to transparency, with many participants seeking clear assurances about how their data would be used. Individuals described their internet history as deeply personal, a window into their habits, interests, and private moments and therefore a reflection of their day-to-day lives. Participants expressed strong concerns about the intrusiveness of using this type of data, emphasizing that it could feel like a violation of privacy if there was a risk of others accessing their data, making its use in research a potentially sensitive issue without clear ethical safeguards in place.

> *I need to be convinced about this and get some reassurances that privacy issues will be dealt with accordingly. Otherwise, I find it intrusive.* [P54, 35-39 years, male, history of cancer]

Interestingly, 6 out of the 20 participants, all those without a history of cancer, reported using incognito mode for reasons such as accessing restricted websites, watching YouTube, avoiding cookies or pop-ups, and conducting private searches related to health or banking.

Some participants expressed a strong desire to maintain autonomy over their personal information. While many would feel comfortable sharing health-related information, they were clear that they did not want to disclose any other personal or sensitive details, frequently citing financial information, political views, or family matters as examples.

> *I like to have this agency over my digital data … I wouldn't like such sensitive data to sneak in my internet use history and shared with irrelevant people.* [P54, 35-39 years, male, history of cancer]

To overcome this, several participants emphasized the importance of having a filtering system that would allow them to share only health-related data, reassuring them that irrelevant or sensitive information, such as bank details or private communications, would not be accessed. This idea of a "health filter" was repeatedly mentioned as something that would make

XSL•FO
RenderX

people feel more comfortable and in control of what was being shared.

> *I do personally think that if this research is medical, then there has to be a way that all you are taking are medical notes.* [P49, 60-64 years, female, history of cancer]

> *I mean obviously no computer programme is 100% going to sift out everything that's medical, but I think it would probably come close to answering my concerns.* [P26, 55-59 years, female, history of cancer]

Participants did not state a preference for prospective or retrospective data collection. When prompted to reflect on whether knowing they would share their data with researchers would change their behavior, only 1 participant felt that they would be on "good behavior" and potentially delete personal activities. Other participants stated they were not concerned as they "had nothing to hide," felt that "everything's big brother as it is," or that they would "forget."

> *I don't believe people would necessarily be mindful about what they look, what they search, knowing that somebody else is checking on it.* [P17, 60-64 years, female, no history of cancer]

Perceived Feasibility and Willingness to Participate in Health Research

Burden of Effort and Digital Literacy

Willingness to share internet browsing history was influenced by perceptions of effort and confidence with digital tools. While many participants were open to the idea, concerns emerged about the time and technical knowledge required to download and share these data. The process was seen by some as potentially tedious or confusing, particularly over a year-long period. One participant joked, "you'd be there all day with mine" [P40, F, 60‐64, history of cancer], reflecting anxieties about data volume and complexity.

Digital literacy varied considerably. While some participants felt comfortable navigating online platforms, others described themselves as "not expert" or relied on others for support. Tools such as Google Takeout were unfamiliar to many, and concerns were raised about storage, filtering content, and whether the process could be made simple and secure. For those less confident, the technical aspects posed as a potential barrier to engagement. These concerns were more commonly expressed by older participants, particularly those aged over 70 years, often unfamiliar with basic terms such as "browser" and relying on others for support; for example, 1 person told us: "I say I'm too old to learn" [P19, 75‐79, F, no history of cancer]. However, confidence did not always coincide with age, as some younger participants also expressed uncertainty.

> *It's too new to me, I didn't know about Google Takeout and I didn't know there was such a thing as a Google account.* [P01, 70-74 years, male, no history of cancer]

> *I don't mind sharing it as long as it's easy for me to actually share it.* [P48, 55-59 years, female, history of cancer]

Personal Experience and Online Health Information

Participants with a history of cancer expressed notably higher levels of privacy concern, emphasizing the sensitive nature of their health data and the need for strict anonymity, whereas participants without a history of cancer, while still valuing privacy, were generally more pragmatic and less emotionally driven in their concerns. These differences suggest that a cancer diagnosis may intensify apprehensions about data sharing and potential identification. Some participants with cancer also highlighted that their main interaction with online health information came after their diagnosis, limiting the utility of their browsing history for early diagnosis ("… for me, … I don't go looking for stuff on-on the internet until it's happened").

Innovation and Societal Impact

Participants recognized the potential for using internet browsing history data to improve individual health outcomes and also contribute to broader societal benefits. During the think-aloud interviews, many participants spoke more broadly about health, cancer, and internet search behavior, with some focusing on how people search for symptoms before diagnosis, while others mentioned improvements in accessing trustworthy information. Many commented on the culture of online health information, expressing frustration over "fake news" and the negative tone of many search results. There was hope that the research could help promote credible sources and bring trustworthy sources such as NHS websites to the forefront.

> *I think it's a really interesting premise, looking at if people are Googling their symptoms before they seek medical help … because intervening earlier might lead to earlier diagnosis and this would be very positive.* [P26, 55-59 years, female, history of cancer]

> *Google search symptoms can be a little bit doom and gloom … it would be great if people actually get more access [to] information … that [is] practical.* [P02, 35-39 years, female, no history of cancer]

Participants expressed strong support for the use of digital tools and online behaviors in health contexts, particularly if related to improving early diagnosis and accessing trustworthy advice.

> *If people can be helped earlier and better, then yes, I support it.* [P17, 60-64 years, female, no history of cancer]

> *Technology can really help people with health stuff.* [P02, 35-39 years, female, no history of cancer]

> *People like to be involved and want to know what effect it might lead to.* [P43, 75-79 years, female, history of cancer]

> *Only by sharing our data can progress be made in medical research.* [P47, 60-64 years, female, history of cancer]

Early diagnosis emerged as a powerful motivator for data sharing and health research participation, with several

XSL•FO
RenderX

participants emphasizing the importance of online data to prompt earlier help-seeking and therefore improve health outcomes. Participants' experiences of cancer shaped their perspectives on the potential value of the research.

> *I always feel really happy if I can help with a study that is going to absolutely make someone move forward with an early diagnosis.* [P49, 60-64 years, female, history of cancer]
>
> *We shouldn't be going to stage three in cancer. I think we should be able to fix it at stage one.* [P46, 45-49 years, female, history of cancer]

However, while people felt the research was worthwhile for early diagnosis, some questioned the usefulness of their own browsing history, describing it as "low value" or highlighting that they did not frequently use the internet to look up health questions.

Discussion

Principal Findings

This study explored the acceptability of using individual-level internet browsing history data for health research focusing on earlier cancer diagnosis. To our knowledge, no previous study has assessed the acceptability of using such data with respect to health inequalities and cancer.

In this study, participants placed greater confidence in public institutions such as the NHS and universities rather than commercial companies with fear of data misuse, discrimination, and profit motives being predominant reasons, similar to other studies [29,38,39]. Transparency played a foundational role in fostering trust, and this could be built and sustained through clearly communicating study aims, respecting participant boundaries, and keeping participants informed throughout the process.

Participants expressed ongoing concerns about the potential for their data to be hacked or misused once shared, reflecting broader mistrust in digital systems and the sensitive and private nature of internet browsing data. This sentiment was particularly strong among those with a history of cancer who linked their online searches to periods of emotional vulnerability and distress, particularly around diagnosis and treatment phases, similar to previous studies [40,41]. Younger participants took steps toward digital privacy, using "incognito mode" or alternative browsers or search providers seen as more secure, presenting a potential barrier for future research which relies on search history data.

To overcome these concerns of data intrusiveness, a key recommendation across interviews was the implementation of a robust filtering mechanism to isolate health-related data from other forms of personal information. Not only did this offer reassurance that sensitive or irrelevant information would not be accessed, but it also served as a practical mechanism to enhance participation by directly addressing concerns around data intrusion and potential misuse. Previous studies have applied health filters using a list of symptoms, diseases, and medications and their synonyms [15,42]. However, patient descriptions of symptoms may vary according to sociodemographic, linguistic, cultural, regional, and other factors [43,44], which may not be taken into account in lists developed from clinical lexicons. The definition of health-relevant information may also vary according to the condition a study focuses on: for example, negatively valenced web pages in individuals with depression [45], or typing speed and error rate in the presence of cognitive changes, such as in multiple sclerosis [46]. Future research approaches using browsing data will need to clearly define and communicate the limitations of any such filtering to participants.

Additionally, the study revealed wide variation in digital literacy, with both older adults and some younger participants expressing confusion around basic digital terms and concerns about making mistakes during the data-sharing process. This uncertainty about the technical aspects of using unfamiliar tools to download internet browsing data (such as Google Takeout) posed as a barrier, especially for those with limited digital confidence, highlighting the need for simple, user-friendly processes and tailored, individual support. Table 2 outlines key recommendations to enhance participation for future studies.

XSL•FO
RenderX

Table . Recommendations for research involving individual-level internet browsing history, as suggested by participants, linked to the study's themes and the adapted theoretical framework of acceptability (ATFA) [32].

Area (ATFA construct)	Recommendations for research involving individual-level internet browsing history
Trust and transparency (ethicality, intervention coherence)	• Provide clear and accessible information about the study and data use • Use NHS[a] or university branding to build trust where possible • Ensure participants have control, including opt-out options • Maintain open communication with participants • Ensure each participant has a point of contact to direct queries • Include clear information on the research team members including name and credentials
Privacy, data control, and digital boundaries (ethicality, trust)	• Implement a robust filtering system to isolate health-related data and exclude nonhealth-related content (eg, finances, politics, schools); potentially including the option to remove browsing activity related to, or on behalf of, others such as friends and family • Reassurance regarding anonymity, data storage, data protection, and confidentiality
Burden of effort and digital literacy (burden, self-efficacy)	• Offer simple, step-by-step guidance on how to share search history • Include visual aids or videos to support less digitally confident users • Provide technical assistance when needed
Innovation and societal impact (affective attitude, perceived effectiveness)	• Emphasize the societal benefits, such as improving early diagnosis • Highlight how the research could help others • Reassure participants that their contribution is meaningful and valued

[a]NHS: National Health Service (United Kingdom).

Comparison With Prior Work

A recent review of harnessing internet search data as a potential tool for medical diagnosis identified ethical, bias, technical, and policy challenges [47]. In our study, several key factors emerged as influencing participants' willingness to share their internet browsing data, consistent with previous research, which broadly categorized these into four key themes: (1) the relationship between the participants and the researchers, (2) fears and harms of data sharing, (3) purposes, and (4) benefits [48]. Similar recommendations were made in studies to share other digital data, such as loyalty card (purchase history) for academic research [49,50].

Across our study and other work, trust and transparency are integral components and prerequisites for willingness to share data. Trust is not a singular concept but multi-layered, operating at 3 distinct levels: interpersonal (trust in individual researchers), institutional (trust in health care or academic organizations), and systemic (trust in the digital infrastructure more broadly) [51]. Key "enablers" and "impediments" within these trust levels include sociodemographic factors and the reputation of the institution fitting into both categories [39]. Understanding these intersecting factors is key to ensuring research utilizing internet browsing history data for health research is equitable: for example, 1 study suggested that internet and social media data may be viewed by the public as more sensitive than electronic health record data [20]; another found that medical records were seen as the most sensitive type of data, but this was willingly shared with health researchers, while sharing banking data (seen as the second most sensitive) raised more concerns [50]. Attitudes to data protection and sharing may also vary by demographic group: for example, younger people and people with more education were more likely to choose an option to pay for a service to keep their data private, regardless of income level [52]. Additionally, people with long-term health conditions and those from ethnic minority backgrounds express higher concern about sharing data in general [53], and trust in researchers may vary by ethnic group [54]. Smart Data Research UK's recent public engagement also found that while information about the purpose and benefits of the research positively influenced willingness for their data to be used in research, there was low trust in government and public institutions for these findings to impact policy [55]. This highlights the need to build trust at all levels to ensure that research using digital data targets those who would benefit the most from the earlier detection of cancer and to prevent the risk of exacerbating inequalities. Participants stressed the personal nature of establishing trust ("I've met you guys"), and prospective studies must consider the practical and resource implications of maintaining these relationships over the course of a study to support sustained participant engagement.

Our study also highlighted potential issues of feasibility related to sharing internet browsing history. Previous research has utilized "Google Takeout" as a method for data porting in this type of research. This study identified potential limitations to this approach: for example, participants used other browsers (eg, did not use Android devices); did not have a Google account or were not logged in. Further fragmentation of the digital ecosystem occurs through social media channels, in-app data access, and the impact of generative artificial intelligence on how people seek and review health information—such as only reading the in-browser artificial intelligence summary or asking

XSL•FO
RenderX

ChatGPT [56]. A review of the use of internet search data in diagnosis research found no studies that integrated data from across different platforms and noted the complexity of standardizing such data for analysis and a need for supporting infrastructure and guidance on best practice.

Participants' individual behavior also impacted feasibility; for example, they avoided "cyberchondria" by not using internet searches for health, used incognito mode to hide searches, and searched for health-related data for others (leading to potentially "messy" data). One study recruiting women with gynecological symptoms reported that one quarter of potential participants did not have a Google account and nearly 30% did not have relevant data or faced technical issues with sharing their data [15]. While the median age was similar in those included and excluded in the study, other potential demographic differences were not discussed. Other studies have identified similar issues, including participants consenting to take part in the study but deciding they did not want to share their Google searches [57]. One study that focused on internet use in older age groups (mean age 81 years) found that even among individuals who used computers, 22/76 (29%) did not conduct any searches during the duration of the study, and a further 12 did not conduct any searches in a particular 3-month period [58]. Some studies have only included individuals with a Google account [59-61], but this may impact the utility of this type of research approach in sociodemographically diverse populations.

This study found that in most domains, participants with and without a history of cancer shared similar perspectives. It is possible that the differences in the use of incognito mode between the participants with and without a history of cancer may relate to individual variability in digital literacy, as other participants were not aware of this function. This was not necessarily limited by age or education level, as those who used it were within age brackets from 20 to 64 years. In contrast, participants with a history of cancer expressed notably higher levels of privacy concern, although they were similarly willing to share their data given adequate safeguards. We speculate that this may be due to their closer experiences with the health system, potentially identifying the nature of rare cancers and their focus on their browsing activity postdiagnosis. Qualitative studies in cancer survivor populations have found similar willingness to share and made similar recommendations for building trust [62], and comparisons of individuals with and without a history of cancer found similar willingness to share information and a higher willingness in those with a history of cancer to share sensitive information (such as genetic information) [63]. This suggests an opportunity to examine in more detail how individual health experiences may impact willingness to share data for research related to specific health risks.

Studies that have involved sharing of internet search logs and health data with researchers have typically been retrospective, and there is limited evidence regarding differences in search patterns in study participants who donate data prospectively [47]. A study of the observer effect on social media use did find differences in posting behavior including frequency and topics [64]. Participants were prompted to discuss this during the interview and reported that it would not impact their (long-term) browsing activities. Nevertheless, this remains an important area to explore.

Strengths and Limitations

This study aimed to explore the acceptability of sharing internet browsing history for research, ensuring representation from groups more at risk of experiencing digital and cancer inequalities. We know that 3% of the population of the United Kingdom do not use the internet, facing barriers such as cost, poor connectivity, and lack of digital skills and that people in this group are more likely to be older and/or earning less [65]. Any future research of this kind would require individuals to use the internet; this required us to find a balance when recruiting and designing the methods. Offering both online and face-to-face interviews (at a convenient location) improved accessibility and allowed those who only use the internet on their phone or on shared devices (eg, library computers) to participate. Additionally, we purposively sampled groups affected by health and digital inequalities, enhancing the relevance of the findings across diverse populations. The participants with a history of cancer were more likely to have a tertiary level of education and showed more variability in whether they lived in more or less deprived areas. However, as limited comparisons were made between the groups, we feel this sample shows sufficient variability across different groups affected by health and digital inequalities. The interviews were restricted to English to support the think-aloud exercises, and as current internet history research has focused on single-language searches. However, future approaches should consider how the internet activity of more diverse populations can be included, particularly as diaspora populations may choose to seek and act on health information from their country of origin [66]. While self-selection bias is possible as participants may have had a preexisting interest in health research or felt comfortable discussing digital data, the diversity and depth of responses suggest that a broad range of perspectives was captured. This was an exploratory study (n=20), and future studies may wish to explore specific barriers and solutions in populations at different risks of exclusion.

Conclusion

This study highlights the steps required to improve the viability of internet browsing data to support earlier diagnosis and enhance public health outcomes. While concerns around privacy, data use, and institutional trust persist, they were not seen as insurmountable. Participants generally expressed positive attitudes toward this innovation, when use is clearly linked to tangible health benefits, provided there are ethical safeguards, simple guidance, and opportunities for consent. The integration of internet browsing data with a patient's medical records may present an opportunity for identifying early signals of disease and holds promise for future research.

Importantly, this research approach offers unique opportunities to reach individuals who may face barriers to accessing traditional health care services, potentially reducing health inequalities and promoting more inclusive approaches to early diagnosis. However, the implementation of this approach requires careful attention to issues of accessibility, digital literacy, and infrastructure; otherwise, the same approach is at

XSL•FO
RenderX

risk of inadvertently exacerbating existing sociodemographic disparities.

To fully harness the benefits of internet search data in health care, systems must become more integrated, secure, and person-centered. As global health pressures grow, such as the rising incidence of cancer, leveraging digital data responsibly offers a path toward more preventative and sustainable health care strategies.

Acknowledgments

We would like to thank patient and public involvement contributors for their valuable inputs during this project. We would also like to thank participants for their time and interest in this study. No artificial intelligence tools were used in the conceptualization, data collection, analysis, or preparation of this manuscript.

Funding

Google provided unrestricted funding for this research through a Privacy Research Faculty Award to SES. SES is also supported by Barts Charity (G-001520; MRC&U0036). The funders had no role in the study design; collection, analysis, or interpretation of data; writing of the paper; and/or decision to submit for publication.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

SES acquired funding. SES and MR contributed to conceptualization and were responsible for project administration. MR and SA undertook the data collection, and NCG and MR undertook the analysis, with input from SES and CD. NCG and MR drafted the original manuscript. All authors contributed to writing – review and editing.
MM and RP are not affiliated with an organization but contributed as patient and public contributors.

Conflicts of Interest

The Data Liberation Front (based at Google) was a collaborator on the study and advised on accurate description about Google Takeout. They had no other input on the study design or conduct. The research team (SES, MR) was the sole decision-makers on research design and outputs.

Multimedia Appendix 1
Participant recruitment, interview guide, and sample participant information sheet.
[DOCX File, 120 KB - cancer_v12i1e82009_app1.docx]

Checklist 1
COREQ checklist.
[PDF File, 440 KB - cancer_v12i1e82009_app2.pdf]

Checklist 2
GRIPP checklist.
[DOCX File, 14 KB - cancer_v12i1e82009_app3.docx]

References

1. Cancer survival by stage at diagnosis for England (experimental statistics): adults diagnosed 2012, 2013 and 2014 and followed up to 2015. Office for National Statistics. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancersurvivalbystageatdiagnosisforenglandexperimentalstatistics/adultsdiagnosed20122013and2014andfollowedupto2015 [accessed 2026-01-14]
2. Crosby D, Bhatia S, Brindle KM, et al. Early detection of cancer. Science 2022 Mar 18;375(6586):eaay9040. [doi: 10.1126/science.aay9040] [Medline: 35298272]
3. DuBrock HM, Germack HD, Gauthier-Loiselle M, et al. Economic burden of delayed diagnosis in patients with pulmonary arterial hypertension (PAH). Pharmacoecon Open 2024 Jan;8(1):133-146. [doi: 10.1007/s41669-023-00453-8] [Medline: 37980316]
4. Burnet NG, Jefferies SJ, Benson RJ, Hunt DP, Treasure FP. Years of life lost (YLL) from cancer is an important measure of population burden—and should be considered when allocating research funds. Br J Cancer 2005 Jan 31;92(2):241-245. [doi: 10.1038/sj.bjc.6602321] [Medline: 15655548]

XSL•FO
RenderX

5. The NHS Long Term Plan. NHS England. 2019. URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20230418155402/https:/www.longtermplan.nhs.uk/publication/nhs-long-term-plan/ [accessed 2026-01-22]
6. Zakkak N, Barclay ME, Swann R, et al. The presenting symptom signatures of incident cancer: evidence from the English 2018 National Cancer Diagnosis Audit. Br J Cancer 2024 Feb;130(2):297-307. [doi: 10.1038/s41416-023-02507-4] [Medline: 38057397]
7. Rubin G, Berendsen A, Crawford SM, et al. The expanding role of primary care in cancer control. Lancet Oncol 2015 Sep;16(12):1231-1272. [doi: 10.1016/S1470-2045(15)00205-3] [Medline: 26431866]
8. Scott SE, Walter FM, Webster A, Sutton S, Emery J. The model of pathways to treatment: conceptualization and integration with existing theory. Br J Health Psychol 2013 Feb;18(1):45-65. [doi: 10.1111/j.2044-8287.2012.02077.x] [Medline: 22536840]
9. Whitaker KL, Scott SE, Wardle J. Applying symptom appraisal models to understand sociodemographic differences in responses to possible cancer symptoms: a research agenda. Br J Cancer 2015 Mar 31;112(Suppl 1):S27-S34. [doi: 10.1038/bjc.2015.39] [Medline: 25734385]
10. Hu X, Bell RA, Kravitz RL, Orrange S. The prepared patient: information seeking of online support group members before their medical appointments. J Health Commun 2012;17(8):960-978. [doi: 10.1080/10810730.2011.650828] [Medline: 22574697]
11. McCarthy DM, Scott GN, Courtney DM, et al. What did you Google? Describing online health information search patterns of ED patients and their relationship with final diagnoses. West J Emerg Med 2017 Aug;18(5):928-936. [doi: 10.5811/westjem.2017.5.34108] [Medline: 28874946]
12. Jadhav A, Andrews D, Fiksdal A, et al. Comparative analysis of online health queries originating from personal computers and smart devices on a consumer health information portal. J Med Internet Res 2014 Jul 4;16(7):e160. [doi: 10.2196/jmir.3186] [Medline: 25000537]
13. Van Riel N, Auwerx K, Debbaut P, Van Hees S, Schoenmakers B. The effect of Dr Google on doctor-patient encounters in primary care: a quantitative, observational, cross-sectional study. BJGP Open 2017 May 17;1(2):bjgpopen17X100833. [doi: 10.3399/bjgpopen17X100833] [Medline: 30564661]
14. Nölke L, Mensing M, Krämer A, Hornberg C. Sociodemographic and health-(care-)related characteristics of online health information seekers: a cross-sectional German study. BMC Public Health 2015 Jan 29;15(1):31. [doi: 10.1186/s12889-015-1423-0] [Medline: 25631456]
15. Barcroft JF, Yom-Tov E, Lampos V, et al. Using online search activity for earlier detection of gynaecological malignancy. BMC Public Health 2024 Mar 11;24(1):608. [doi: 10.1186/s12889-024-17673-0] [Medline: 38462622]
16. White RW, Horvitz E. Evaluation of the feasibility of screening patients for early signs of lung carcinoma in web search logs. JAMA Oncol 2017 Mar 1;3(3):398-401. [doi: 10.1001/jamaoncol.2016.4911] [Medline: 27832243]
17. Paparrizos J, White RW, Horvitz E. Screening for pancreatic adenocarcinoma using signals from web search logs: feasibility study and results. J Oncol Pract 2016 Aug;12(8):737-744. [doi: 10.1200/JOP.2015.010504] [Medline: 27271506]
18. Asch JM, Asch DA, Klinger EV, Marks J, Sadek N, Merchant RM. Google search histories of patients presenting to an emergency department: an observational study. BMJ Open 2019 Feb 20;9(2):e024791. [doi: 10.1136/bmjopen-2018-024791] [Medline: 30787088]
19. Hirst Y, Stoffel ST, Brewer HR, Timotijevic L, Raats MM, Flanagan JM. Understanding public attitudes and willingness to share commercial data for health research: survey study in the United Kingdom. JMIR Public Health Surveill 2023 Mar 23;9:e40814. [doi: 10.2196/40814] [Medline: 36951929]
20. Grande D, Mitra N, Iyengar R, et al. Consumer willingness to share personal digital information for health-related uses. JAMA Netw Open 2022 Jan 4;5(1):e2144787. [doi: 10.1001/jamanetworkopen.2021.44787] [Medline: 35072717]
21. Thompson M, Chan C, Daniels E, et al. Systematic review of health research using internet search data. Research Square. Preprint posted online on Jun 4, 2024. [doi: 10.21203/rs.3.rs-4456499/v1]
22. van Kessel R, Wong BLH, Clemens T, Brand H. Digital health literacy as a super determinant of health: more than simply the sum of its parts. Internet Interv 2022 Mar;27:100500. [doi: 10.1016/j.invent.2022.100500] [Medline: 35242586]
23. Jia X, Pang Y, Liu LS. Online health information seeking behavior: a systematic review. Healthcare (Basel) 2021 Dec 16;9(12):1740. [doi: 10.3390/healthcare9121740] [Medline: 34946466]
24. Wang X, Cohen RA. Health information technology use among adults: United States, July-December 2022. : U.S. Department of Health and Human Services and CDC; 2023 Oct 31 URL: https://www.cdc.gov/nchs/data/databriefs/db482.pdf [accessed 2026-01-14]
25. Yoon H, Jang Y, Vaughan PW, Garcia M. Older adults' internet use for health information: digital divide by race/ethnicity and socioeconomic status. J Appl Gerontol 2020 Jan;39(1):105-110. [doi: 10.1177/0733464818770772] [Medline: 29661052]
26. Wong C, Harrison C, Britt H, Henderson J. Patient use of the internet for health information. Aust Fam Physician 2014 Dec;43(12):875-877. [Medline: 25705739]
27. Adults' media use and attitudes: reports. Ofcom. 2025. URL: https://www.ofcom.org.uk/media-use-and-attitudes/media-habits-adults/adults-media-use-and-attitudes [accessed 2026-01-14]
28. The Lancet Public Health. Inequalities in cancer: a major public health concern. Lancet Public Health 2024 Mar;9(3):e147. [doi: 10.1016/S2468-2667(24)00029-X]

XSL•FO
RenderX

29. Cascini F, Pantovic A, Al-Ajlouni YA, Puleo V, De Maio L, Ricciardi W. Health data sharing attitudes towards primary and secondary use of data: a systematic review. EClinicalMedicine 2024 May;71:102551. [doi: 10.1016/j.eclinm.2024.102551] [Medline: 38533128]
30. Benevento M, Mandarelli G, Carravetta F, et al. Measuring the willingness to share personal health information: a systematic review. Front Public Health 2023;11:1213615. [doi: 10.3389/fpubh.2023.1213615] [Medline: 37546309]
31. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
32. Deja E, Peters MJ, Khan I, et al. Establishing and augmenting views on the acceptability of a paediatric critical care randomised controlled trial (the FEVER trial): a mixed methods study. BMJ Open 2021 Mar 10;11(3):e041952. [doi: 10.1136/bmjopen-2020-041952] [Medline: 33692177]
33. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444]
34. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25. [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
35. Whitty JA, Walker R, Golenko X, Ratcliffe J. A think aloud study comparing the validity and acceptability of discrete choice and best worst scaling methods. PLoS ONE 2014;9(4):e90635. [doi: 10.1371/journal.pone.0090635] [Medline: 24759637]
36. Search history and health research (SHARE) study. OSF. URL: https://osf.io/9jgp4/overview [accessed 2026-12-23]
37. English indices of deprivation 2019. Ministry of Housing CLG. URL: https://imd-by-postcode.opendatacommunities.org/imd/2019 [accessed 2026-01-14]
38. Shabani M, Bezuidenhout L, Borry P. Attitudes of research participants and the general public towards genomic data sharing: a systematic literature review. Expert Rev Mol Diagn 2014 Nov;14(8):1053-1065. [doi: 10.1586/14737159.2014.961917] [Medline: 25260013]
39. Adjekum A, Blasimme A, Vayena E. Elements of trust in digital health systems: scoping review. J Med Internet Res 2018 Dec 13;20(12):e11254. [doi: 10.2196/11254] [Medline: 30545807]
40. Phillips CA, Hunt A, Salvesen-Quinn M, et al. Health-related Google searches performed by parents of pediatric oncology patients. Pediatr Blood Cancer 2019 Aug;66(8):e27795. [doi: 10.1002/pbc.27795] [Medline: 31069926]
41. Mueller J, Jay C, Harper S, Todd C. The role of web-based health information in help-seeking behavior prior to a diagnosis of lung cancer: a mixed-methods study. J Med Internet Res 2017 Jun 8;19(6):e189. [doi: 10.2196/jmir.6336] [Medline: 28596146]
42. Yom-Tov E, Borsa D, Hayward AC, McKendry RA, Cox IJ. Automatic identification of Web-based risk markers for health events. J Med Internet Res 2015 Jan 27;17(1):e29. [doi: 10.2196/jmir.4082] [Medline: 25626480]
43. Parola A, Lin JM, Simonsen A, et al. Speech disturbances in schizophrenia: assessing cross-linguistic generalizability of NLP automated measures of coherence. Schizophr Res 2023 Sep;259:59-70. [doi: 10.1016/j.schres.2022.07.002] [Medline: 35927097]
44. Ma A, Patitsas E, Sterne J. You sound depressed: a case study on sonde health's diagnostic use of voice analysis AI. 2023 Jun 12 Presented at: FAccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency; Jun 12-15, 2023; Chicago, IL, USA p. 639-650. [doi: 10.1145/3593013.3594032]
45. Kelly CA, Sharot T. Web-browsing patterns reflect and shape mood and mental health. Nat Hum Behav 2025 Jan;9(1):133-146. [doi: 10.1038/s41562-024-02065-6] [Medline: 39572688]
46. Chen MH, Leow A, Ross MK, et al. Associations between smartphone keystroke dynamics and cognition in MS. Digit Health 2022;8:20552076221143234. [doi: 10.1177/20552076221143234] [Medline: 36506490]
47. Downing GJ, Tramontozzi LM, Garcia J, Villanueva E. Harnessing internet search data as a potential tool for medical diagnosis: literature review. JMIR Ment Health 2025 Feb 11;12:e63149. [doi: 10.2196/63149] [Medline: 39813106]
48. Howe N, Giles E, Newbury-Birch D, McColl E. Systematic review of participants' attitudes towards data sharing: a thematic synthesis. J Health Serv Res Policy 2018 Apr;23(2):123-133. [doi: 10.1177/1355819617751555] [Medline: 29653503]
49. Dolan EH, Shiells K, Goulding J, Skatova A. Public attitudes towards sharing loyalty card data for academic health research: a qualitative study. BMC Med Ethics 2022 Jun 7;23(1):58. [doi: 10.1186/s12910-022-00795-8] [Medline: 35672737]
50. Shiells K, Di Cara N, Skatova A, et al. Participant acceptability of digital footprint data collection strategies: an exemplar approach to participant engagement and involvement in the ALSPAC birth cohort study. Int J Popul Data Sci 2022;5(3):1728. [doi: 10.23889/ijpds.v7i1.1728] [Medline: 35519823]
51. Naeem I, Quan H, Singh S, et al. Factors associated with willingness to share health information: rapid review. JMIR Hum Factors 2022 Feb 9;9(1):e20702. [doi: 10.2196/20702] [Medline: 35138263]
52. Skatova A, McDonald R, Ma S, Maple C. Unpacking privacy: valuation of personal data protection. PLOS ONE 2023;18(5):e0284581. [doi: 10.1371/journal.pone.0284581] [Medline: 37134067]

XSL•FO
RenderX

53. What we know from engaging with the public on data: September 2025. Office for National Statistics. 2023 Jun. URL: https://www.ons.gov.uk/aboutus/usingpublicdatatoproducestatistics/peoplesattitudestodata/whatweknowfromengagingwiththepublicondatajune2023 [accessed 2026-01-22]
54. Hu WT, Bergren SM, Dychtwald DK, Ma Y, Dong X. Variations in racial and ethnic groups' trust in researchers associated with willingness to participate in research. Humanit Soc Sci Commun 2023;10(1):466. [doi: 10.1057/s41599-023-01960-z] [Medline: 38650745]
55. Smart Data Research UK: public dailogue 2025. : Smart Data Research UK; 2025 URL: https://www.sdruk.ukri.org/wp-content/uploads/2025/05/SDR-UK-Public-Dialogue-report-2025-1.pdf [accessed 2026-01-14]
56. Lopez-Lopez E, Abels CM, Holford D, Herzog SM, Lewandowsky S. Generative artificial intelligence-mediated confirmation bias in health information seeking. Ann N Y Acad Sci 2025 Aug;1550(1):23-36. [doi: 10.1111/nyas.15413] [Medline: 40716036]
57. Areán PA, Pratap A, Hsin H, et al. Perceived utility and characterization of personal google search histories to detect data patterns proximal to a suicide attempt in individuals who previously attempted suicide: pilot cohort study. J Med Internet Res 2021 May 6;23(5):e27918. [doi: 10.2196/27918] [Medline: 33955838]
58. Austin J, Hollingshead K, Kaye J. Internet searches and their relationship to cognitive function in older adults: cross-sectional analysis. J Med Internet Res 2017 Sep 6;19(9):e307. [doi: 10.2196/jmir.7671] [Medline: 28877864]
59. Zaman A, Zhang B, Silenzio V, Hoque E, Kautz H. Individual-level anxiety detection and prediction from longitudinal YouTube and Google search engagement logs. arXiv. Preprint posted online on Jul 1, 2020. [doi: 10.48550/arXiv.2007.00613]
60. Zhang B, Zaman A, Silenzio V, Kautz H, Hoque E. The relationships of deteriorating depression and anxiety with longitudinal behavioral changes in Google and YouTube use during COVID-19: observational study. JMIR Ment Health 2020 Nov 23;7(11):e24012. [doi: 10.2196/24012] [Medline: 33180743]
61. Birnbaum ML, Kulkarni PP, Van Meter A, et al. Utilizing machine learning on internet search activity to support the diagnostic process and relapse detection in young individuals with early psychosis: feasibility study. JMIR Ment Health 2020 Sep 1;7(9):e19348. [doi: 10.2196/19348] [Medline: 32870161]
62. Johnston EA, Bourdaniotis XE, Ayre SK, Zajdlewicz L, Beesley VL, Goodwin BC. Data sharing in cancer research: a qualitative study exploring community members' preferences. Cancer Med 2024 Aug;13(15):e70084. [doi: 10.1002/cam4.70084] [Medline: 39126200]
63. Grande D, Asch DA, Wan F, Bradbury AR, Jagsi R, Mitra N. Are patients with cancer less willing to share their health information? Privacy, sensitivity, and social purpose. J Oncol Pract 2015 Sep;11(5):378-383. [doi: 10.1200/JOP.2015.004820] [Medline: 26265174]
64. Saha K, Gupta P, Mark G, Kiciman E, De Choudhury M. Observer effect in social media use. 2024 May 11 Presented at: CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems; May 11-16, 2024. [doi: 10.1145/3613904.3642078]
65. UK consumer digital index and essential digital skills report 2024. : Lloyds Banking Group; 2024 URL: https://www.lloydsbankinggroup.com/assets/pdfs/media/consumer-digital-index/2024-consumer-digital-index-report.pdf [accessed 2026-01-14]
66. Ramasawmy M, Sunkersing D, Persson DR, et al. "If it all goes digital, we'll have to learn": facilitators and barriers to uptake of digital health in British South Asians with cardiometabolic disease. BMC Digit Health 2024;2(1):32. [doi: 10.1186/s44247-024-00083-y]

Abbreviations

NHS: National Health Service
PPI: patient and public involvement

Edited by N Cahill; submitted 19.Aug.2025; peer-reviewed by AM Yacob, D Wong, L Raymond Guo; revised version received 18.Dec.2025; accepted 18.Dec.2025; published 02.Feb.2026.

Please cite as:
Gradwell NC, Ramasawmy M, Arora S, Mistry M, Punjabi R, Derksen C, Scott SE
Acceptability of Sharing Internet Browsing History for Cancer Research: Think-Aloud and Interview Study
JMIR Cancer 2026;12:e82009
URL: https://cancer.jmir.org/2026/1/e82009
doi:10.2196/82009

© Nicola Cara Gradwell, Mel Ramasawmy, Sanjula Arora, Manoj Mistry, Reshma Punjabi, Christina Derksen, Suzanne E Scott. Originally published in JMIR Cancer (https://cancer.jmir.org), 2.Feb.2026. This is an open-access article distributed under the

XSL•FO
RenderX

terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Health Care Providers' Perceptions of Unmet Needs Among African American Cancer Caregivers: Qualitative Investigation Among US Medical Professionals

Brad Love[1], PhD; Gerold Dermid[1], MBA; Sean Upshaw[1], PhD; Amy Stark[2], MA

[1]Center for Health Communication, The University of Texas at Austin, Austin, United States

[2]Gryt Health, Rochester, NY, United States

Corresponding Author:
Brad Love, PhD
Center for Health Communication
The University of Texas at Austin
300 W Dean Keeton St
Austin, 78712
United States
Phone: 1 5125572965
Email: lovebrad@utexas.edu

Abstract

Background: African American caregivers are more likely to be sole unpaid caregivers, spend more hours on caregiving tasks, and receive less external support compared to White caregivers; yet, limited research focuses on their specific needs. Even less attention has been paid to health care provider perspectives on how to better support this population, despite providers' critical role in connecting caregivers to resources and implementing systems-level changes.

Objective: This study aimed to understand health care providers' experiences supporting African American cancer caregivers and to identify actionable recommendations for improving care. Specific objectives were to (1) identify unmet needs that providers observe among African American cancer caregivers, (2) explore barriers preventing these needs from being met, and (3) elicit provider recommendations for interventions to enhance caregiver support.

Methods: Between January and May 2023, we conducted semistructured online interviews with 12 health care providers across 7 US states. Providers were purposively sampled from facilities serving patient populations with ≥20% African American representation. Participants included physicians (n=7), social workers (n=2), nurses (n=2), and other providers (n=1), with 58% identifying as Black or African American and 83% having more than 15 years of clinical experience. Interviews lasted ~60 minutes and were conducted via Zoom (Zoom Video Communications, Inc) with audio recording. Data were analyzed using condensed thematic analysis guided by the McKillip needs assessment framework and socioecological model.

Results: Thematic analysis revealed 2 overarching categories of findings. First, providers identified three types of unmet needs among African American cancer caregivers (1) practical needs, including transportation, financial constraints, and competing family obligations; (2) social-emotional needs, including stress, burnout, and fear; and (3) cultural barriers, including medical mistrust rooted in historical trauma, "superhero Black woman" expectations, tensions between faith and medical treatment, and stigma around mental health. Second, providers offered four themes of recommendations for transformational change: (1) formal acknowledgment and compensation of caregiving as essential work; (2) integration of caregivers as equal members of multidisciplinary care teams; (3) recognition and leveraging of cultural assets, including strong family networks, community values, and faith-based support; and (4) strengthening providers' roles as hubs for individual-level support and systems-level advocacy.

Conclusions: Health care providers readily identify substantial unmet needs among African American cancer caregivers and offer practice-based recommendations that extend beyond individual-level support to emphasize structural and systems transformation. Findings suggest that meaningful improvement requires multilevel intervention. This includes policy changes to formalize and compensate caregiving work, organizational restructuring to integrate caregivers into care teams, provider training in cultural humility and asset-based approaches, and institutional commitment to addressing historical trauma and rebuilding trust with African American communities. This novel provider-focused approach offers actionable pathways for clinical settings to reduce disparities and improve outcomes for African American cancer caregivers and the patients they support.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e76266) doi:10.2196/76266

KEYWORDS

cancer caregivers; African-American health; Black health; cancer survivorship; provider perceptions; health care disparities; support needs; qualitative research; cultural competence; health care systems

Introduction

Cancer Caregiving

As cancer survivorship increases in the United States due to medical advances, the burden of long-term caregiving has grown substantially [1-5]. This burden falls disproportionately on communities of color, who face higher rates of chronic disease, greater caregiving responsibilities, and significant structural barriers, including limited health care access, lower socioeconomic status, and systemic medical mistrust rooted in historical trauma [6-8].

Caregivers play a crucial role in addressing the health and social needs of individuals with chronic conditions, providing essential support that significantly impacts patient outcomes [5]. For instance, cancer caregivers dedicate an average of 32.9 hours per week to caregiving tasks, with 32% reporting more than 40 hours of caregiving weekly [6]. Informal caregiving, which involves providing care without formal employment or financial compensation, is particularly common [4]. However, the burden of informal caregiving falls disproportionately on communities of color [7,8].

Systematic reviews and meta-analyses focused on caregiving within communities of color consistently reveal negative effects on caregivers' physical and mental health [9-11]. One meta-analysis identified a significant association between caregiving and increased levels of self-reported depression and stress, along with reduced general well-being [10]. Additionally, caregivers were found to have lower odds of engaging in preventive health services and personal wellness activities compared to noncaregivers, often prioritizing the needs of those they care for over their own health [12,13].

Studies of access to medical support also show significant gaps. For instance, 54% of caregivers reported having no contact with a health care team in the past year [14]. Among cancer caregivers, many performed complex medical tasks without prior training, and few addressed their own self-care needs during medical visits [6].

Prior Work on African American Caregivers

The prevalence of caregiving is higher among African American (28.1%) and Hispanic (21.9%) populations compared to White populations (19.8%) [7]. However, these groups often face greater health disparities, including higher rates of negative health outcomes, limited access to health care, lower health literacy, and lower income and education levels than White populations [15]. Furthermore, caregivers from diverse backgrounds frequently receive less empathetic communication from health care providers, obtain less medical information, are less involved in medical decision-making, and feel less confident in requesting training or social assistance from health care professionals [16,17].

African Americans are more likely to be the sole informal and unpaid caregivers compared to White caregivers (55% vs 44%), more likely to live full-time with the individuals they care for (45% vs 36%), and more frequently report receiving no external support for their caregiving duties (41% vs 30%) [7]. Additionally, they tend to spend more hours per week on caregiving tasks and are more likely to assist with both activities of daily living, such as managing incontinence and bathing, and instrumental activities of daily living, including medication and financial management. Despite these increased responsibilities, African American caregivers are only half as likely to receive respite services compared to their White counterparts [7].

Caregiving can negatively impact caregivers' health, with only 34% of African American caregivers rating their health as excellent or very good, compared to 45% of White caregivers [7]. African American caregivers also face greater financial challenges, reporting lower household incomes and higher levels of financial strain than their White counterparts [18-21]. Among African American caregivers, 74% have had to modify their employment status due to difficulties balancing work and caregiving responsibilities [19]. Social values, such as familyism—which emphasizes prioritizing family needs over individual needs—are prevalent in African American and other communities of color, adding further strain and contributing to poorer health outcomes for caregivers [12,19]. Additional complexities can arise from the fact that only 3% of oncologists in the United States are Black [22], limiting the potential positive effects of racial concordance on patient-provider (and caregiver-provider) interactions during care visits [23,24].

The Critical Role of Provider Perspectives in Understanding Caregiver Needs

Health care provider perspectives on African American cancer caregiver needs remain understudied. This represents a critical gap given providers' roles as gatekeepers to resources, care coordinators, and drivers of institutional change. Providers like physicians, nurses, and social workers occupy a unique structural position at the intersection of individual patient and caregiver needs and institutional resources and policies [25]. Unlike caregivers who experience their own individual challenges, providers are in a unique position to observe patterns across multiple patients and caregivers, allowing them to identify common issues, effective strategies, and systemic gaps that may not be apparent from individual caregiver accounts.

Provider awareness, attitudes, and behaviors directly influence whether and how caregivers receive support during clinical encounters [26]. Providers who recognize caregiver burden and prioritize caregiver needs are more likely to initiate conversations about support services, make appropriate referrals, and engage caregivers as partners in care planning. Conversely, when providers lack awareness of caregiver challenges or hold implicit biases about caregiver capacity based on race or

XSL•FO
RenderX

socioeconomic status, caregivers may be overlooked in care processes despite having substantial unmet needs [27].

Providers also serve as the primary conduit through which caregivers access formal support services, including social work, care navigation, palliative care, respite services, and community resources [28]. One of the strongest predictors of patients and caregivers accessing support services is provider recommendation. Understanding what providers observe about caregiver needs and what improvement recommendations they offer directly informs how to optimize referral pathways, given their position to influence institutional and systems-level change [29]. They often hold positional authority that allows them to champion policy changes, implement new care models, secure resources, and influence organizational priorities.

Further, understanding provider perspectives can identify discrepancies between caregiver-reported needs and provider perceptions, revealing potential blind spots in current health care delivery across specialties and well beyond oncology [28,30]. When providers underestimate caregiver burden, misunderstand cultural factors shaping caregiver experiences, or fail to recognize specific issues faced by African American caregivers, these gaps in awareness translate directly into inadequate support. No studies, to our knowledge, have specifically examined provider perspectives on the unique challenges and support needs of African American cancer caregivers, despite well-documented disparities in both cancer outcomes and caregiving experiences within this population [31].

Goals of This Study

Existing research addresses caregivers' general needs, but there is a gap regarding African American cancer caregivers' specific needs and even more so considering how providers might play a supportive and constructive role in improving the caregiving experience. This study aims to engage with this understudied area by describing the unique challenges faced by these caregivers in three key ways: (1) within the context of cancer caregiving, (2) specific to the experiences of African American caregivers, and (3) from the provider's viewpoint, rather than solely from the caregiver's perspective. This study seeks to offer valuable insights into how health care professionals can better assist and work alongside African American cancer caregivers, guided by the following research questions:

1. What are health care providers' experiences in supporting the needs of African American cancer caregivers?
2. What unmet needs do health care providers observe among African American cancer caregivers, and what barriers prevent African American cancer caregivers from having their needs met?
3. What recommendations do providers offer for improving support for African American cancer caregivers at individual, organizational, and systems levels?

This provider-focused approach offers unique insights into systemic barriers and solutions that can inform practice improvements and policy changes to better support African American cancer caregivers.

Innovation and Implications for Participatory Cancer Care

This study offers methodological and conceptual innovations aligned with advancing participatory, patient-centered cancer care. Methodologically, we apply a needs assessment framework through the provider lens to explicitly link identified needs to actionable, multilevel solutions [32]. Unlike traditional needs assessments that document gaps without pathways to change, our approach positions providers as both informants about needs and architects of solutions, recognizing their capacity to translate insights into implementable interventions.

Conceptually, this research advances participatory care models by identifying mechanisms to formally integrate caregivers as essential members of multidisciplinary care teams rather than peripheral figures. By centering provider perspectives on caregiver integration, we identify specific strategies for operationalizing truly collaborative care that includes patients, caregivers, and health care professionals as partners.

Our findings also have direct implications for designing technology-enabled interventions responsive to African American caregiver needs. Provider recommendations regarding accessibility, communication, and cultural responsiveness can inform the development of telehealth platforms, mobile health (mHealth) apps, patient and caregiver portals, and digital care coordination tools that address rather than perpetuate existing disparities.

Finally, this work directly addresses health equity priorities in cancer care by focusing on an underserved population and identifying concrete, systems-level strategies for improvement, providing actionable pathways for health care organizations committed to reducing disparities in cancer survivorship.

Methods

Study Context and Background: Research Context and Rationale

This research emerged from identified gaps in understanding the specific needs and experiences of African American cancer caregivers within US cancer care systems. Provider perspectives are essential for several reasons. First, providers are positioned at the intersection of individual patient and caregiver needs and institutional resources, making them uniquely situated to identify both clinical-level and systems-level barriers and potential solutions. Second, provider awareness and prioritization of caregiver needs directly influences whether and how caregivers receive support during clinical encounters. Third, providers observe patterns across multiple patients and caregivers that individual caregivers may not recognize, offering a broader view of common challenges and effective strategies. Finally, understanding provider perspectives illuminates opportunities for professional education, clinical practice improvement, and institutional policy change that can create more supportive environments for African American cancer caregivers.

Institutional and Geographic Context

Participants in this study represented diverse cancer care settings, including academic medical centers, community

XSL•FO
RenderX

hospitals, comprehensive cancer centers, and outpatient oncology clinics, representing the range of settings where cancer care is delivered. This diversity of institutional contexts provides insights into how different care delivery models and resource environments shape both the challenges caregivers face and the support providers can offer.

Temporal and Sociopolitical Context

Data collection occurred between January and May 2023, a period marked by heightened national attention to health equity and racial disparities in health care. This timeframe followed widespread recognition during the COVID-19 pandemic of the disproportionate health impacts on African American and other marginalized communities, as well as growing public discourse about systemic racism in health care institutions. National initiatives from organizations including the American Cancer Society, the National Cancer Institute, and the Centers for Disease Control and Prevention were emphasizing the urgent need to address cancer disparities affecting African American populations. Professional medical organizations were increasingly calling for cultural humility training, diversification of the health care workforce, and structural changes to reduce disparities.

This sociopolitical context is relevant to understanding provider perspectives captured in this study. Providers interviewed during this period were practicing within an environment of increased awareness, institutional pressure, and professional discourse around health equity. Their insights may reflect not only their direct experiences but also evolving consciousness about systemic factors contributing to disparities and the imperative for change. This context may have made providers more willing to engage in critical reflection about shortcomings in current care delivery and more open to discussing unmet needs, structural barriers, and potential solutions. Throughout this paper, we use “unmet needs” to describe gaps in support or resources required by caregivers and “barriers” to describe specific obstacles preventing those needs from being met.

Funding and Study Development Context

This study represents a secondary analysis of interview data originally collected under funding from EMD Serono to Gryt Health, a patient and caregiver support organization. The original data collection was designed to inform the development of culturally responsive support programs and resources for African American cancer caregivers. The partnership between a pharmaceutical company (EMD Serono), a patient advocacy organization (Gryt Health), and academic researchers (University of Texas at Austin) reflects a collaborative model increasingly common in patient-centered outcomes research, wherein industry funding supports community-focused research with academic rigor and independence.

The current manuscript represents an independent academic analysis conducted by the research team coordinated at the University of Texas at Austin under appropriate institutional review board approval and data use agreement (though not all authors are affiliated with the university). While the original project produced internal reports to guide program development at Gryt Health, the current analysis employs different theoretical frameworks, such as the McKillip needs assessment model [32] and social-ecological theory [15]. Also, this analysis uses analytical approaches such as condensed thematic analysis with multilevel intervention classification [33] to generate novel insights for academic dissemination and broader health care system application. This secondary analysis allows extraction of maximum knowledge value from rich qualitative data while adhering to ethical principles of research efficiency by avoiding redundant data collection that would place additional burden on health care providers already stretched thin by clinical demands.

The research team maintained complete independence in conducting analysis, interpreting findings, and preparing this manuscript for publication. EMD Serono and Gryt Health had no role in the design of the secondary analysis, interpretation of findings reported here, manuscript preparation, or the decision to submit for publication.

Recruitment and Data Collection

Semistructured interviews were selected as the optimal method to explore provider experiences and elicit detailed, contextually rich insights that cannot be captured through quantitative approaches. Qualitative methods are particularly appropriate for this study because they enable exploration of under-researched topics where validated measurement instruments do not yet exist, allow investigation of the “how” and “why” questions underlying observed patterns, and generate rich contextual understanding needed to inform culturally responsive interventions [33-35]. Also, the semistructured interview format allowed providers to share experiences in their own words while ensuring core topics were consistently addressed across interviews. Twelve health care providers participated in online key informant interviews conducted via Zoom (Zoom Video Communications, Inc) between January and May 2023. Participants represented a range of professional roles with diverse experiences and firsthand knowledge of supporting African American cancer caregivers in their clinical practice.

Sampling Strategy and Inclusion Criteria

Participants were recruited via purposive sampling with snowball recruitment according to explicit inclusion criteria. To be eligible, participants had to be 18 years of age or older and self-identify as health care providers who support African American cancer caregivers in their professional role. Additionally, providers had to work in facilities serving patient populations with at least 20% African American representation. This threshold was established to ensure that providers had substantial, sustained experience working with African American patients and caregivers rather than relying on limited or anecdotal encounters. The 20% threshold exceeds the overall US African American population proportion (approximately 13%), ensuring recruitment of providers with above-average exposure to these communities and deeper understanding of the specific challenges and strengths characteristic of African American cancer caregiving experiences. We acknowledge this threshold is somewhat arbitrary and represents a pragmatic research decision rather than an empirically derived cutoff.

XSL•FO
RenderX

The research team engaged in directed outreach through systematic review of LinkedIn (LinkedIn Corporation) profiles and affiliated faculty websites of providers who met the inclusion criteria. This directed outreach focused on facilities providing cancer care with significant African American patient populations and, when possible, providers with lived experiences as people of color themselves. After each interview, participants were invited to refer others in their professional networks to the study in a snowball sampling approach. In total, 25 providers were contacted for participation and 12 completed interviews (48% response rate).

The sample size of 12 was established a priori based on qualitative research guidelines indicating that samples of 6-15 participants are sufficient for focused research questions [36,37]. This sample size was deemed appropriate given the study's high "information power" (the richness of data from knowledgeable participants) [37] stemming from the narrow study aim (provider experiences with a specific caregiver population), sample specificity (participants with substantial experience supporting African American cancer caregivers), established theoretical framework [32], quality of dialogue (in-depth interviews), and systematic analysis strategy [36]. Data saturation was observed beginning at approximately the ninth interview, with no fundamentally new themes emerging in the final 3 interviews. The sample size was further constrained by the limited pool of eligible providers; fewer than 12% of oncologists in the United States are from underrepresented minorities, and fewer than 3% are Black, limiting the available recruitment pool [22].

Positionality Statement

Qualitative research is inherently shaped by the perspectives and experiences of researchers. To promote transparency, we acknowledge how our research team's characteristics may have influenced this study. The research team included BL, PhD (health communication researcher with 20 years of experience in cancer caregiving research and personal experience as a cancer caregiver); AS, MA (anthropology background focused on marginalized groups, 10 years amplifying patient voices); GD (doctoral student in community health with personal caregiving experience and diverse urban upbringing); and SU, PhD (health communication professor with 10+ years of qualitative research at a university cancer center).

Our team's racial and professional diversity provided multiple perspectives during data collection and analysis. Team members brought both insider perspectives (shared racial and ethnic identities with many participants and caregivers, personal caregiving experiences) and outsider perspectives (different racial backgrounds, academic positions) that shaped our interpretations. We acknowledge this study involved layered interpretation; providers interpreted caregiver experiences, and we then interpreted provider accounts. This twice-removed relationship to caregivers' lived experiences means both provider perspectives and our interpretations are filtered through our own frameworks and potential blind spots.

Our research was informed by commitments to health equity, asset-based rather than deficit-based framing of communities, and the belief that providers can play active roles in addressing systemic inequities. Two team members' affiliation with Gryt Health reflected engagement with patient and caregiver communities while potentially biasing us toward pragmatically feasible solutions. To mitigate bias, we employed multiple independent coders, maintained reflexive memos throughout analysis, actively sought contradictory evidence, and grounded findings in direct quotations that allow readers to assess our interpretations. We recognize our positionalities likely influenced which themes drew our attention and how we interpreted provider statements, particularly regarding structural solutions and cultural assets.

Interview Procedures and Data Collection

All interviews were conducted online via Zoom video conferencing software. The online format was selected to remove geographic limits to participation and allow recruitment of providers from diverse locations across the United States.

Interviews were conducted by 2 members of the research team. BL, PhD (Associate Professor of Health Communication and Chief Research Officer at Gryt Health), conducted 75% (9/12) of interviews, and AS, MA (Senior Vice President for Outcomes and Impact at Gryt Health), conducted the remaining 25% (4/12) of interviews. Both interviewers have extensive training in qualitative research methods through their graduate education and substantial experience with cancer caregiver research (see Positionality Statement for detailed backgrounds).

Interviews were scheduled at various times throughout the week and day to accommodate the complex and variable schedules of health care providers. Interview duration ranged from 45 to 60 minutes (mean 54 minutes, SD 4). This timeframe was selected to respect providers' time constraints while allowing sufficient depth of exploration. The shorter duration compared to many qualitative interviews (often 60-90 minutes) was a deliberate accommodation to the realities of provider availability and was deemed appropriate given the focused nature of the research questions and the providers' ability to offer rich, experience-based insights efficiently.

Interview Guide Development and Pilot Testing

The semistructured interview guide (Multimedia Appendix 1) was developed collaboratively by the research team through an iterative process. Development was informed by the McKillip needs assessment framework [32], which emphasizes identifying community needs, available resources, and gaps requiring intervention, as well as by a comprehensive literature review on caregiver experiences, African American health disparities, and provider perspectives in cancer care. The guide was designed to balance structure, coverage of key topics, and flexibility to allow providers to share unexpected insights and follow emergent themes.

To ensure clarity, cultural sensitivity, and appropriateness of language, the interview guide underwent rigorous review and pilot testing before full implementation. The guide was first reviewed by 2 community stakeholders, a cancer survivor and a community health manager, both with lived experience relevant to cancer caregiving in underserved communities. Their feedback informed refinements to question wording, sequencing, and cultural framing. Following this review, the guide was pilot tested with 2 health care providers who met the study inclusion

XSL•FO
RenderX

criteria but were not part of the final sample. Pilot testing revealed that some questions required refinement to reduce interview length and that certain prompts could be reworded to encourage more open-ended, expansive responses rather than brief confirmatory answers. Based on pilot feedback, the research team made minor adjustments to shorten the overall guide while maintaining coverage of essential topics and revised question phrasing to enhance elicitation of detailed narratives. No major structural changes were required, as pilot testers confirmed the guide was effective in exploring the intended topics.

Audio Recording and Data Capture

All interviews were audio recorded using Zoom's recording function with participants' informed consent (as described in Ethical Considerations). To enhance data quality and capture nuances that might not be fully represented in automated transcription, each interview was accompanied by a dedicated note taker who documented detailed observations in real time. The note taker recorded not only the content of responses but also observed emphases, emotional tone, pauses, and other contextual elements that provide interpretive depth.

Following each interview, Zoom's automated transcription feature was used to generate initial transcripts of the audio recordings. These automated transcripts, along with the detailed notes taken during interviews, were then systematically verified for accuracy against the original audio recordings by research team members, who corrected transcription errors, verified observations, and began considering emerging themes to ensure comprehensive and accurate data capture. This dual approach of combining Zoom automated transcription with human-verified contextual notes provided both efficiency and rigor in data documentation.

Data Analysis Procedures

Interview data were stored and coded using Atlas.ti software (ATLAS.ti Scientific Software Development GmbH) for Mac (RRID:SCR_022920). A condensed thematic analytic approach was used to identify key themes. The methods of this analysis were guided by the McKillip model of needs assessment [32], which emphasizes identifying specific community needs, available resources, unmet needs, and linking each topic to community-driven solutions. The analysis also employed a socioecological framework [38,39] to categorize themes at micro (individual), meso (organizational or community), or macro (structural or policy) levels of intervention.

A socioecological framework allows for conceptualization of health determinants and interventions across multiple interconnected levels (1) micro level (individual-level factors affecting individual caregivers and their immediate support systems), (2) meso level (organizational and community-level factors including health care systems, community organizations, and social networks), and (3) macro level (structural and policy-level factors including health care policy, reimbursement structures, and societal norms). This multilevel framework recognizes that comprehensive solutions require coordinated change across individual, organizational, and structural domains. This classification helps identify which stakeholders and systems must be engaged to address each category of unmet needs.

Preliminary analysis was completed by a PhD-level graduate research assistant and a member of the publication team, who together created a coding scheme through reading 3 interview transcripts independently and then using an open coding technique to identify themes within the transcripts following a modified thematic analysis informed by Strauss and Corbin [40]. The analysis included both inductive and deductive reasoning that followed open coding, axial coding, and then selective coding. Emerging themes related to the research questions were established.

Provider experiences with African American cancer caregivers were coded using the thematic analysis approach to better understand the lived experiences of providers in supporting the unique needs of these caregivers. Notes taken by the research team during the interview process were reviewed for each interview, and memos were included with the data for the initial development of a codebook. Codes were developed into a themed codebook starting with the "thickest" file and then building upon the initial codebook with every subsequent transcript and note analysis. Subcodes were created and evidence through cited quotations of interviewees was placed within the matrix.

Trustworthiness and Rigor

To promote scientific rigor, we used multiple strategies aligned with Lincoln and Guba's criteria for qualitative research trustworthiness: credibility, transferability, dependability, and confirmability [41].

Credibility was established through purposive sampling of experienced providers with substantial exposure to African American cancer caregivers (facilities with ≥20% African American patient populations), detailed field notes documenting content and contextual elements during interviews, dual data capture combining Zoom transcription with human verification against audio recordings, and probing follow-up questions to deepen understanding and verify interpretation of participants' meanings in real time.

Transferability was enhanced through description enabling readers to assess applicability to their own contexts, including comprehensive participant characterization (provider types, racial or ethnic backgrounds, years of experience, and geographic locations), explicit articulation of sampling procedures and inclusion criteria rationale, description of temporal and sociopolitical context, and transparent reporting of all procedures from recruitment through analysis.

Dependability was addressed through systematic documentation of all research procedures and maintenance of an audit trail, including recruitment contacts and response rates, interviewer assignments, data verification procedures, dated codebook versions, coding decisions and rationale, team meeting notes, and reflexive memos. Atlas.ti software provided systematic, transparent organization of coded data. The semistructured interview guide ensured consistency across interviews while allowing flexibility for emergent themes.

XSL•FO
RenderX

Confirmability was enhanced through independent coding by 2 team members who compared codes, discussed discrepancies, and developed the codebook iteratively through regular team meetings until consensus was reached. Atlas.ti provided transparency in coding decisions, allowing verification by other team members. Reflexive memos documented interpretive decisions, assumptions, and potential biases, which were discussed in team meetings to make implicit processes explicit and subject to critical examination. Both inductive and deductive coding approaches balanced exploratory discovery with theoretical structure, reducing the risk of purely subjective interpretation. Extensive use of direct quotations in reporting allows readers to assess whether interpretations are well-grounded in participant language.

Data saturation, as noted in the recruitment and data collection section, was observed beginning at approximately the ninth interview, with core themes clearly emerging and subsequent interviews reinforcing rather than expanding these themes. No fundamentally new themes emerged in the final 3 interviews. While sample size was constrained by the limited pool of eligible providers (fewer than 12% of US oncologists are from underrepresented minorities, and <5% are Black), the observation of saturation combined with focused research questions and high "information power" of experienced, knowledgeable participants provides confidence that findings meaningfully capture provider perspectives.

Ethical Considerations

This study received exempt determination from the Institutional Review Board at the University of Texas at Austin (Protocol #STUDY00005543). The research met criteria for exemption under 45 CFR 46.104(d)(2), as it involved interviews with health care professionals about their professional experiences and observations in their roles as providers, rather than direct research involving vulnerable populations or sensitive personal information about the participants themselves. Although the study involved discussion of patient and caregiver populations, participants (providers) were asked to reflect on their professional practices and aggregate experiences without sharing identifiable information about specific patients or caregivers under their care.

The data used in this manuscript were originally collected by Gryt Health under funding from EMD Serono. The current analysis represents a secondary analysis of these interview data, conducted under an independent research protocol at the University of Texas at Austin. Data sharing between institutions is governed by the Nondisclosure Data Use License Agreement UTAUS-DUA00001286, which establishes terms for appropriate use, storage, and reporting of the data in accordance with research ethics standards and participant protections established during the original data collection.

All participants provided verbal informed consent before interview participation. Given the online nature of the interviews conducted via Zoom, verbal consent was deemed appropriate and was audio recorded as documentation. Prior to beginning each interview, the interviewer reviewed the following elements with each participant the purpose of the study (to understand health care providers' experiences supporting African American cancer caregivers and to identify recommendations for improved support), study procedures (participation would involve a single 45-60 minutes semistructured interview conducted via Zoom), voluntary nature of participation (participants could decline to answer any question and could withdraw from the study at any time without penalty), recording practices (the interview would be audio recorded using Zoom's recording function for transcription and analysis purposes), data use (deidentified interview data would be used for research analysis, academic publications, and to inform development of caregiver support resources), right to ask questions (participants were invited to ask questions before, during, or after the interview), and contact information (participants were provided with contact information for the research team and the institutional review board).

Participants verbally confirmed their understanding of these elements and their willingness to participate before interviews commenced. For this secondary analysis, the original informed consent obtained during data collection explicitly allowed for analysis and publication of deidentified data for research purposes. The University of Texas at Austin Institutional Review Board reviewed the scope of the original consent and confirmed that it was sufficient to permit the secondary analysis reported in this manuscript without requiring additional consent from participants.

Multiple measures were implemented to protect participant privacy and maintain confidentiality of research data. All interview data were deidentified, with participants assigned numeric codes used in analysis files. Any potentially identifying information mentioned during interviews was redacted. During interviews, participants were explicitly instructed not to share identifying information about patients, caregivers, or colleagues. All audio recordings and analysis files are stored on password-protected, encrypted servers maintained by Box (Box, Inc), with access limited exclusively to research team members who completed human subject research training and signed confidentiality agreements. Following verification of data accuracy and quality, all audio recordings were permanently deleted from Zoom's cloud storage and from local research team devices. Only deidentified data and analysis files are retained for the duration of the research project and the required institutional retention period.

Data sharing between Gryt Health (original data collector) and the University of Texas at Austin research team is governed by the nondisclosure agreement mentioned above, specifying appropriate data use, establishing data security requirements, limiting who may access the data, and prohibiting reidentification efforts or sharing of data with unauthorized parties. In reporting results, illustrative quotations from provider interviews are presented without attribution to specific participants to prevent potential identification based on distinctive speech patterns, roles, or institutional characteristics. Given the relatively small sample size (n=12) and the specialized nature of providers who work extensively with African American cancer populations, additional care was taken to avoid presenting combinations of demographic characteristics that might enable deductive identification.

XSL•FO
RenderX

Participants did not receive financial compensation for their participation in this study. The decision not to provide compensation was based on several considerations. First, participation involved health care professionals sharing insights from their professional practice, which is commonly considered part of professional development and service to the field rather than an activity requiring payment. Second, the interview duration (45-60 minutes) was relatively brief and scheduled at participants' convenience to minimize disruption to their work schedules. Third, many participants expressed intrinsic motivation to contribute to research aimed at improving support for underserved caregiver populations, indicating that the opportunity to shape future interventions served as a meaningful nonfinancial incentive. Finally, the absence of financial compensation helped ensure that participants' decision to participate was based on genuine interest in the research topic rather than external inducement, which may enhance the authenticity and thoughtfulness of responses.

Results

Participant Characteristics

Participants included physicians (7/12, 58.3%), social workers (2/12, 16.7%), nurses (2/12, 16.7%), and other providers (1/12, 8.3%). More than half of providers (7/12, 58.3%) identified as Black or African American, with additional representation from Latinx and Hispanic (3/12, 25%) and White (2/12, 16.7%) providers. Most participants (10/12, 83.3%) had more than 15 years of experience in their field. Providers practiced across 7 US states: California, Florida, Georgia, North Carolina, Ohio, Pennsylvania, and Tennessee. Detailed participant characteristics are presented in Table 1.

Table 1. Participant characteristics (N=12).

Category and subcategory	Values, n (%)
Provider type	
Social worker	2 (16.7)
Nurse	2 (16.7)
Provider or Physician	7 (58.3)
Other	1 (8.3)
Race and ethnicity	
White	2 (16.7)
African American or Black	7 (58.3)
Latinx or Hispanic	3 (25)
Gender identity	
Male	5 (41.7)
Female	7 (58.3)
Years of experience, (years)	
0-5	0 (0)
6-10	1 (8.3)
11-15	1 (8.3)
16-20	5 (41.7)
21-25	4 (33.3)
26-30	1 (8.3)
Geographic location	
California	2 (16.7)
Florida	1 (8.3)
Georgia	2 (16.7)
North Carolina	3 (25)
Ohio	1 (8.3)
Pennsylvania	2 (16.7)
Tennessee	1 (8.3)

XSL•FO
RenderX

Evaluation Outcomes

Thematic analysis revealed two overarching categories of findings: (1) unmet needs and barriers faced by African American cancer caregivers, and (2) opportunities and recommendations for improving caregiver support. The first category, unmet needs and barriers, comprises 3 subthemes practical barriers (transportation, financial constraints, housing insecurity, and competing obligations); social-emotional barriers (stress, guilt, burnout, fear, and unexpressed emotions); and cultural barriers (medical mistrust, superhero expectations, faith or spirituality tensions, and stigma around mental health and medications).

The second category, opportunities and recommendations, comprises four subthemes representing providers' suggestions for transformational change: (1) acknowledgment and formal compensation of caregiving work, (2) integration of caregivers into multidisciplinary care teams, (3) recognition and leveraging of cultural assets, and (4) strengthening the provider's role as a hub for change. These themes and subthemes are discussed below with supporting quotations from provider interviews.

The first theme of unmet needs and barriers covers 3 major categories practical, social-emotional, and cultural. Practical needs included lack of transportation, lack of food, lack of childcare, lack of available sick time for the caregiver, housing insecurity, and competing family obligations of partners and children. Additionally, several providers noted that African American cancer caregivers they have encountered cannot afford to sustain the caregiving role over the long term, as they lose income from the time they are performing their caregiving roles. Social-emotional challenges discussed included providers observing feelings of guilt or shame for not being able to be there for the family member as much as they would like. Also, providers discussed the range of unique emotions of African American cancer caregivers that they have observed, including anger, fear, grief, and sadness; both the ones that they express and other emotions that are not outwardly expressed but still noticed.

The third and most influential theme discussed relates to the cultural barriers unique to African American cancer caregivers. Providers frequently observed overwhelming stress among African American cancer caregivers. This stress stemmed both from the demands of caregiving and from previous negative experiences with the health care system, which contributed to medical mistrust and reluctance to participate in care decisions. Providers also discussed the cultural expectations of the "superhero Black woman" who takes on multiple roles in the family beyond cancer caregiving. Several providers discussed unspoken norms in African American communities that cancer caregiving is a family responsibility and that responsibility is often expected to occur in the home where there is limited support. Providers also discussed the role of faith and spirituality with African American communities, which is often complex when trying to provide care in a manner that is also responsive to a commonly held view of putting care "in God's hands." Finally, providers discussed cultural stigma at length, focusing on how African American communities can view caregiving, medications, and emotional labor. These thematic categories with illustrative quotes are listed in Table 2.

The second thematic category, opportunities and recommendations, comprises four subthemes representing providers' suggestions for transformational change and centers around (1) acknowledgment and compensation of the caregiver role, (2) incorporation of the caregiver into an expanded caregiving team, (3) reliance on the use of cultural assets as protective factors, and (4) the active role of the provider in providing support and facilitating systems change. Themes and selected illustrative quotes are listed in Table 3 and then discussed below in the context of their social-ecological placement to contextualize the findings as nested opportunities for change.

XSL•FO
RenderX

Table 2. Thematic barriers of African American cancer caregivers.

Theme	Example	Illustrative Quote
Practical barriers	Money, paid family sick leave, transportation, food, housing, competing obligations	• "If I call out means I don't get paid. Who's gonna watch the kids? How long is this gonna take? Is this a half day of work? Is this a full day of work? You see them doin' the calculations and you're understandin' because people need to take care of themself as well." • "It's essential they need to work to pay their bills, to survive. You understand the struggle." • "They understand that it's important. They want to be there and part of their care and their treatment plans, but they also have to do what they need to do to survive." • "Housing and then followed by transportation, followed by food, those seem to be the big three, but in that, I think transportation just gets the most, because that's the one we see quickly when someone doesn't show up for their appointment. I think underlying that, there seems to be a larger challenge with housing." • "That's a great need for a lot of, um, our minority patients, you know, African American patients, where you're having to take care of different generations. So it's not just that you have maybe a spouse or you have a parent, but you also have younger generations for whom you're caring for, and so you are torn." • "The wife was not only responsible for caring for her husband who was in the terminal process, but she had grandchildren in the home who she was caring for. She was trying to use the income that she did have to support adult children in the home." • "They've been tryin' to pay these bills. They've been tryin' to find daycare. They've been tryin' to find transportation. They're at the point where they have to go to work. They have to pick up a extra shift. They have to work two jobs. Now they can't be that caregiver because they have to do what they need to do for themselves."
Social emotional barriers	Stress, guilt, burnout, anxiety, fear, sadness, anger	• "They wanna protect their loved one and that they're not trusting. They may also pour that on to their loved one, the patient that have those reservations and fear. I think medical mistrust is just a huge blanket of an issue that should be addressed across African American community as a whole." • "That lead caregiver, begins to just take the rings and starts making really hard concrete decisions much to a loss to themselves because they forget, or they have to shelve their own emotion so that they take care of everyone else." • "That guilt and that anger, they're angry. Like, hey, I've put so many things on the backburner. I've sacrificed. That thought of still losing their family member even all that they've sacrificed is heartbreaking. They do get angry." • "A lot of the times they're not vocal. They kinda just deal with it. They're like, oh, I'm just used to doin' what we have to do. A lotta the times when they meet their breakin' point when they can't come anymore, they can't drop 'em off, they can't do this. Lots of times it's we're at that point where it's just so much that it's a lot too late where we don't have the grants or resources to help them through because they've been tryin' to get past it." • "There's definitely this duty, and I'm not quite sure if there's a lot of coping. The duty to get more information, the duty to be there, even if it means hurting themselves and not caring for themselves."

XSL•FO
RenderX

Theme	Example	Illustrative Quote
Cultural barriers	Mistrust or distrust, stigma and discrimination in health care, limited reflective providers, historical trauma, feels of exclusion or isolation, norms of caregiving expectations, cultural stigma of medications, stigma or taboo of mental health concerns or emotional needs, feeling of inferiority	• "This superhero Black woman thing is incredibly burdensome on caregivers because not only does she have to be the one for her father—and she loves it. She loves her father and will do it, But I'm sure she's the one who is the pillar of her kids and her husband, and at work she's taking on many roles. I think people put all these things on themselves." • "In the African American community, um, you know, family support is such a great thing. Um, and sometimes when we think about those patients that are hospitalized and have to go home or they need to go to, like, a rehab, sometimes they may not want their loved one to go to rehab or a skilled nursing facility. They would rather have them home. But when they're home, um, there may not be available equipment." • "In the African-American…the thought of going on hospice means that they are giving up, and they are not doing everything they can to live." • "Typically, in my experience in African American families, faith is very strong. When you're dealing with trauma or turmoil, the common notion is to put it in God's hands. God's got this, I put it in God's hands, and if I claim this disease, that means I'm going against the word of God." • "We have to look at the historical trauma and the mistrust that's developed over many, many, many years because of the disparities in care, the health inequities, the poor treatment that African Americans have and consistently still do experience. We're overcoming maybe a situation that we never had anything to do with, but we still have to recognize and honor it in the place that it is." • "When it comes to things like medications, there can be a stigma against pain medication, for example, I've had scenarios where there maybe was substance abuse in the family so there's this avoidance of, we're not gonna take anything but an aspirin for stage four metastatic cancer to the bones." • "My experience in the Black community is that it's uncommon or not a comfort space to admit that you're feeling vulnerable or to admit that you're struggling with anxiety or depression or fear, sadness, grief, any of those things because an acknowledgment of that can be seen as either weak or there's a lot of stigma associated with mental health and mental health illness. I think what that does is leave some really glaring gaps for our caregivers." • "My experience with African Americans looks different, from my experience with African Americans in the South. Some of what, like I brought up earlier, I think some of that might be colored by systemic racism that's pervasive in regions, that may not have anything to do with cultural challenges." • "What I've seen in the past is that, particularly among minorities and African Americans, they would come with very limited expectations, so whatever the doctor said was fine. They didn't ask questions among all of my other cultural components of patients. They didn't ask as many questions, because they assumed whatever was given was enough." • "If we think about mental health, and traditionally that's been a very taboo thing to address in the African American community. It's been very taboo for us to even admit, 'Hey, I need help. I think something's wrong.' It's not something that traditionally African Americans seek services for." • "'Just pray about your cancer, God don't hear you have your cancer.' Sometimes there's a disconnect with a person being seen as spiritually weak if they are seeking man for clinical services."

XSL•FO
RenderX

Table 3. Opportunities and aspirations for African American cancer caregivers.

Theme	Level of intervention	Quotes
Theme 1: acknowledgment of African American cancer caregiving as formal work with adequate compensation and support	Macro (structural)	• "We don't really think of what the caregivers go through and talk about the grief and everything. They always focus on the patients, all the emotions they go through, but the caregivers go through so much as well as far as all the emotions. Understand when they're angry it's valid because they don't get to voice that a lot of the times. We have so many mental health resources for our patients, but we don't really have a lot as far as the caregivers when it comes to oncology clinics." • "African Americans, I think more than any other community, because of the historical mistrust, should allow us first to start with being able to embrace their voices. Everybody's voice, they've gotta be able to own it." • "We've got social workers, we've got psychotherapists, we've got caregiver support groups, we've got art classes, we have tai chi, we have healing touch, we have all of what people call the soft and fuzzy stuff, but it's open to not just the individual with the diagnosis, but the caregiver too, knowing that the caregiver is also feeling the pain and the suffering, and is also walking through the journey and also carrying the extra weight of being in charge of the care, that's a herculean task." • "They're trying to hold the entire ocean in their hands. Just acknowledging, man, this is a lot. You're going through so many things, you're doing and pointing out the things that are going really well like you handled that beautifully. I don't know that I could have done so well."
Theme 2: acknowledgment and integration of African American cancer caregiving in a core multidisciplinary care team alongside providers and navigators.	Macro (structural)	• "I believe that it's absolutely, um, important to engage caregivers, for them to be part of the care team, so that this is not just, um, the providers and the patient, um, but allowing caregivers to be part of decision-making." • "The social worker will say, this caregiver is burned out, this caregiver needs this, and I'm always so happy because that's usually the perspective that the social worker can find quickly. I would say, I don't know that providers always have the best perspective on that, but I think having another person on the team see that, is probably even more useful." • "You know, caregivers expect that not only are we going to be, um, a team that provides the best care available to our patients, um, but also availing all the resources to make that successful." • "I think everyone needs to be listened to very closely, the provider and the patient. I think that caregivers are that link that translate the patient to us. They add information that we may not get." • "I think it's very important with just educating caregivers as well that you're an integral part of this journey as well. I want to empower them to ask questions, to question things because they need to understand what's goin' on with the patient as well so that they can help better take care of them." • "Yeah, it's important for—I think my way of doctoring is one that is very much—I'm gonna use this word—it's collaborative, in the sense that—or inclusive, or it is important that we're on the same team, that we have the same—that we're working together toward this."

XSL•FO
RenderX

Theme	Level of intervention	Quotes
Theme 3: recognition of African American diverse cultural and social assets as key outcomes to success.	Meso (community)	• "Sometimes, the best supportive care is providing hospice, and a lotta times, in the African-American community, they (don't) wanna do that either. They feel that they either have very good networks in their church, their religious communities that may be able to provide things." • "I think the first thing is trying to develop trust and trying to show that you care about them as a person, them as a believer in their faith and in what they actually feel is strong in their own persona and trying to not appear to be trying to coerce them or push them into one thing or the other and allow them as many options as possible." • "I think fortunately in the African-American community, usually that bond of family is very strong, and I think that helps." • "Churches in the African-American community are oftentimes a very central player in the lives of patients and the caregivers, and a huge source of strength and the support of being able to go to a church as well as knowing that everyone's praying for them, prayer circles, different things like that, that I feel are almost more specific to African Americans than some of our other groups." • "You only survive if there's a family network. There's no housing. There's no health care. There's no services. But if you have a family that is tight-knit that takes care of the elderly and the young, they take care of the addicted and the not addicted, people do well." • "If the pastor says to do it, people do it. Again, oftentimes women are very strong in their households. The matriarch of the family, if she says do it if grand-mama says do it, everybody's gonna do it. It's understanding the dynamics of what the relationships are in those families but resources like the faith nurse ministry, wonderful resource, because they're often embedded within their own community and providing not just health education, but the empathy side of things as well." • "Someone comes and someone else is there, and if the patient doesn't ask, a caregiver is asking, and that doesn't have to be a traditional caregiver, and I like that." • "They're bringing a niece, they're bringing a neighbor, they're bringing somebody a part of the faith that was big in the south, a minister or someone from church. I think that that's huge, and I think that's where the shift has been. I think where we are making strides." • "There's a large church just a few miles from my home that offers health ministry. People and their families can come to those and get some support and have a community of people who understand what they're going through."
Theme 4: solidifying the role of the provider as the hub for both individual level and transformational systems level change to better meet African American cancer caregiving needs.	Micro, Meso, Macro (individual, community, structural)	

XSL•FO
RenderX

Theme	Level of intervention	Quotes
		• "I play connect the dots every day, and sometimes I collect dots of names and organizations and resources because I may not need 'em today, but at some point, I'm gonna encounter someone that needs me to connect that next dot. I'm a professional dot collector, also known as a junk drawer because I keep all those things knowing that someone at some point may reach out and I wanna be able to help them to weave their way through the tsunami that's not only a health diagnosis but the system itself." • "She was so focused on the assessment papers that she had as her task and her duty and her job that she was missing that golden opportunity to just sit alongside this family and embrace where they are and listen to what they needed to share and express. I think there's so much we can learn from people's stories if we as providers leave our agenda behind." • "I realized in my case that thinking about, particularly when African Americans and Latinos, that the adherence to treatments depends upon us showing the commitment to the patient…I learned that if I give my cellphone to my patients, even though they don't call me, they show up to clinic. If I don't, they don't show up to clinic. I was able to increase adherence to the treatment just by giving my cellphone." • "I think that we need to make ourselves available to caregivers to make their job easier. We need to make more telemedicine available. We have to make ourselves more available. We have to answer questions more readily. We have to have that, but we also probably need to reach out and see what they need. I don't think we're asking enough." • "I feel like there's not a lot offered for caregivers in the clinical setting that I'm working in." • "I feel like we have a brief intervention during the clinic visit where we may turn attention to a caregiver's concern, but just acknowledging and really supporting that caregiver and having anything to offer them at all, I don't feel like that's something that I'm aware of or that I was trained to do during all of these years of medical practice and training. I don't feel like that's something that has ever been magnified or highlighted or focused on." • "I think that we need to make ourselves available to caregivers to make their job easier. We need to make more telemedicine available. We have to make ourselves more available. We have to answer questions more readily. We have to have that, but we also probably need to reach out and see what they need. I don't think we're asking enough." • "In the same way that we have thermometers of stress in patients, we should have thermometers of stress on caregivers. That would be very easy to deliver. It could be an after-visit survey that you can give out, then the institution can then appoint services to that." • "I think also understanding even the caregivers may feel overwhelmed, and they may not feel like they have the right to ask so they themselves may not be straightforward. I think that we go so fast that we just—okay. It's all a monologue, 'I'm telling you X, X, X, X. Okay. Bye.' Next. It's really a conversation. This was lovely. Do you need anything else in the future, or if you wanted to talk to caregivers, I could find some."

Theme 1. Acknowledgment of Cancer Caregiving as Formal Work With Adequate Compensation and Support (Macro)

All providers interviewed acknowledged cancer caregiving as work. Providers discussed the need for formal acknowledgment and compensation for African American cancer caregivers, with changes being made at a structural or governmental level. Providers discussed that caregivers need to be compensated at appropriate levels or minimum "living wages." Providers stressed that this formal employment should include benefits that acknowledge workplace hazards and stress and provide resources to support social and emotional health. This formal acknowledgment, compensation, and support are particularly salient for African American cancer caregivers who face unique practical, social-emotional, and cultural barriers to cancer caregiving.

Theme 2. Acknowledgment and Integration of African American Cancer Caregivers in a Core Multidisciplinary Care Team Alongside Providers and Navigators (Macro)

Providers overwhelmingly discussed the need to integrate caregivers as part of the cancer care team, along with the primary care provider, navigator, and social worker. This care team should be collectively trained, oriented, and prepared to meet the needs of the patient. African American cancer caregivers should play an essential and equal role in the care team as they serve as "transformational interpreters" for the patient with cancer. This newly realized care team should include open communication and mutual respect to transform the health outcomes of patients. The thematic analysis highlights

XSL•FO
RenderX

key attributes that would be necessary for the provider role, including having providers that are reflective of the community with shared lived experiences. Therefore, initiatives that seek to diversify providers are necessary. Many interviewed participants also highlighted the essential role of the navigator, especially when serving African American communities. Navigators can provide value to the patient and facilitate better outcomes for the entire team. This defined care team has the potential to be especially effective, as data show that supported African American caregivers show less stress in their caregiving role than other races and ethnicities and are well equipped to respond as part of the care team.

Theme 3. Recognition of African American Diverse Cultural and Social Assets as Key Outcomes to Success (Meso)

Participants noted that African American cancer caregivers have unique assets within their families, communities, and culture that can drive transformational change. Participants noted the diverse support systems and assets among African American communities, such as the cultural norm of taking pride in the caregiving roles and valuing a strong family support network. Participants also observed the African American community having a unique advantage of having a collectivist culture where community caregiving is the norm and often performed by a variety of people, including partners, children, friends, neighbors, pastors, and faith leaders. Religion and spirituality serve a unique and influential role in having the potential to positively affect the emotional health, well-being, and health outcomes of both the patient and the caregiver.

Theme 4. Solidifying the Role of the Provider as the Hub for Both Individual-Level and Transformational Systems–Level Change to Better Meet African American Cancer Caregivers' Needs (Micro, Macro)

Participants who were interviewed also acknowledged the contributing and transformative role that providers could and should play in supporting African American cancer caregivers and their networks. Participants noted that providers are intentionally situated as an intermediary between individual (micro) and organizational and managerial (macro) influence to make these suggested changes to better care for caregivers. Due to this level of influence, participants emphasized that it is essential for providers to be trustworthy, humble, sincere, authentic, caring, and empathetic throughout this process. The provider serves as a central hub for "resource capture" and ensures the functioning of the system at all levels. Participants noted that to be successful, providers must listen, support, and communicate effectively. A key attribute to this success is being available both in appointments and by providing cues of being open and available. Participants shared personal successes of sharing emails or phone numbers that signaled availability and care to caregivers, showing, "I am here. I care. You can reach me."

Discussion

Principal Findings

This study reveals how providers identify and support the needs of African American cancer caregivers, highlighting both the unique challenges these caregivers face and opportunities for improvement. Within health practice settings, this research presents an opportunity for care providers to work with African American cancer caregivers to meet their needs and improve the health outcomes for both themselves and those they care for.

Findings from this study highlight issues faced by African American cancer caregivers, as reported by health care professionals experienced working with these populations. The insights offered cover practical and social-emotional topics as well as culturally specific challenges unique to the African American cancer caregiving experience. The data offered by providers highlight the increasing need for the acknowledgment and support for the caregiving role, especially among African American communities. Further, the findings support structural interventions, including the establishment of the caregiving role as one that is respected, compensated, and formally included in the care team system.

Providers interviewed for this study identified multiple cultural assets unique to African American communities that they observed serving as protective factors for both caregivers and patients. As detailed in Theme 3 and Table 3, providers described witnessing the pride and dedication African American caregivers demonstrated in their caregiving roles, the centrality of family networks in mobilizing support, the influential role of religion and faith communities in providing emotional support and practical assistance, the collectivist orientation where caregiving was shared among extended family and community members rather than falling to one individual, and the strength derived from these interconnected support systems.

These observations represent providers' interpretations of cultural strengths they witnessed in their clinical practices and offer important insights into community resources that health care systems might better recognize and leverage in supporting African American caregivers, particularly drawing from extensive quotations tied to Theme 3. Finally, providers identified their influential and important role that they need to take to support African American cancer caregivers and transform systems to better meet needs. Health care providers have the opportunity to work directly with caregivers to ensure their needs are met and their status is elevated in the caregiving role, though doing so would reasonably need additional administrative and organizational support.

Implications for Innovation in Participatory Cancer Care

This research offers important implications for innovation in cancer care delivery, particularly regarding participatory care models, technology implementation, and health equity initiatives.

XSL•FO
RenderX

Advancing Participatory Care Models

Our findings provide a roadmap for operationalizing participatory cancer care. Provider recommendations for formally integrating caregivers into multidisciplinary care teams (Theme 2) offer specific mechanisms defining caregivers as official team members with designated roles, including caregivers in care planning meetings, providing training and resources, and creating communication channels that position caregivers as partners. Health care organizations committed to patient- and family-centered care can use these provider-generated recommendations to redesign care delivery models and team structures.

The emphasis on formal acknowledgment and compensation of caregiving work (Theme 1) has implications for health care policy and payment models. As value-based care increasingly recognizes the importance of care coordination and outcomes beyond acute treatment, there are opportunities to reimagine how caregiving contributions are valued and supported. Provider recommendations point toward policy innovations such as caregiver stipends, paid family leave expansions, and coverage for family caregiver training and support.

Informing Technology and Digital Health Innovation

Provider insights have significant implications for designing digital health tools responsive to African American caregiver needs. Providers emphasized accessibility, cultural responsiveness, and addressing practical barriers such as time constraints and competing obligations. These insights can inform the development of telehealth platforms designed with African American caregiver needs in mind (addressing medical mistrust through culturally concordant providers, flexible scheduling, and technical support); mHealth apps addressing practical needs while providing emotional support and culturally relevant resources; patient and caregiver portals that recognize caregivers as partners with appropriate information access and care coordination tools; and remote monitoring systems that engage both patients and caregivers sustainably.

Critically, providers highlighted the importance of recognizing and leveraging cultural assets such as faith communities, extended family networks, and collectivist values (Theme 3). Digital innovations that facilitate rather than replace these existing support systems—for example, by enabling coordination among multiple family caregivers or connecting caregivers with faith-based resources—may be more culturally acceptable and effective.

Reducing Disparities Through Systems Innovation

Providers' emphasis on systems-level change (Theme 4) points toward structural innovations, including workforce diversification initiatives to increase representation of underrepresented minority providers and navigators, institutional policies requiring caregiver assessment and support planning with specific attention to barriers affecting communities of color, care navigator programs specifically designed to address barriers faced by African American caregivers, provider training incorporating cultural humility and asset-based approaches, and quality metrics including caregiver support as a measured component of cancer care quality.

Translation to Practice and Policy

The multilevel nature of our findings provides several entry points for intervention. Health care organizations can begin with feasible changes such as provider training and caregiver assessment tools while advocating for policy changes like caregiver compensation and employment protections. Importantly, our asset-based findings (Theme 3) offer a counter-narrative to deficit-focused approaches, providing a foundation for interventions that leverage community strengths while addressing systems-level barriers, an important shift toward equity-oriented innovation.

Limitations

Limitations of this research include the fact that it employs a purposive sample of 12 participants. This sample was established a priori due to the limited number of qualifying providers in the United States. The research represents the lived experiences of a group of providers reporting interactions with African American cancer caregivers and may not be fully reflective of the caregiving experience. Further, this data does not represent the lived experiences of African American cancer caregivers but provides some external insights into the health care experiences of this population and how participants in systems interact with these caregivers. Our purposive sampling strategy, while methodologically appropriate for exploratory qualitative research [40,42], introduces potential selection bias that limits generalizability. Providers who responded to recruitment outreach and agreed to participate may differ systematically from nonresponders in important ways, including greater awareness of or commitment to health equity issues, more positive experiences working with African American caregivers, higher comfort discussing issues of race and health care disparities, or more flexible schedules allowing research participation. Additionally, recruitment via LinkedIn profiles and institutional websites may have skewed toward more digitally engaged or academically oriented providers working at larger institutions. These selection effects mean our findings may not fully represent the perspectives of providers who work in smaller community practices, have less experience or comfort working with African American populations, hold views less aligned with patient-centered care models, or face greater time constraints that prevented participation. Readers should interpret findings as representing the perspectives of a motivated, experienced subset of providers rather than all providers caring for African American patients and caregivers with cancer. Despite these limitations, the rich, detailed insights from experienced providers provide valuable direction for intervention development even if not necessarily generalizable. Significant accessibility limitations must also be noted. Recruitment and data collection were completed online. Potential participants without a significant online presence or less familiar with Zoom may not have participated. These limitations may result in a younger, more technologically savvy, and more urban sample. Following best practices of qualitative research, the interview guide was flexible but not substantially adapted during the iterative process of data analysis. Many culturally sensitive and stigmatized topics were discussed in the research. Participants may not have fully shared experiences related to prejudices or medical experiences due to distrust or social desirability.

XSL•FO
RenderX

Comparison With Prior Work

Our findings confirm and extend prior research on African American caregiver experiences while offering novel insights from the provider perspective. Previous literature documented that African American caregivers are disproportionately burdened as sole caregivers, often receiving no external support [7]. Our provider interviews corroborated these disparities, with participants observing the "superhero Black woman" phenomenon, where caregivers assume multiple family roles beyond cancer care, often without adequate support systems.

Additionally, our provider observations aligned with prior research identifying that only a few African American caregivers rate their health as excellent or very good [7]. Our findings here note significant social-emotional barriers, including stress, guilt, and burnout among African American caregivers (Table 2).

Existing literature highlighted that African American caregivers face greater financial strain, with 74% modifying employment status due to caregiving demands [19]. Our provider participants confirmed these practical barriers, emphasizing transportation, housing, and competing family obligations as primary challenges. Importantly, providers observed how financial pressures create impossible choices: "They have to pick up an extra shift...Now they can't be that caregiver because they have to do what they need to do for themselves."

Previous studies noted limited provider diversity, with only 3% of oncologists being Black [22], potentially limiting racial concordance benefits. Our findings suggest this deficit profoundly impacts care delivery, with providers emphasizing the need for workforce diversification as essential for building trust and improving communication.

Novel Contributions

While prior research documented caregiver challenges through direct caregiver accounts, our study uniquely reveals provider perspectives on these same issues, offering novel insights. First, we identified cultural assets that providers observe as protective factors, including strong family networks, faith communities, and collectivist caregiving approaches, each of which has been underexplored in deficit-focused caregiver literature. Second, our findings reveal specific recommendations from providers for systems-level change, including formal caregiver compensation and integration into care teams, which existing literature has not addressed. Third, we document how medical mistrust, previously identified as a barrier [15,16], manifests in provider-caregiver interactions and affects care delivery from the provider perspective.

These provider insights complement existing caregiver-focused research by identifying actionable intervention points within health care systems and revealing potential misalignments between caregiver experiences and provider perceptions that could inform targeted improvements in care delivery. Our study also extends the understanding of caregiving by revealing that providers often witness caregivers suppressing emotions ("a lot of the times they're not vocal, they kinda just deal with it"), suggesting the health impacts may be even more severe than self-reported measures indicate.

Conclusions

This research lays a foundation for understanding and responding to the needs unique to African American cancer caregivers. It provides key data from providers to support the need for health care transformation and better meet the needs of African American cancer caregivers. With increases in the number of cancer caregivers within the African American population, it is essential to understand mechanisms of support in a culturally responsive manner.

Caregiver support is always and everywhere a systems phenomenon; individual interventions cannot address structural inequities that require institutional transformation. Data shows that this population faces significant health disparities coupled with numerous barriers within the physical, social, cultural, and political environments. This research can positively contribute to the long-term goals of addressing unmet needs and their practical barriers, supporting an assets-based approach to meeting the needs of this population through transformational systems change.

Also, this research will contribute to a better understanding of how providers can specifically help to meet the unique needs of the African American cancer caregiver population. Caregivers cannot be peripheral to care teams when they are central to patient outcomes, as providers note in this project. Future work should seek to provide evidence-informed guidelines for the implementation of these suggested strategies in addition to adding the perspectives of caregivers themselves so more of the care team is included in the dialogue. Full implementation of these strategies will require substantial organizational commitment and resources, which may pose challenges.

Additional research is needed that (1) provides an evidence-informed guide to organizational transformation to better meet the needs of African American cancer caregivers, (2) explores the readiness and resources of health care organizations and their staff to implement such changes, (3) explores the readiness of African American cancer caregivers to be more formally acknowledged and included in the caregiving process, and (4) evaluates the effectiveness of the suggested transformation.

Acknowledgments

All providers interviewed acknowledged cancer caregiving as work. Providers discussed the need for formal acknowledgment and compensation for African American cancer caregivers, with changes being made at a structural or governmental level. Providers discussed that caregivers need to be compensated at appropriate levels or minimum "living wages." Providers stressed that this formal employment should include benefits that acknowledge workplace hazards and stress and provide resources to

XSL•FO
RenderX

support social and emotional health. This formal acknowledgment, compensation, and support are particularly salient for African American cancer caregivers who face unique practical, social-emotional, and cultural barriers to cancer caregiving.

Funding

The data used in this research comes from a project funded by EMD Serono in a contract with Gryt Health as part of the initial stages of planning the creation of culturally responsive support programs and resources for African American caregivers. The authors received no specific funding for this secondary data analysis.

Data Availability

The data generated in this study are available upon request from the corresponding author. Interested researchers may contact the corresponding author to discuss access arrangements that maintain ethical obligations to participants while supporting research transparency. Interview transcripts contain potentially identifying information about participants and the health care settings in which they work, so any arrangement would need to respect participant anonymity, also recognizing the need for the depth of the qualitative data.

Authors' Contributions

Author contributions are reported according to the Contributor Roles Taxonomy (CRediT).
Conceptualization: BL (lead), AS (equal)
Data curation: AS (lead), GD (equal)
Formal analysis: GD (lead), BL (equal), SU (supporting)
Funding acquisition: AS (lead), BL (supporting)
Investigation: BL (lead), AS (equal)
Methodology: BL (lead), AS (equal)
Project administration: AS (lead), BL (equal), GD (supporting)
Resources: GD (lead), AS (supporting)
Supervision: BL (lead)
Validation: SU (lead), BL (supporting)
Writing – original draft: BL (lead), GD (equal), AS (supporting), SU (supporting)
Writing – review & editing:BL (lead), GD (supporting), AS (supporting), SU (supporting)

Conflicts of Interest

BL is a co-founder at Gryt Health, where he is a member of the governing board and serves as chief research officer. AS was senior vice president for outcomes and impact at Gryt Health at the time of this work. Both have equity shares in Gryt Health.

Multimedia Appendix 1
Interview guide.
[DOCX File , 12 KB - cancer_v12i1e76266_app1.docx]

References

1. Life expectancy increases, however suicides up in 2022. Centers for Disease Control and Prevention. 2023. URL: https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2023/20231129.htm [accessed 2025-11-27]
2. Skolnik R. Global Health 101. Burlington, MA: Jones & Bartlett Learning; 2015.
3. Singer PA, Taylor AD, Daar AS, Upshur REG, Singh JA, Lavery JV. Grand challenges in global health: the ethical, social and cultural program. PLoS Med 2007;4(9):e265 [FREE Full text] [doi: 10.1371/journal.pmed.0040265] [Medline: 17850175]
4. Chronic conditions and caregiving. Johns Hopkins Medicine. URL: https://www.hopkinsmedicine.org/health/wellness-and-prevention/chronic-disease-caregiver-resources [accessed 2025-11-27]
5. Marín-Maicas P, Corchón S, Ambrosio L, Portillo MC. Living with Long Term Conditions from the Perspective of Family Caregivers. A Scoping Review and Narrative Synthesis. Int J Environ Res Public Health 2021 Jul 08;18(14):7294 [FREE Full text] [doi: 10.3390/ijerph18147294] [Medline: 34299745]
6. Abazari A, Chatterjee S, Moniruzzaman M. Understanding cancer caregiving and predicting burden: an analytics and machine learning approach. AMIA Annu Symp Proc 2023;2023:243-252 [FREE Full text] [Medline: 38222371]
7. Tan KR, Waters AR, Chen Q, Hendricks BA, Coombs LA, Kent EE. Inequities Among Cancer Caregivers with Diverse Identities: A Review of the Literature and Future Directions. Curr Oncol Rep 2023 Jul;25(7):803-812 [FREE Full text] [doi: 10.1007/s11912-023-01415-3] [Medline: 37043117]

XSL•FO
RenderX

8. Bom J, Bakx P, Schut F, van Doorslaer E. The Impact of Informal Caregiving for Older Adults on the Health of Various Types of Caregivers: A Systematic Review. Gerontologist 2019 Sep 17;59(5):e629-e642 [FREE Full text] [doi: 10.1093/geront/gny137] [Medline: 30395200]
9. Liu C, Badana A, Burgdorf J, Fabius C, Roth D, Haley W. Systematic Review and Meta-Analysis of Racial and Ethnic Differences in Dementia Caregivers' Well-Being. Gerontologist 2021 Jul 13;61(5):e228-e243 [FREE Full text] [doi: 10.1093/geront/gnaa028] [Medline: 32271380]
10. Kim H, Mahmood A, Kedia S, Ogunsanmi DO, Sharma S, Wyant DK. Impact of Residential Segregation on Healthcare Utilization and Perceived Quality of Care Among Informal Caregivers in the United States. J Racial Ethn Health Disparities 2025 Jun;12(3):1887-1899. [doi: 10.1007/s40615-024-02018-9] [Medline: 38758399]
11. Whitney RL, Bell JF, Kilaberia TR, Link BM, Choula RB, Reinhard SC, et al. Diverse demands and resources among racially/ethnically diverse caregivers. Ethn Health 2023 Aug;28(6):912-931 [FREE Full text] [doi: 10.1080/13557858.2023.2179022] [Medline: 36858966]
12. Navaie-Waliser M, Feldman PH, Gould DA, Levine C, Kuerbis AN, Donelan K. When the caregiver needs care: the plight of vulnerable caregivers. Am J Public Health 2002 Mar;92(3):409-413. [doi: 10.2105/ajph.92.3.409] [Medline: 11867321]
13. Cook SK, Snellings L, Cohen SA. Socioeconomic and demographic factors modify observed relationship between caregiving intensity and three dimensions of quality of life in informal adult children caregivers. Health Qual Life Outcomes 2018 Aug 29;16(1):169 [FREE Full text] [doi: 10.1186/s12955-018-0996-6] [Medline: 30157852]
14. Howe RJ, Bell JF, Agnoli A, Sullivan JL, Bidwell JT. Home alone or connected: Caregiver communication and training from health providers. J Am Geriatr Soc 2024 Feb;72(2):361-368. [doi: 10.1111/jgs.18686] [Medline: 38006285]
15. Schillinger D. Social Determinants, Health Literacy, and Disparities: Intersections and Controversies. Health Lit Res Pract 2021 Jul;5(3):e234-e243 [FREE Full text] [doi: 10.3928/24748307-20210712-01] [Medline: 34379549]
16. Vandecasteele R, Robijn L, Willems S, De Maesschalck S, Stevens PAJ. Barriers and facilitators to culturally sensitive care in general practice: a reflexive thematic analysis. BMC Prim Care 2024 Oct 23;25(1):381. [doi: 10.1186/s12875-024-02630-y] [Medline: 39443846]
17. Stiefel F, Bourquin C, Salmon P, Achtari Jeanneret L, Dauchy S, Ernstmann N, ESMO Guidelines Committee. Electronic address: clinicalguidelines@esmo.org. Communication and support of patients and caregivers in chronic cancer care: ESMO Clinical Practice Guideline. ESMO Open 2024 Jul;9(7):103496 [FREE Full text] [doi: 10.1016/j.esmoop.2024.103496] [Medline: 39089769]
18. Longacre ML, Miller MF, Fang CY. Racial and ethnic variations in caregiving-related physical, emotional, and financial strain during COVID-19 among those caring for adult cancer patients. Support Care Cancer 2021 Jul;29(7):4137-4146 [FREE Full text] [doi: 10.1007/s00520-020-05933-9] [Medline: 33404809]
19. Gazaway S, Odom JN, Herbey I, Armstrong M, Underwood F, Heard TV, et al. Cultural Values Influence on Rural Family Caregivers' Decision-Making for Ill Older Adult Loved Ones. J Pain Symptom Manage 2024 Jul;68(1):86-95. [doi: 10.1016/j.jpainsymman.2024.04.012] [Medline: 38641135]
20. Okai A, Elter B, Williams M. Understanding Diverse Caregiver Experiences and Their Impact on Caregiver Burden in Multiple Sclerosis. Int J MS Care 2025 Apr;27(Q2):125-133 [FREE Full text] [doi: 10.7224/1537-2073.2024-025] [Medline: 40304015]
21. Smith M, Kew C, Washington T, Bergeron C, Merianos A, Sherman L, et al. Caregiver strain among African American and Hispanic male caregivers with chronic conditions. Innov Aging 2022;6(Suppl 1):100. [doi: 10.1093/geroni/igac059.398]
22. Winkfield KM, Flowers CR, Patel JD, Rodriguez G, Robinson P, Agarwal A, et al. American Society of Clinical Oncology Strategic Plan for Increasing Racial and Ethnic Diversity in the Oncology Workforce. J Clin Oncol 2017 Aug 01;35(22):2576-2579. [doi: 10.1200/JCO.2017.73.1372] [Medline: 28459634]
23. Shen MJ, Peterson EB, Costas-Muñiz R, Hernandez MH, Jewell ST, Matsoukas K, et al. The Effects of Race and Racial Concordance on Patient-Physician Communication: A Systematic Review of the Literature. J Racial Ethn Health Disparities 2018 Feb;5(1):117-140 [FREE Full text] [doi: 10.1007/s40615-017-0350-4] [Medline: 28275996]
24. Lin OM, Reid HW, Fabbro RL, Johnson KS, Batch BC, Olsen MK, et al. Association of Provider Perspectives on Race and Racial Health Care Disparities with Patient Perceptions of Care and Health Outcomes. Health Equity 2021;5(1):466-475 [FREE Full text] [doi: 10.1089/heq.2021.0018] [Medline: 34316530]
25. Berry LL, Dalwadi SM, Jacobson JO. Supporting the Supporters: What Family Caregivers Need to Care for a Loved One With Cancer. J Oncol Pract 2017 Jan;13(1):35-41. [doi: 10.1200/JOP.2016.017913] [Medline: 27997304]
26. Senger A, Venetis MK, Greene K, Catona D, Devine KA. Healthcare provider assessments of caregiver communication behaviors during gynecologic Cancer treatment appointments. PEC Innov 2024 Dec;4:100259 [FREE Full text] [doi: 10.1016/j.pecinn.2024.100259] [Medline: 38347863]
27. Schatz A, Brooks-Coley K, Harrington E, Murray MS, Carlson RW. Patient, Caregiver, and Oncologist Experiences With and Perceptions of Racial Bias and Discrimination in Cancer Care Delivery. J Natl Compr Canc Netw 2022 Oct;20(10):1092-1098.e2. [doi: 10.6004/jnccn.2022.7051] [Medline: 36240848]
28. Griffin JM, Riffin C, Bangerter LR, Schaepe K, Havyer RD. Provider perspectives on integrating family caregivers into patient care encounters. Health Serv Res 2022 Aug;57(4):892-904 [FREE Full text] [doi: 10.1111/1475-6773.13932] [Medline: 34957543]

XSL•FO
RenderX

29. Litzelman K. Caregiver Well-being and the Quality of Cancer Care. Semin Oncol Nurs 2019 Aug;35(4):348-353 [FREE Full text] [doi: 10.1016/j.soncn.2019.06.006] [Medline: 31229346]
30. Riffin C, Wolff JL, Estill M, Prabhu S, Pillemer KA. Caregiver Needs Assessment in Primary Care: Views of Clinicians, Staff, Patients, and Caregivers. J Am Geriatr Soc 2020 Jun;68(6):1262-1270 [FREE Full text] [doi: 10.1111/jgs.16401] [Medline: 32167573]
31. Fenton ATHR, Ornstein KA, Dilworth-Anderson P, Keating NL, Kent EE, Litzelman K, et al. Racial and ethnic disparities in cancer caregiver burden and potential sociocultural mediators. Support Care Cancer 2022 Nov;30(11):9625-9633 [FREE Full text] [doi: 10.1007/s00520-022-07367-x] [Medline: 36190556]
32. McKillip J. Need Analysis: A Guide to Assessing Needs for Health and Social Programs. Beverly Hills, CA: Sage Publications; 1987.
33. Creswell JW, Poth CN. Qualitative Inquiry And Research Design: Choosing Among Five Approaches. Thousand Oaks, CA: Sage Publications; 2016.
34. Pope C, Mays N. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. BMJ 1995 Jul 01;311(6996):42-45 [FREE Full text] [doi: 10.1136/bmj.311.6996.42] [Medline: 7613329]
35. Patton MQ. Qualitative Evaluation And Research Methods. Thousand Oaks, CA: Sage Publication; 2015.
36. Hennink M, Kaiser B. Conducting Qualitative Research: A Practical Guide. London: Sage Publications; 2022.
37. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies: Guided by Information Power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
38. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15(4):351-377. [doi: 10.1177/109019818801500401] [Medline: 3068205]
39. Stokols D. Translating social ecological theory into guidelines for community health promotion. Am J Health Promot 1996;10(4):282-298 [FREE Full text] [doi: 10.4278/0890-1171-10.4.282] [Medline: 10159709]
40. Service R. Organizational Research Methods. Thousand Oaks, CA: Sage Publications; Sep 05, 2008:614-617.
41. Lincoln YS, Guba EG. Naturalistic Inquiry. Beverly Hills, CA: Sage Publications; 1985.
42. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]

Abbreviations

mHealth: mobile health

Edited by N Cahill; submitted 19.Apr.2025; peer-reviewed by TK Torres, K Kleszynski; comments to author 11.Sep.2025; accepted 25.Nov.2025; published 08.Jan.2026.

Please cite as:
Love B, Dermid G, Upshaw S, Stark A
Health Care Providers' Perceptions of Unmet Needs Among African American Cancer Caregivers: Qualitative Investigation Among US Medical Professionals
JMIR Cancer 2026;12:e76266
URL: https://cancer.jmir.org/2026/1/e76266
doi:10.2196/76266
PMID:

©Brad Love, Gerold Dermid, Sean Upshaw, Amy Stark. Originally published in JMIR Cancer (https://cancer.jmir.org), 08.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Scalp Cooling as a Biopsychosocial Intervention in Patients Receiving Highly Alopecia-Inducing Chemotherapy: Prospective Single-Arm Study

Dörthe Schaffrin-Nabe[1], Dr med; Anke Josten-Nabe[1], Dr rer nat; Adrian Heinze[2,3], MSc; Andrea Tannapfel[4], Prof Dr Med; Merle Schaffrin[1], Dr. med; Rudolf Voigtmann[1], Prof Dr Med

[1]Praxis für Hämatologie und Onkologie, Universitätsstr. 110 e, Bochum, Germany

[2]Independent Research Consultant, Hamburg, NRW, Germany

[3]University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[4]Bürkle-de-la-Camp-Platz 1, Pathologie Ruhr-Universität Bochum, 44789, NRW, Germany

Corresponding Author:
Dörthe Schaffrin-Nabe, Dr med
Praxis für Hämatologie und Onkologie, Universitätsstr. 110 e, Bochum, Germany

Abstract

Background: Chemotherapy-induced alopecia is among the most psychologically distressing adverse effects of systemic cancer therapy. Although scalp cooling is increasingly used to mitigate hair loss, it is still largely perceived as a cosmetic intervention. Its broader psychological relevance and the biological basis of treatment success, particularly the preservation of follicular integrity under ongoing cytotoxic exposure, remain insufficiently explored.

Objective: This study aimed to reconceptualize scalp cooling beyond visible hair preservation by examining its psychological impact on patients receiving highly alopecia-inducing chemotherapy, while integrating quantitative objective hair preservation metrics with structural and ultrastructural analyses of hair follicle damage to identify avenues for improving follicular integrity and scalp-cooling efficiency.

Methods: A total of 82 patients undergoing highly alopecia-inducing chemotherapy consisting of a sequential anthracycline-taxane regimen (4 cycles of epirubicin and cyclophosphamide followed by 12 weekly paclitaxel applications) received standardized scalp cooling. Objective hair preservation was quantified using the hair mass index (HMI) as a standardized and reproducible measure of hair retention. Structural and ultrastructural follicular integrity was assessed using light microscopy as well as scanning and transmission electron microscopy. Objective hair preservation metrics were analyzed in relation to patient-reported quality-of-life outcomes (EORTC [European Organisation for Research and Treatment of Cancer]–based measures), subjective treatment burden, and cognitive appraisal of the scalp-cooling experience. Multivariable regression models were applied to identify determinants of posttherapeutic quality of life.

Results: Visible chemotherapy-induced alopecia was successfully prevented in more than half of the treated patients. Scalp cooling resulted in substantial objective hair preservation as quantified by the HMI. However, HMI values showed only a limited association with posttherapeutic quality-of-life outcomes. In contrast, the cognitive appraisal of scalp cooling emerged as a central determinant of posttherapeutic quality of life, independent of the degree of objective hair retention. Structural and ultrastructural analyses demonstrated that the preservation of follicular integrity was closely associated with successful macroscopic hair retention under ongoing cytotoxic exposure, supporting a biological basis for the clinical effectiveness of scalp cooling.

Conclusions: The clinical relevance of scalp cooling extends beyond objective and visible hair preservation and appears to reside predominantly in its psychological impact on patients undergoing highly alopecia-inducing chemotherapy. Importantly, the identification of structural and ultrastructural markers of follicular vulnerability provides a mechanistic foundation for the future optimization of scalp-cooling approaches and for the development of adjunct follicle-directed protective strategies to enhance follicular integrity and support patient well-being during cytotoxic therapy.

(JMIR Cancer 2026;12:e94635) doi:10.2196/94635

KEYWORDS

scalp cooling; chemotherapy-induced alopecia; hair follicle biology; ultrastructural analysis; quality of life; cognitive appraisal; supportive oncology; patient-reported outcomes

XSL•FO
RenderX

Introduction

Background

Chemotherapy-induced alopecia (CIA) remains one of the most visible and psychologically distressing adverse effects of systemic cancer therapy. Beyond a transient cosmetic change, hair loss functions as a public marker of disease, disrupts personal identity, and continuously reinforces the illness experience for patients and their families [1,2]. Despite substantial advances in oncologic treatment, the psychosocial burden of CIA remains inadequately addressed.

Scalp cooling is currently the only US Food and Drug Administration–approved intervention shown to mitigate CIA and has gained broad clinical acceptance. However, its efficacy is highly heterogeneous: while some patients achieve near-complete hair preservation, others experience pronounced alopecia despite identical chemotherapy regimens and standardized cooling protocols [3,4]. This interindividual variability represents a major limitation of current approaches and indicates that key determinants of treatment response remain insufficiently understood.

Although direct clinical evidence remains limited, experimental human hair-follicle models and scalp-cooling outcome studies support the biological plausibility that intrinsic follicular characteristics, including hair-shaft integrity and pigmentary-unit stability, may contribute to interindividual variation in chemotherapy-induced damage and scalp-cooling response [5-10]. At the same time, the subjective experience of hair loss and individual coping strategies substantially shape quality of life, highlighting CIA as an interface between biological vulnerability and psychosocial resilience.

In this context, this study examines scalp-cooling efficacy in a homogeneously treated cohort receiving highly alopecia-inducing chemotherapy. By integrating objective follicular parameters with patient-reported outcomes (PROs), this work aims to contribute to a more nuanced, biopsychosocial understanding of hair preservation and to enable a more individualized approach to supportive care strategies [6,11].

Study Population

The study enrolled 82 patients with early-stage breast cancer, all of whom received curative-intent chemotherapy. A uniform treatment regimen was applied, consisting of epirubicin and cyclophosphamide followed by weekly paclitaxel. This standardized approach ensured comparable cytobiological effects on hair follicles across the cohort and provided a consistent basis for evaluating scalp-cooling efficacy. Given the predictable and uniform severity of alopecia associated with this regimen, a control group without scalp cooling was not included.

Scalp cooling was performed with a sensor-controlled device designed to maintain a consistent scalp temperature below 22 °C. By ensuring reproducibility in both chemotherapy and cooling parameters, the study sought to isolate patient-specific and follicle-specific determinants of hair preservation. Patients were excluded from scalp cooling in cases of clinically manifest scalp metastases, generalized hematologic malignancies such as leukemia, scalp injuries, or cryoglobulinemia.

Objectives

The primary objective of the study was to evaluate the efficacy of scalp cooling for hair preservation during chemotherapy, using the hair mass index (HMI) as an objective measure obtained through cross-sectional trichometry.

Three primary efficacy end points were defined:

1. Posttreatment HMI as a measure of residual hair mass,
2. Avoidance of visible alopecia, defined as HMI ≥50,
3. Absolute hair loss, calculated as the difference between posttreatment and pretreatment HMI (ΔHMI).

Secondary objectives included the characterization of treatment-associated trichological and ultrastructural hair alterations, the exploratory assessment of biological and clinical factors associated with hair preservation, and the evaluation of patient-reported treatment experiences and the psychosocial burden related to hair loss and scalp cooling. Trichological and biological secondary outcomes included pretreatment and posttreatment assessments of hair density, anagen rate, and hair shaft and bulb diameter, as well as ultrastructural hair-shaft surface alterations. Particular emphasis was placed on melanosomal features, with transmission electron microscopy (TEM) used to quantify melanosome morphology and density.

Potential predictors and influencing factors of hair preservation were evaluated across multiple domains, including general health parameters (eg, age, comorbidities, alopecia-inducing co-medications, nicotine use, and menopausal status), individual hair characteristics and practices (eg, hair color, length, structure, and hair treatment), and treatment-related biological vulnerability markers. These included hematologic toxicity during chemotherapy (eg, neutrophil nadirs) as well as laboratory parameters reflecting liver function, renal function, and thyroid status. In female patients, endocrine profiling (follicle-stimulating hormone, luteinizing hormone, and estradiol) was explored to contextualize gonadal function.

PROs were collected to assess the subjective experience of scalp cooling, including perceived burden and treatment acceptance. Psychosocial impact was evaluated before and after the intervention using selected items derived from the EORTC QLQ-C30 (European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire), complemented by study-specific questions to capture treatment-related psychosocial effects associated with hair loss and the cooling procedure.

Methods

Overview

To determine hair density, measurements were carried out at the frontal vertex before and after treatment using the Cross Section Trichometer (Cohen HMI device) [12]. This tool enables standardized quantification of hair mass within a 4-cm² scalp area. The procedure was designed to deliver consistent, reproducible results at single time points, assuming the examiner had received appropriate training. Only hairs longer than 2.5 cm were considered. Outcomes were expressed as the HMI, defined as the surface area of hair in mm² per cm² of scalp,

XSL•FO
RenderX

multiplied by 100. Values typically fall between 75 and 100; a score ≥50 was predefined as the threshold for the absence of clinically visible alopecia consistent with established HMI methodology, indicating that visible thinning typically emerges below this level [12], while values below 20 correspond to advanced alopecia [12]. For analysis, around 50 follicles were extracted from the selected 4 cm^2 scalp region and positioned between 2 glass slides with the roots exposed, sampling for reproducible trichogram assessment. At the same time, the number of extracted follicles was intentionally kept as low as possible to minimize additional hair removal in a psychologically vulnerable patient population, with the explicit aim of hair preservation [12].

Light microscopy was applied to capture images of untreated hair samples at magnifications of 10× and 40× using a Leica DM3000 (Leica Microsystems) microscope. Measurements included bulb diameter (typical range: 150 - 200 µm) and shaft thickness (normal for the frontal vertex: 75 - 85 µm) [13]. The growth phase of the follicles was assessed based on the characteristic bulb morphology: according to established trichogram and hair cycle morphology criteria, anagen follicles showed normal bulb dimensions, while catagen and telogen follicles presented with smaller bulbs and club-shaped ends. Structural abnormalities such as shaft irregularities, breakage, and deformation were interpreted as signs of follicular dystrophy [14].

To study melanosome distribution, TEM was applied to ultrathin hair shaft sections obtained from 52 participants before and after treatment (Figure 1). These cases represented the first consecutively available samples for which electron microscopic processing was feasible. The subset was not selected according to clinical outcomes, degree of alopecia, treatment response, or any other biological characteristics. Rather, the number of evaluable cases was determined by the substantial technical complexity, limited pathology resources, and the labor-intensive nature of ultrastructural tissue preparation, image acquisition, and quantitative analysis. Consequently, no intentional enrichment for responders, nonresponders, or specific morphological patterns was performed, minimizing the risk of selection bias. Two types of organelles were identified: pheomelanosomes, typically smaller and more rounded, and eumelanosomes, which tend to be elongated and fibrillar in form. The quantitative assessment showed that blond hair contained relatively few melanosomes (about 9 per 100 $µm^2$), while markedly higher densities were observed in darker hair, averaging around 66 per 100 $µm^2$ in brown and 70 per 100 $µm^2$ in black samples [15]. For preparation, follicles were fixed in glutaraldehyde, dehydrated through successive steps, and embedded in Epon. Sections approximately 50-nm thick were cut, stained with uranyl acetate and lead citrate, and examined at magnifications ranging between 1500× and 15,000× (JEOL 2100).

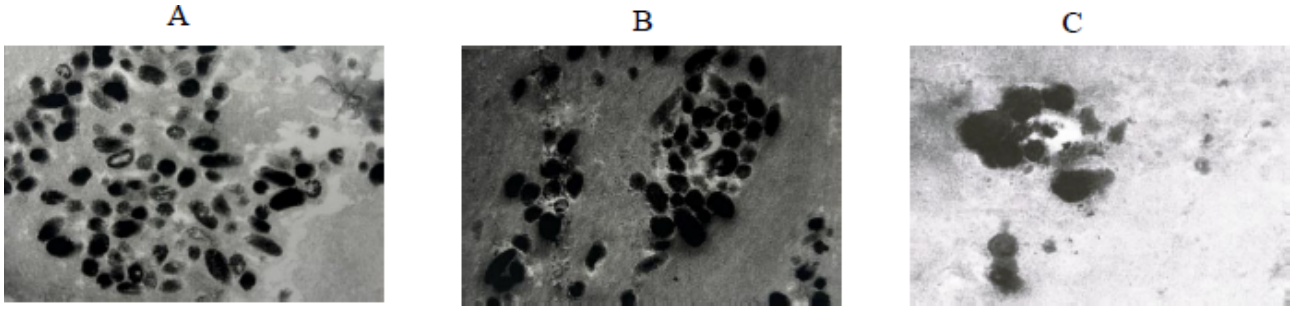
Scanning electron microscopy (SEM) was applied to evaluate surface alterations of the hair shaft in untreated and treated samples. Examinations were performed with a Zeiss Gemini 982 (ZEISS Microscopy) field emission SEM at magnifications ranging from 200× to 2000×. Prior to imaging, the hair was sputter-coated with a thin gold layer. The extent of structural impairment was categorized using a system comparable to the Kim scale, which is commonly used for classifying extrinsically induced hair damage. For illustration, representative patient samples showing different levels of hair surface deterioration are displayed in Figure 2 [16].

The Kim scale (adapted) is as follows:

- Grade 0: intact, untreated hair with a uniform cuticle layer.
- Grade 1: slight irregularities of the cuticle, including shallow dents or uneven overlap, without cracks or loss of material.
- Grade 2: pronounced lifting of cuticle layers, with fissures or small perforations, but the cortex is still covered.
- Grade 3: partial exposure of the underlying cortex.
- Grade 4: complete loss of the cuticle layer, leaving the cortex unprotected.

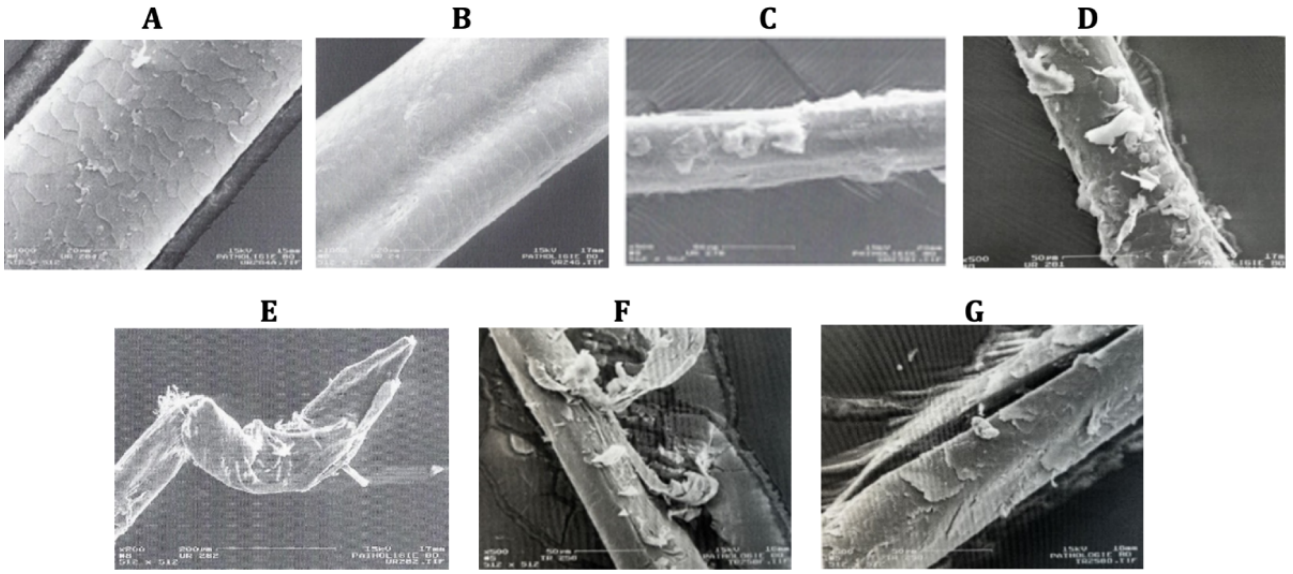
Sections of hair shafts from 52 participants were examined to confirm consistency in the detected alterations. The analysis concentrated on grading intrinsic surface damage from 0 to 4 in relation to the therapeutic procedure.

Figure 1. Exemplary transmission electron microscopy photos. (A) Patient with brown dense hair (pre-HMI=100). Pretherapeutic pigmentary status: 89 total melanosomes, 44 pheomelanosomes, 45 eumelanosomes, and ratio of eumelanosomes to pheomelanosomes=1.02. Scale: approximately 8000:1. (B) Same patient with no visible hair loss (post-HMI=58). Posttherapeutic pigmentary status: 75 total melanosomes, 31 pheomelanosomes, 44 eumelanosomes, and ratio of eumelanosomes to pheomelanosomes=1.42. Scale: approximately 8000:1. (C) Patient with bad hair retention beginning of visible alopecia (post-HMI=35). Pigmentary status: disordered melanogenesis characterized by clumping of melanin. Scale: approximately 12,000:1. HMI: hair mass index.



XSL•FO
RenderX

Figure 2. Different patterns and grades of hair shaft damage. (A) Normal roof tile scale pattern; (B) Dents, grade 1; (C) Deposits on hair shaft, grade 1; (D) Lift-ups of cuticle, grade 2; (E) Distortion and crack of hair shaft, grade 3; (F) Severe lift-ups with partial exposure of cortex; (G) Extensive exposure of cortex with dents, deposits, and cracks, grade 4.



Patient-Centered Outcomes (Self-Report Measures)

Patient-centered outcomes were assessed using self-report measures obtained at baseline and after the intervention. The subjective appraisal of treatment and the overall quality of life were rated on Likert-type scales ranging from 1 (very poor) to 7 (very good). Treatment-related burden and discomfort were assessed on a 4-point scale from 1 (none) to 4 (very strong).

Quality-of-life domains were evaluated using selected items from the EORTC QLQ-C30, complemented by additional study-specific questions addressing treatment experience and procedure-related aspects.

Patient-reported questionnaires were used to capture patients' subjective treatment experiences, perceived burden, and overall well-being. Psychosocial data were collected prospectively from the study's initiation onward as part of the observational study protocol. Consequently, analyses examining associations between cognitive appraisal, hair preservation, and quality of life were conducted using prospectively collected data.

Statistical Procedures

Descriptive statistics (median, IQR, and percentage) were used to summarize patient and treatment characteristics, as well as hair retention outcomes. Multiple linear and binary-logistic regression models were conducted as exploratory analyses to assess associations with hair retention and quality of life. Because some variables included in the models were measured after treatment, the models were not intended to strictly establish causal or temporal effects. Baseline variables were interpreted as potential predictors, whereas posttreatment parameters were interpreted as correlates of the corresponding outcome. The joint approach was used to examine these associations while simultaneously adjusting for baseline characteristics and posttreatment morphological parameters.

All key model assumptions were assessed and met for all models, including linearity, normality, and homoscedasticity of residuals, absence of influential outliers, and acceptable multicollinearity. For the 3 outcome variables of hair retention (post-HMI, HMI≥50, and HMI difference), clinically relevant and biologically plausible correlates were identified within thematic domains and entered into multivariable models if they were at least borderline significant (P=.006). Only standardized regression coefficients are reported, and the intercept was omitted from the models' tables because it is not interpretable in standardized form. Finally, we used Spearman correlation coefficients to test associations between melanosomes and damage degree because of their ordinal scale and potential outliers. Statistical significance was defined as P=.003.

Ethical Considerations

This study was approved by the Research Ethics Committee (2016 - 011-f-S Ethikkommission Westfalen-Lippe, Münster, Germany). All research data were deidentified and coded with unique identifiers, and the data were transmitted using secure encrypted protocols and stored on encrypted disks on secure servers. The study included only patients who had agreed to the use of their medical information by signing the respective consent form. No compensation was provided to participants for survey completion. This paper does not include any images or information that could identify individual participants.

Results

Among 82 patients, 80 (97.6%) were women, and the median age was 56.5 (IQR 49 - 65) years. The baseline median HMI score was 75 (IQR 58 - 92) and declined after treatment to a median of 51 (IQR 17 - 58). At follow-up, 52% (n=43) of participants still had an HMI score of 50 or higher, corresponding to the absence of visible hair loss. For patients with initially low hair density, maintaining the HMI may be of particular clinical importance, highlighting the role of relative HMI preservation as a meaningful end point. The median decline in the HMI score from baseline to posttreatment was −27 (IQR −45 to −10). In those starting with an HMI below 50 (n=12),

XSL•FO
RenderX

the median reduction amounted to −19 (IQR −27 to −5). Comprehensive baseline characteristics, treatment-related factors, and posttreatment clinical parameters together with variables derived from light microscopy, TEM, and SEM are summarized in Table S1 in Multimedia Appendix 1. Complete ultrastructural data were available for a quality-controlled subset of 52 patients.

While longitudinal changes in HMI provided a robust quantitative measure of treatment-associated alterations in scalp hair density, this global metric does not fully capture the underlying biological mechanisms at the follicular level. To contextualize these numerical changes, we also examined the microstructural integrity of individual hair follicles using complementary morphological approaches. These microscopic assessments capture complementary aspects of follicular and hair shaft integrity, encompassing surface morphology—particularly ultrastructural damage features—as well as pigmentation patterns, which cannot be adequately reflected by global quantitative indices alone.

In order to comprehensively characterize hair retention outcomes, 3 complementary HMI–based efficiency parameters were evaluated: (1) posttreatment HMI as a measure of residual scalp hair density, (2) the proportion of patients maintaining an HMI ≥50, corresponding to the absence of visible hair loss, and (3) the absolute change in HMI from baseline, capturing individual longitudinal dynamics and relative hair preservation. Together, these parameters allow for a differentiated assessment of treatment-associated effects on scalp hair beyond absolute posttreatment values.

The initial modeling steps are summarized in Tables S2-S5 in Multimedia Appendix 1, showing multivariable regression analyses thematically grouped by general clinical characteristics, as well as light, TEM, and SEM variables to identify independent factors associated with each of the 3 HMI efficiency parameters.

These supplementary tables provide the structural and biological context for the subsequent efficiency-focused analyses and are presented separately to maintain clarity and readability of the main results. Only variables meeting the predefined relevance threshold of $P=.08$ were retained in the final multivariable models for posttreatment HMI, HMI ≥50, and absolute HMI change and are therefore presented in Tables 1-3.

Table . Final exploratory linear regression for associations with posttreatment hair mass index[a].

Variables	β (95% CI)	*P* value
Comorbidity	−.28 (−0.58 to 0.02)	.07
Neutrophil	.17 (−0.12 to 0.46)	.25
Anagen post	.3 (−0.01 to 0.61)	.06
Pheomelanosome difference	.17 (−0.12 to 0.46)	.25
Damage difference	−.29 (−0.58 to -0.01)	.046

[a]Model fit was adjusted R^2=0.32.

Table . Final exploratory binary-logistic regression for associations with hair mass index ≥50[a].

Influencing variables	Odds ratio (95% CI)	*P* value
Comorbidity	0.05 (0.01 - 0.75)	.05
Neutrophil	2.46 (0.60 - 25.69)	.32
Hair shaft pre	1.12 (1.04 - 1.27)	.02
Bulb size pre	0.98 (0.95 - 1.01)	.26
Bulb size post	1.02 (0.98 - 1.06)	.39
Damage difference	0.28 (0.06 - 1.06)	.08

[a]Model fit was Nagelkerke R^2=0.65.

Table . Final exploratory linear regression for associations with the hair mass index difference[a].

Variables	β (95% CI)	*P* value
Age	.31 (0.05 to 0.58)	.02
Medication	−.35 (−0.61 to −0.09)	.009
Bulb size pre	.05 (−0.21 to 0.31)	.69
Pheomelanosome difference	.34 (0.09 to 0.59)	.009
Eumelanosome difference	.27 (0.01 to 0.53)	.04

[a]Model fit was adjusted R^2=0.33.

XSL•FO
RenderX

To address the temporal structure of the data, baseline variables and posttreatment variables were conceptually distinguished. Baseline variables were analyzed as potential predictors of treatment outcomes, whereas posttreatment variables were included to characterize biological features associated with treatment responses. Accordingly, the models including posttreatment variables should be interpreted as exploratory analyses of associations rather than predictive models in a strict temporal sense.

When analyzing factors associated with posttreatment HMI (Table 1), the degree of hair shaft damage difference showed a statistically significant negative association (β=−.29, *P*=.046), indicating that greater increases in hair shaft damage were linked to lower posttreatment HMI values. Comorbidity exhibited a borderline negative association (β=−.28, *P*=.066), suggesting a tendency toward reduced hair retention in patients with comorbid conditions. Posttreatment anagen frequency showed a borderline positive association (β=.30, *P*=.06), indicating a trend toward improved posttreatment HMI with higher anagen proportions.

For HMI ≥50 (Table 2), the baseline hair shaft diameter was significantly associated with visible hair retention (odds ratio [OR] 1.12, 95% CI 1.04-1.27; *P*=.02), indicating that a larger baseline hair shaft diameter was linked to higher odds of maintaining visible hair. Comorbidity showed a borderline negative correlation (OR 0.05, 95% CI 0.01-0.75; *P*=.05), suggesting a tendency toward a reduced likelihood of HMI-defined hair preservation (HMI≥50). A greater increase in hair shaft damage was also related to lower odds of preventing visible hair loss, although this did not reach statistical significance (OR 0.28, 95% CI 0.06-1.06; *P*=.08). No relevant associations were observed for neutrophil count or bulb size parameters. Overall, model performance was very good (Nagelkerke R^2=0.65).

Finally, in the linear model for HMI difference (Table 3), age (β=.31, *P*=.02), pheomelanosome difference (β=.34, *P*=.009), and eumelanosome difference (β=.27, *P*=.04) were significantly associated with the HMI change, indicating that older age and lower degrees of melanosome loss were linked to a smaller decline in the HMI. Medication use showed a significant inverse relationship to greater HMI reduction among patients receiving alopecia-inducing medication (β=−.35, *P*=.009). Overall, the model fit was moderate (adjusted R^2=0.33).

Given the exploratory design and the number of tested influencing factors across models, no formal adjustment for multiple testing was applied; the results should therefore be interpreted as exploratory. CIs were wide for some estimates, reflecting the limited precision associated with the sample size.

In addition to biological and treatment-related variables, psychosocial parameters were specifically examined for their association with hair retention outcomes. This analysis, summarized in Table 4, addresses the psychological dimension of hair preservation and its contribution to the overall treatment experience and quality of life. Finally, in a linear regression model using posttherapeutic quality of life as the dependent variable, we entered pretreatment quality of life, subjective appraisal of scalp cooling, emotional burden, and HMI difference as candidate variables. QoL pre (β=.41, *P*<.001) and scalp-cooling appraisal (β=.53, *P*<.001) showed significant positive associations with moderate-to-strong effect sizes, indicating that higher baseline QoL and a more favorable appraisal of scalp-cooling outcomes were linked to better QoL after therapy. Age, burden, and HMI change were not independently associated with posttreatment quality of life. Overall, the model fit was high (adjusted R^2=0.48).

Table . Final exploratory linear regression for associations with posttreatment quality of life (QoL)[a].

Variables	β (95% CI)	*P* value
Age	.04 (−0.17 to 0.26)	.69
QoL pre	.41 (0.21 to 0.62)	<.001
Appraisal	.53 (0.27 to 0.79)	<.001
Burden	.05 (−0.19 to 0.29)	.67
HMI[b] difference	−.01 (−0.24 to 0.23)	.97

[a]Model fit was adjusted R^2=0.48.

[b]HMI: hair mass index.

Discussion

Principal Findings

This study provides an integrated analysis of scalp cooling that extends beyond the traditional dichotomy of hair preservation versus alopecia. In a cohort exposed to highly alopecia-inducing chemotherapy, visible hair loss was avoided in 53% (n=43) of patients, representing a clinically meaningful level of effectiveness that exceeds purely cosmetic expectations. Importantly, this efficacy signal was consistently accompanied by distinct structural and ultrastructural features of follicular integrity, indicating that treatment response is biologically modulated rather than binary. These findings further suggest that the effectiveness of scalp cooling is not solely determined by temperature reduction per se but reflects a more complex biological response at the level of follicular resilience. While these observations are biologically plausible, they should be interpreted as exploratory. This study was not designed to establish causal mechanistic relationships but rather to identify clinically relevant associations. In particular, relative melanosome stability and hair shaft surface integrity appear to

XSL•FO
RenderX

represent a coherent biomarker constellation of follicular health, supporting the concept that scalp cooling modulates intrinsic follicular vulnerability to chemotherapy-induced damage rather than acting as a purely physical cooling intervention. However, these structural features were assessed in a limited subset of samples using high-resolution ultrastructural techniques, reflecting their technically demanding nature and should therefore be interpreted with caution.

Previous large-scale registry data, such as the Dutch Scalp Cooling Registry, including more than 7400 patients, have consistently demonstrated that the type of chemotherapy regimen is a key determinant of scalp-cooling efficacy [17]. Consistent with these findings, we have similarly observed variability in scalp-cooling efficacy across different treatment settings in previous work [18]. In addition, we have reported variable success rates across different regimens in prior work. In contrast, this study applied a uniform chemotherapy protocol consisting of epirubicin and cyclophosphamide followed by weekly paclitaxel, a regimen invariably associated with complete alopecia in the absence of scalp cooling. This approach was deliberately chosen to minimize treatment-related heterogeneity and to allow for the assessment of scalp-cooling effects within a highly standardized clinical setting. However, the absence of a control group represents a methodological limitation, as comparative conclusions regarding absolute treatment effects cannot be drawn. At the same time, the use of a uniform chemotherapy regimen allowed for a more focused analysis of interindividual variability in follicular response under comparable treatment conditions. Furthermore, establishing an untreated control group was considered ethically challenging, given the expected near-complete alopecia associated with this regimen.

Although prior investigations have primarily focused on the chemotherapy regimen type and general patient-related factors, our analysis extends this perspective by systematically addressing structural and ultrastructural determinants of follicular preservation.

Our findings indicate that the relative preservation of melanosome abundance—particularly pheomelanosomes—together with a larger pretreatment hair shaft diameter and minimal posttherapeutic shaft surface damage, represents independent correlates of successful hair preservation.

Importantly, these observations should be interpreted as biologically plausible associations rather than strictly causal relationships. Nevertheless, they consistently point toward a central role of melanogenesis integrity and keratinocyte proliferative capacity in maintaining follicular resilience under cytotoxic stress. In this context, chemotherapy-induced impairment of melanogenesis and keratinocyte turnover appears to directly compromise follicular homeostasis. While scalp cooling may attenuate these effects, it does not fully prevent them, thereby highlighting both the biological potential and the inherent limitations of the intervention.

Collectively, these findings support the concept that scalp-cooling efficacy is not solely determined by physical temperature modulation, but rather reflects a biologically modulated preservation of follicular integrity.

From a pathophysiological perspective, anthracyclines and taxanes preferentially target rapidly proliferating keratinocytes within the hair matrix while simultaneously inducing the formation of reactive oxygen species (ROS). This dual mechanism not only disrupts keratinocyte turnover but also functionally impairs components of the follicular pigmentary unit, including melanocytes and associated sebocytic structures [19-21].

Even in the presence of scalp cooling, the residual exposure of follicular melanocytes to cytotoxic agents may induce a prematurely dysregulated state of melanogenesis. Morphologically, this may manifest as increased melanin aggregation, reduced melanosome density, and a compositional shift from pheomelanin toward eumelanin synthesis [19,22].

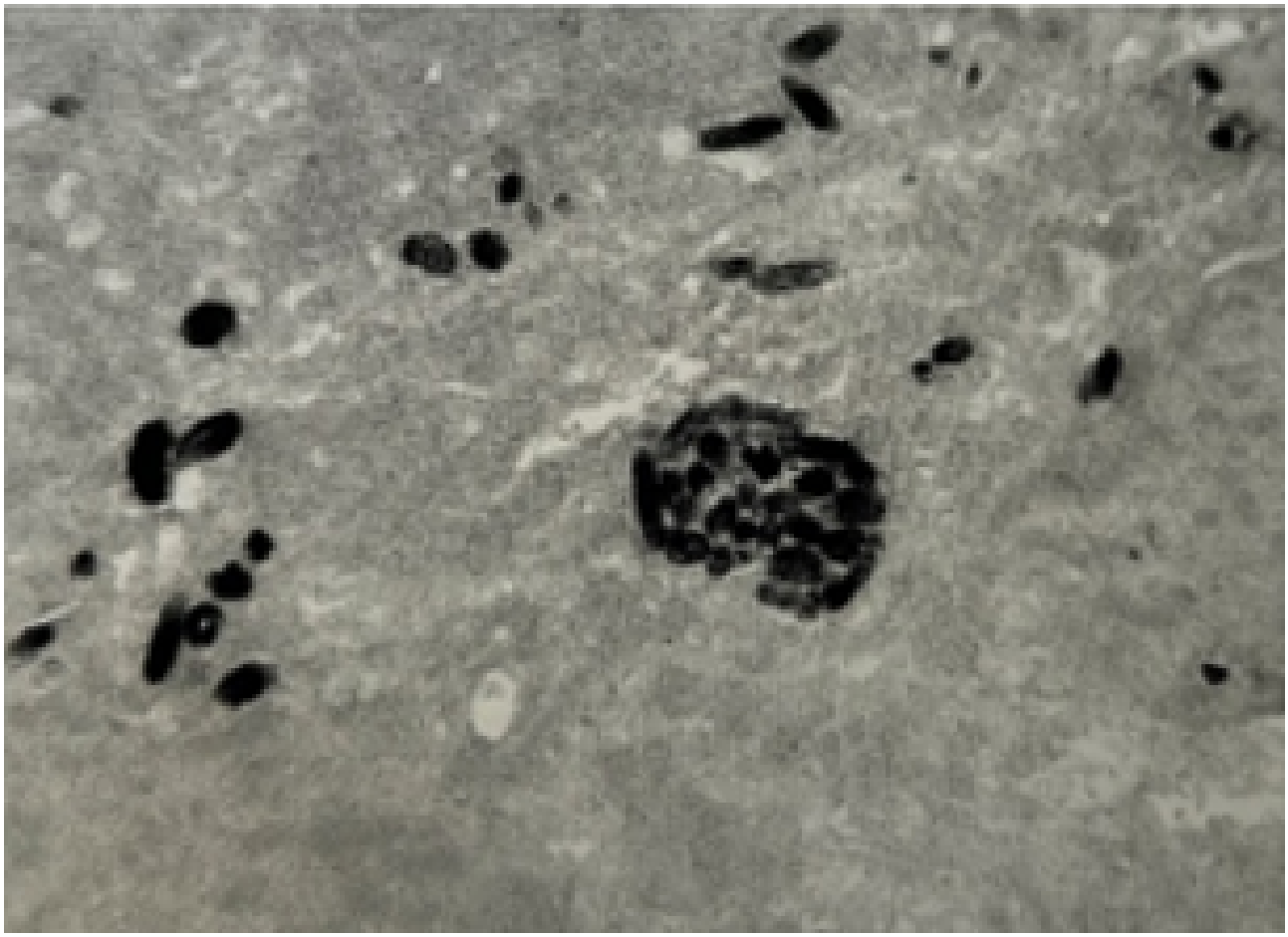
This shift is mechanistically plausible in the context of oxidative stress, as cysteine—a critical substrate for pheomelanin synthesis—is increasingly diverted toward glutathione production under conditions of elevated ROS. Consequently, the balance of melanogenic pathways is altered in favor of eumelanogenesis, thereby reflecting an adaptive but functionally disruptive antioxidative response [23-25].

These mechanisms provide a molecular basis for the observed reduction in melanosome density and the elevated eumelanin-to-pheomelanin ratio post treatment (Mdn_{pre}=0.34 to Mdn_{post}=0.52; Table S1 in Multimedia Appendix 1). Importantly, the role of melanosomes extends far beyond the mere determination of hair color. Their perinuclear distribution around keratinocyte nuclei confers both photoprotective and antioxidative benefits, safeguarding keratinocyte genomic stability and ensuring controlled proliferative activities that ultimately drive proper hair shaft development [26-28].

Posttherapeutically, occasional melanosome clustering (Figure 3) was observed. However, this feature did not show a significant association with hair retention, suggesting that clustering represents a secondary or epiphenomenal structural alteration rather than a determinant of follicular resilience.

XSL•FO
RenderX

Figure 3. Cluster (more than 5 touching pigments).



In contrast, preserved melanosome abundance and the relative stability of melanosomal organization emerged as a potential surrogate marker associated with more favorable hair preservation. Minor qualitative shifts in melanosome composition were observed but were not formally quantified and are therefore reported descriptively. These exploratory and hypothesis-generating observations may reflect lower treatment-associated cellular stress and could represent a morphological correlate of relative follicular resilience. Preservation of melanosomal architecture thus extends beyond cosmetic relevance and may contribute to the maintenance of hair shaft integrity. Melanocytes and keratinocytes exist within a tightly coordinated functional unit characterized by reciprocal paracrine signaling and direct cellular cross talk. Mediators such as stem cell factors, endothelin-1, and basic fibroblast growth factor contribute to maintaining melanogenic activity and broader follicular homeostasis. This coordinated interaction provides a biological framework, linking the preservation of melanosomal architecture with follicular stability, and may help explain why the maintenance of an intact melanogenic process is associated with improved hair preservation [26,29]. Although both melanosome stability and shaft-surface integrity contribute significantly to hair retention under scalp cooling, statistical modeling identified them as independent correlates. This phenomenon can be explained by their distinct biological roles: the relative preservation of melanosome density provides intracellular photoprotection and antioxidative buffering, whereas shaft-surface conditions primarily determine the mechanical resilience of the emergent hair fiber. Therefore, each factor contributes additively, but not redundantly, to follicular protection, highlighting the multifactorial nature of CIA [27,28,30].

Increased oxidative cellular stress resulting from residual exposure to cytostatic agents can compromise both the hair bulb and the hair shaft surface. In contrast to metabolically active cells, the extracutaneous compartments of the hair shaft lack sophisticated antioxidative defense mechanisms, including enzymatic and nonenzymatic radical scavenging systems. As a consequence, structural components of the hair shaft remain particularly vulnerable to oxidative degradation, thereby amplifying the impact of chemotherapy-induced injury [31-33].

ROS compromise the hair's protective lipid envelope, most notably the 18-methyleicosanoic acid monolayer, which is crucial for shaft cohesion. Loss of this barrier renders the cuticle highly susceptible to endogenous and environmental damage. Under chemotherapy, excessive ROS not only damage cuticular elements but also affect cortical structures, leading to fibrillary misalignment. Subsequent oxidative events, including keratin modification, disruption of disulfide cross-links, and alteration of amino acid residues, irreversibly deteriorate the cortical architecture, resulting in fragile, dull, and fracture-prone hair fibers [32,34-37].

From a clinical perspective, the significance of hair shaft surface damage cannot be overstated. SEM demonstrated that a higher

XSL•FO
RenderX

damage grade is associated with visible alopecia or lower HMI in general, and descriptive results appear to suggest that a damage grade of 3 (cortical exposure) may constitute a critical threshold for visible alopecia (Table S6 in Multimedia Appendix 1). This observation parallels earlier findings in cosmetic dermatology, where chemical cuticle disruption progressing to cortical exposure invariably results in irreversible hair fragility [16,37].

Within the follicular pigmentary unit, sebocytes act as integral partners to keratinocytes and melanocytes. Their secretory activity not only supports hair shaft emergence but also contributes to the antioxidative and immunomodulatory milieu of the follicle. Anthracycline-based chemotherapy may impair the sebocyte function, thereby reducing the protective capacity of sebum and increasing hair shaft vulnerability [14,19,38,39].

The lipid-rich sebum film reduces friction, minimizes breakage, and provides a hydrophobic barrier that limits transepidermal water loss, chemical irritation, and microbial penetration [40]. Acting as a lipid matrix, sebum enriched with antioxidants such as vitamin E and squalene may facilitate the scavenging of ROS and the absorption of ultraviolet radiation, indirectly reducing oxidative and inflammatory stress on keratinocytes and melanocytes [5,41-43].

After characterizing the impact of chemotherapy on the hair follicle and shaft quantitatively by hair loss assessed through the HMI, structurally by shaft alterations ranging from tapering detectable by light and SEM to ultrastructural surface damage, as well as indirectly by changes in melanosome density and patterning as surrogate markers of oxidative stress within the follicular pigmentary unit, these parameters may explain the individual variability in hair retention observed under scalp cooling.

However, further lowering of scalp-cooling temperatures has not improved efficacy in preclinical models when combined with highly cytotoxic regimens such as TAC (simultaneously administered docetaxel, doxorubicin, and cyclophosphamide), underscoring the need for additional protective or restorative strategies [44,45]. Consequently, recent efforts have shifted toward dermatological approaches designed to preserve the follicular pigmentary unit during chemotherapy under scalp cooling, thereby enhancing hair preservation.

Potential strategies to augment scalp-cooling outcomes include topical antioxidants (eg, melatonin and vitamin C encapsulated in nanovesicles) [46]. Another experimental concept involves transient pharmacological cell cycle arrest to reduce chemotherapy susceptibility of follicular keratinocytes. Preclinical studies have demonstrated that topical palbociclib induces temporary G1 arrest in follicular keratinocytes, protecting them from cytotoxic apoptosis while maintaining the subsequent anagen activity [20]. By reducing local perfusion and metabolic activity, scalp cooling limits follicular exposure to cytotoxic agents and lowers chemotherapy-induced oxidative stress. In this context, controlled oxidative stress may paradoxically induce p53-dependent cell cycle arrest rather than apoptosis, consistent with the protective effects of scalp cooling [47], providing a plausible, though not causally established, mechanistic explanation.

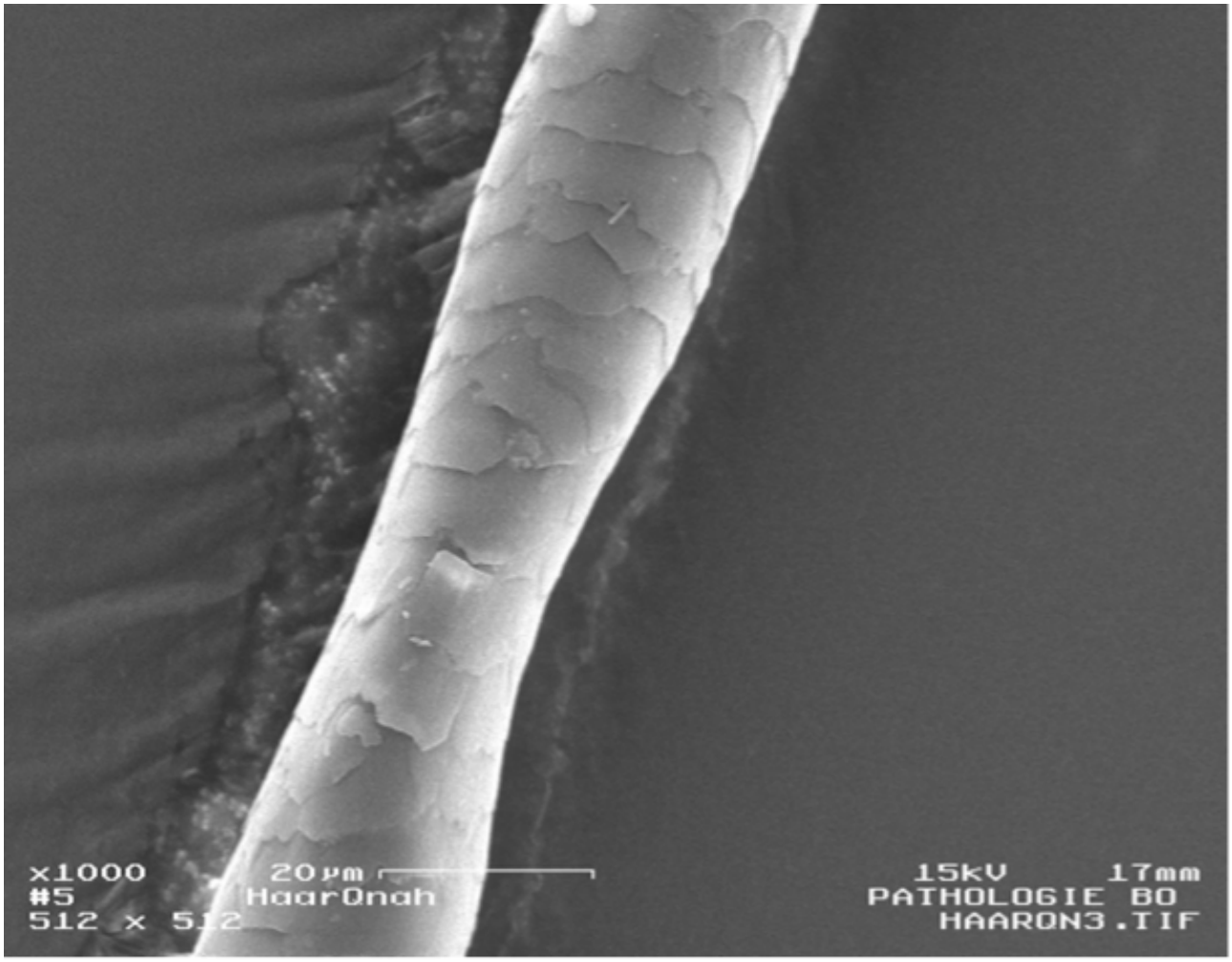
Another promising direction involves targeting the transforming growth factor-beta pathway. The chemotherapy-induced upregulation of transforming growth factor-beta 2 promotes catagen transition and apoptosis. The inhibition of this pathway by agents such as cartogenin has shown encouraging results in preclinical models, extending anagen duration and improving follicular resilience. While these findings are promising, clinical validation in human settings is currently lacking. Within this context, the integration of such pharmacological approaches represents a rational translational avenue that warrants further investigation [48,49].

Conceptually, ultrastructural analysis by SEM or TEM provides the most detailed assessment of follicular morphology.

In this study, ultrastructural analyses by SEM and TEM were available for a subset of 52 patients (Figures 4 and 5). This restriction did not reflect a biological or clinical selection criterion but was primarily driven by technical and logistical constraints related to the availability of specialized pathology resources. Consequently, these findings should be interpreted as hypothesis-generating rather than fully representative, while still providing valuable mechanistic insights into follicular structural integrity. However, such techniques are not feasible in routine clinical practice. In everyday settings, macroscopic and light microscopic observations must therefore guide the selection of topical interventions to avoid unnecessary polypharmacy. Light microscopy can reveal surface fragility through cortical irregularities, fissuring, or shaft tapering, features that may reflect oxidative stress as well as impaired matrix activity. Indicators of oxidative stress may further manifest as irregular or diminished pigmentation, uneven melanosome distribution, or translucent cortical segments. Collectively, these structural features may serve as practical surrogate markers of underlying biological vulnerability. However, their predictive validity requires prospective validation before routine clinical applications. Such insights may help stratify patient expectations, inform clinical decision-making regarding scalp cooling, and guide the use of adjunctive restorative strategies to further enhance follicular protection.

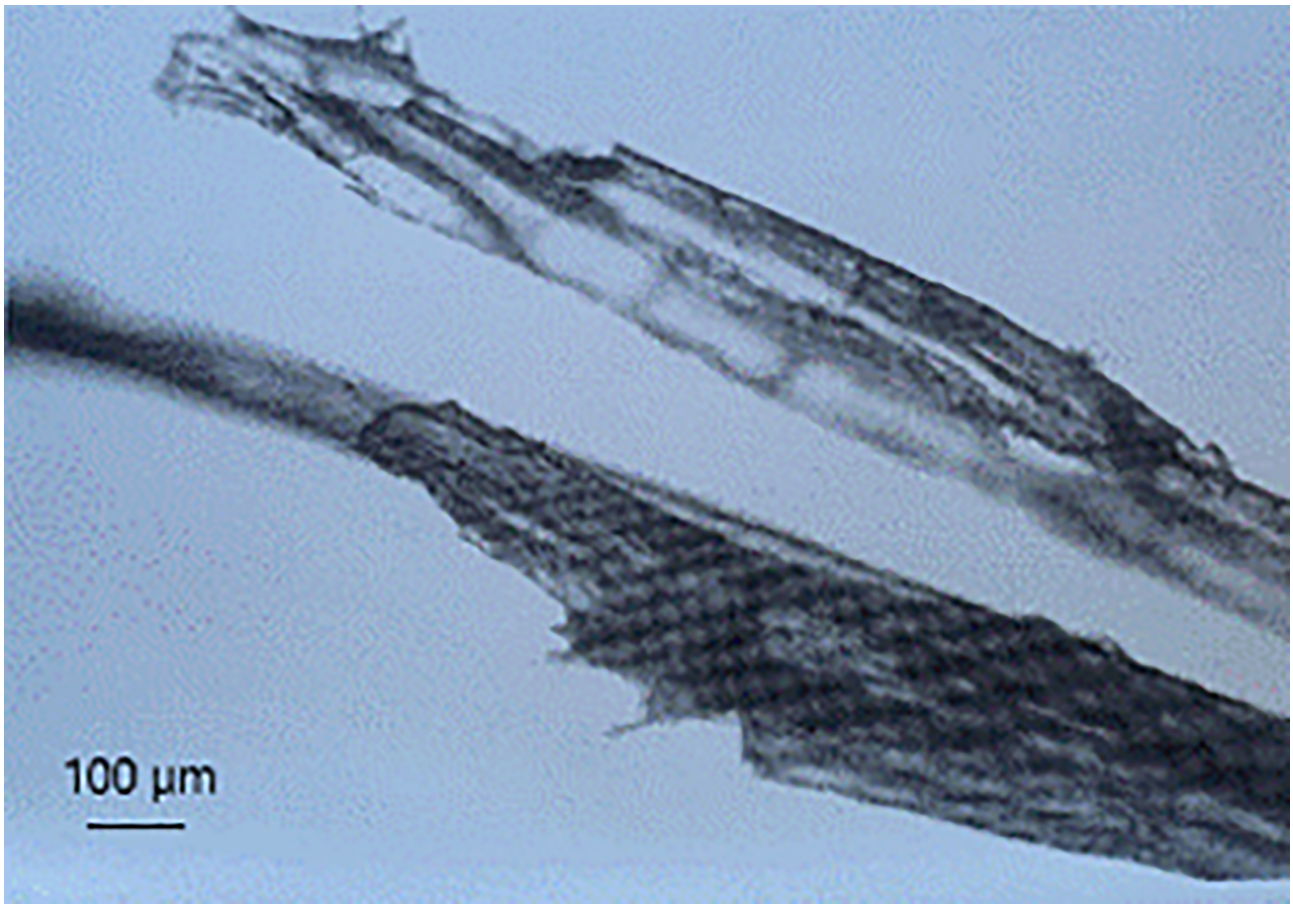
XSL•FO
RenderX

Figure 4. Tapering of hair shaft.



XSL•FO
RenderX

Figure 5. Light microscopic image of the hair shaft showing structural irregularities such as cuticle lifting, focal cuticle disruption, and longitudinal cortical translucencies.



Preventing visible hair loss, defined as an HMI greater than 50 and achieved in 53% of patients, represents not only a biological endpoint but also a meaningful psychological resource, highlighting the broader relevance of scalp cooling beyond esthetics with respect to self-image, quality of life, and potentially treatment engagement. At the same time, the observational single-arm design precludes the causal attribution of psychosocial effects specifically to scalp cooling, as quality-of-life outcomes are influenced by multiple clinical and individual factors.

These observations suggest that patients' subjective interpretation and personal evaluation of scalp cooling may, in some individuals, be more closely associated with posttreatment quality-of-life measures than the objectively quantified extent of hair preservation itself. However, these relationships should be interpreted cautiously and as associative rather than causally established within the limitations of the present observational design. Accordingly, these findings are best understood as preliminary interpretive observations that require further validation before broader psychosocial conclusions can be drawn. Any potential links between biological hair preservation, ultrastructural follicular findings, and psychosocial outcomes should therefore currently be regarded as descriptive and not as evidence of a directly established mechanistic relationship.

From a psychological perspective, these findings may be interpreted within the framework of cognitive appraisal theory [50]. According to this model, individuals continuously evaluate situations in terms of personal relevance and controllability. Within this context, a positive appraisal of scalp cooling may contribute to a greater sense of security and subjective control during treatment, both of which may act as protective psychological resources. Such resources reduce perceived burden by mitigating feelings of helplessness and uncertainty, while acting as a buffer against stress. In this way, hair preservation under scalp cooling contributes not only to external appearance but also to psychological stability, reinforcing its role as a cornerstone of supportive cancer care [51-53].

In addition, the relative stability of quality of life between preintervention and postintervention assessments may partly reflect the influence of underlying psychological dispositions. However, this pattern is unlikely to be determined by attitudes alone and may also be shaped by patients' experiences during treatment. These factors are closely linked to body image and psychosocial self-perception, both of which are well-established determinants of quality of life in oncological care [45].

Taken together, these observations suggest that cognitive appraisal represents not only a statistical correlate but also a potentially clinically relevant psychological mechanism. Positive appraisal may be associated with greater resilience and more adaptive coping strategies and could contribute to psychological well-being throughout treatment. Thus, beyond the biomedical effects of scalp cooling, the subjective evaluation of the intervention may represent an important mediator of psychosocial outcomes [50,54-56].

XSL•FO
RenderX

Within the framework of cognitive appraisal, it is conceivable that the integration of adjunctive strategies targeting the follicular pigmentary unit may further support patients' sense of control and bodily integrity. Interventions aimed at stabilizing pigmentary functions and thereby enhancing the effectiveness of scalp cooling could contribute to a more favorable subjective treatment experience. We hypothesize that patients receiving not only scalp cooling but also adjunctive follicle-directed supportive measures could potentially experience improved coping with the overall burden of cancer therapy. Such an integrative approach may therefore contribute to psychological well-being and maintain quality of life throughout treatment [5,45,57].

The study has several limitations, including its single-center design and the relatively small cohort of 82 patients. The availability of electron microscopy data was limited to the first 52 consecutively enrolled patients. This restriction was determined by the labor-intensive and technically demanding nature of ultrastructural processing and the finite pathology resources available for electron microscopy analyses, rather than by clinical, biological, or outcome-related selection criteria. Consequently, the risk of selection bias within the ultrastructural subgroup is considered low. Although menopausal status was assessed exclusively in female participants, no association with hair preservation was observed in our cohort. Therefore, the absence of an equivalent classification for male participants is unlikely to have substantially influenced the interpretation of our findings and thus represents only a minor limitation.

Consistent with previous reports, no increased risk of scalp metastasis was observed, supporting scalp cooling as a safe option in cancer treatment [58-60].

In summary, this study broadens the current understanding of scalp cooling by identifying the stability of the follicular pigmentary unit and the preservation of shaft integrity as biologically relevant markers of follicular resilience. The findings support a translational perspective in which adjunctive follicle-directed protective strategies may complement scalp cooling. Potential approaches ranging from antioxidant support to the modulation of cell cycle dynamics, for example, by stabilizing or prolonging the anagen phase, remain exploratory and require prospective validation. Collectively, such concepts may contribute to improving the consistency of hair preservation beyond the currently observed response rates of approximately 50% across heterogeneous patient populations. Ultimately, the clinical relevance of these approaches may lie in achieving more reliable treatment-related hair preservation and supporting patient-centered outcomes.

Conclusion

Taken together, these findings suggest that the clinical relevance of scalp cooling may extend beyond the binary end point of hair preservation. Preserved hair mass was associated with more favorable PROs and appeared to interact with patients' appraisal of treatment-related changes, supporting a broader biopsychosocial perspective on the treatment experience. While causal relationships cannot be established from this observational study, the results indicate that the impact of scalp cooling may not be fully captured by hair-retention thresholds alone.

By integrating the ultrastructural evidence of follicular vulnerability with patient-reported psychosocial outcomes, this study highlights the multidimensional nature of treatment-related hair loss. These findings support a more nuanced, patient-centered evaluation of scalp cooling and provide a rationale for future strategies aimed at improving both the biological protection of the hair follicle and patients' experiences during chemotherapy.

Acknowledgments

The authors declare the use of generative AI in the research and writing process. Responsibility for the final manuscript lies entirely with the authors.

Funding

Statistical analysis provided by Adrian Heinze was funded by D Schaffrin-Nabe. For all other authors, no financial support was received for this work.

Data Availability

The data are mentioned in the article and can otherwise be obtained by contacting josten-nabe@onkologie-bochum.com.

Authors' Contributions

Investigation: DS-N
Methodology: DS-N, AJ-N, AH, AT, MS, RV
Validation: DS-N
Visualization: DS-N, AJ-N, AH, AT, MS, RV
Writing – original draft: DS-N
Writing – review and editing: DS-N, AJ-N, AH, AT, MS, RV
All authors read and approved the final manuscript.

XSL•FO
RenderX

Conflicts of Interest

Adrian Heinze received compensation from D Schaffrin-Nabe for statistical analyses related to this study, performed in the capacity of an independent research consultant.

Multimedia Appendix 1
Comprehensive baseline characteristics, treatment-related factors, and posttreatment clinical parameters together with variables derived from light microscopy, transmission electron microscopy, and scanning electron microscopy.
[DOCX File, 55 KB - cancer_v12i1e94635_app1.docx]

References

1. Balagula Y, Rosen ST, Lacouture ME. The emergence of supportive oncodermatology: the study of dermatologic adverse events to cancer therapies. J Am Acad Dermatol 2011 Sep;65(3):624-635. [doi: 10.1016/j.jaad.2010.06.051] [Medline: 21777992]
2. Choi EK, Kim IR, Chang O, et al. Impact of chemotherapy-induced alopecia distress on body image, psychosocial well-being, and depression in breast cancer patients. Psychooncology 2014 Oct;23(10):1103-1110. [doi: 10.1002/pon.3531] [Medline: 24664939]
3. Dua P, Heiland MF, Kracen AC, Deshields TL. Cancer-related hair loss: a selective review of the alopecia research literature. Psychooncology 2017 Apr;26(4):438-443. [doi: 10.1002/pon.4039] [Medline: 26594010]
4. Gregory RP, Cooke T, Middleton J, Buchanan RB, Williams CJ. Prevention of doxorubicin-induced alopedia by scalp hypothermia: relation to degree of cooling. Br Med J (Clin Res Ed) 1982 Jun 5;284(6330):1674. [doi: 10.1136/bmj.284.6330.1674] [Medline: 6805650]
5. Ibraheem K, Smith A, Collett A, Georgopoulos NT. Prevention of chemotherapy drug-mediated human hair follicle damage: combined use of cooling with antioxidant suppresses oxidative stress and prevents matrix keratinocyte cytotoxicity. Front Pharmacol 2025;16:1558593. [doi: 10.3389/fphar.2025.1558593] [Medline: 40693274]
6. Haslam IS, Smart E. Chemotherapy-induced hair loss: the use of biomarkers for predicting alopecic severity and treatment efficacy. Biomark Insights 2019;14:1177271919842180. [doi: 10.1177/1177271919842180] [Medline: 31037027]
7. Carelle N, Piotto E, Bellanger A, Germanaud J, Thuillier A, Khayat D. Changing patient perceptions of the side effects of cancer chemotherapy. Cancer 2002 Jul 1;95(1):155-163. [doi: 10.1002/cncr.10630] [Medline: 12115329]
8. McGarvey EL, Baum LD, Pinkerton RC, Rogers LM. Psychological sequelae and alopecia among women with cancer. Cancer Pract 2001;9(6):283-289. [doi: 10.1046/j.1523-5394.2001.96007.x] [Medline: 11879330]
9. Münstedt K, Manthey N, Sachsse S, Vahrson H. Changes in self-concept and body image during alopecia induced cancer chemotherapy. Support Care Cancer 1997 Mar;5(2):139-143. [doi: 10.1007/BF01262572] [Medline: 9069615]
10. Boehmke MM, Dickerson SS. Symptom, symptom experiences, and symptom distress encountered by women with breast cancer undergoing current treatment modalities. Cancer Nurs 2005;28(5):382-389. [doi: 10.1097/00002820-200509000-00008] [Medline: 16192830]
11. Slominski A, Paus R, Plonka P, et al. Pharmacological disruption of hair follicle pigmentation by cyclophosphamide as a model for studying the melanocyte response to and recovery from cytotoxic drug damage in situ. J Invest Dermatol 1996 Jun;106(6):1203-1211. [doi: 10.1111/1523-1747.ep12348479] [Medline: 8752658]
12. Wikramanayake TC, Mauro LM, Tabas IA, Chen AL, Llanes IC, Jimenez JJ. Cross-section trichometry: a clinical tool for assessing the progression and treatment response of alopecia. Int J Trichology 2012 Oct;4(4):259-264. [doi: 10.4103/0974-7753.111221] [Medline: 23766610]
13. Whiting DA. Histology of the human hair follicle. In: Blume-Peytavi U, Tosti A, Trüeb R, editors. Hair Growth and Disorders: Springer; 2008:107-123. [doi: 10.1007/978-3-540-46911-7_7]
14. Paus R, Cotsarelis G. The biology of hair follicles. N Engl J Med 1999 Aug 12;341(7):491-497. [doi: 10.1056/NEJM199908123410706] [Medline: 10441606]
15. Roldan-Kalil J, Zueva L, Alves J, Tsytsarev V, Sanabria P, Inyushin M. Amount of melanin granules in human hair defines the absorption and conversion to heat of light energy in the visible spectrum. Photochem Photobiol 2023;99(4):1092-1096. [doi: 10.1111/php.13744] [Medline: 36403200]
16. Kim YD, Jeon SY, Ji JH, Lee WS. Development of a classification system for extrinsic hair damage: standard grading of electron microscopic findings of damaged hairs. Am J Dermatopathol 2010 Jul;32(5):432-438. [doi: 10.1097/DAD.0b013e3181c38549] [Medline: 20414091]
17. Brook TS, Seetsen T, Dercksen MW, et al. Results of the Dutch Scalp Cooling Registry in 7424 patients: analysis of determinants for scalp cooling efficacy. Oncologist 2024 Oct 3;29(10):e1386-e1395. [doi: 10.1093/oncolo/oyae116] [Medline: 38869252]
18. Schaffrin-Nabe D, Schmitz I, Josten-Nabe A, von Hehn U, Voigtmann R. The influence of various parameters on the success of sensor-controlled scalp cooling in preventing chemotherapy-induced alopecia. Oncol Res Treat 2015;38(10):489-495. [doi: 10.1159/000440636] [Medline: 26451590]

XSL•FO
RenderX

19. Paus R, Haslam IS, Sharov AA, Botchkarev VA. Pathobiology of chemotherapy-induced hair loss. Lancet Oncol 2013 Feb;14(2):e50-e59. [doi: 10.1016/S1470-2045(12)70553-3] [Medline: 23369683]
20. Purba TS, Ng'andu K, Brunken L, et al. CDK4/6 inhibition mitigates stem cell damage in a novel model for taxane-induced alopecia. EMBO Mol Med 2019 Oct;11(10):e11031. [doi: 10.15252/emmm.201911031] [Medline: 31512803]
21. Choi M, Kim MS, Park SY, et al. Clinical characteristics of chemotherapy-induced alopecia in childhood. J Am Acad Dermatol 2014 Mar;70(3):499-505. [doi: 10.1016/j.jaad.2013.10.034] [Medline: 24355411]
22. Bodó E, Tobin DJ, Kamenisch Y, et al. Dissecting the impact of chemotherapy on the human hair follicle: a pragmatic in vitro assay for studying the pathogenesis and potential management of hair follicle dystrophy. Am J Pathol 2007 Oct;171(4):1153-1167. [doi: 10.2353/ajpath.2007.061164] [Medline: 17823286]
23. Hassan YF, Shabaan DA. Effect of N-acetylcysteine on hair follicle changes in mouse model of cyclophosphamide-induced alopecia: histological and biochemical study. Histochem Cell Biol 2024 Jun;161(6):477-491. [doi: 10.1007/s00418-024-02282-0] [Medline: 38641701]
24. del Marmol V, Ito S, Bouchard B, et al. Cysteine deprivation promotes eumelanogenesis in human melanoma cells. J Invest Dermatol 1996 Nov;107(5):698-702. [doi: 10.1111/1523-1747.ep12365591] [Medline: 8875952]
25. Galván I, Wakamatsu K, Camarero PR, Mateo R, Alonso-Alvarez C. Low-quality birds do not display high-quality signals: the cysteine-pheomelanin mechanism of honesty. Evolution 2015 Jan;69(1):26-38. [doi: 10.1111/evo.12549] [Medline: 25330349]
26. Slominski A, Tobin DJ, Shibahara S, Wortsman J. Melanin pigmentation in mammalian skin and its hormonal regulation. Physiol Rev 2004 Oct;84(4):1155-1228. [doi: 10.1152/physrev.00044.2003] [Medline: 15383650]
27. Solano F. Photoprotection and skin pigmentation: melanin-related molecules and some other new agents obtained from natural sources. Molecules 2020 Mar 27;25(7):1537. [doi: 10.3390/molecules25071537] [Medline: 32230973]
28. Benito-Martínez S, Salavessa L, Raposo G, Marks MS, Delevoye C. Melanin transfer and fate within keratinocytes in human skin pigmentation. Integr Comp Biol 2021 Oct 14;61(4):1546-1555. [doi: 10.1093/icb/icab094] [Medline: 34021340]
29. Wang JX, Fukunaga-Kalabis M, Herlyn M. Crosstalk in skin: melanocytes, keratinocytes, stem cells, and melanoma. J Cell Commun Signal 2016 Sep;10(3):191-196. [doi: 10.1007/s12079-016-0349-3] [Medline: 27553358]
30. Fellows AP, Casford MTL, Davies PB. Nanoscale molecular characterization of hair cuticle cells using integrated atomic force microscopy-infrared laser spectroscopy. Appl Spectrosc 2020 Dec;74(12):1540-1550. [doi: 10.1177/0003702820933942] [Medline: 32462900]
31. Du F, Li J, Zhang S, Zeng X, Nie J, Li Z. Oxidative stress in hair follicle development and hair growth: signalling pathways, intervening mechanisms and potential of natural antioxidants. J Cell Mol Med 2024 Jun;28(12):e18486. [doi: 10.1111/jcmm.18486] [Medline: 38923380]
32. Hirai T, Ikeda-Imafuku M, Tasaka N, et al. Human hair keratin responds to oxidative stress via reactive sulfur and supersulfides. Adv Redox Res 2024 Apr;10:100091. [doi: 10.1016/j.arres.2023.100091]
33. Lee WS. Photoaggravation of hair aging. Int J Trichology 2009 Jul;1(2):94-99. [doi: 10.4103/0974-7753.58551] [Medline: 20927230]
34. Coderch L, Alonso C, García MT, Pérez L, Martí M. Hair lipid structure: effect of surfactants. Cosmetics 2023;10(4):107. [doi: 10.3390/cosmetics10040107]
35. Breakspear S, Smith JR, Luengo G. Effect of the covalently linked fatty acid 18-MEA on the nanotribology of hair's outermost surface. J Struct Biol 2005 Mar;149(3):235-242. [doi: 10.1016/j.jsb.2004.10.003] [Medline: 15721577]
36. Grosvenor AJ, Deb-Choudhury S, Middlewood PG, et al. The physical and chemical disruption of human hair after bleaching—studies by transmission electron microscopy and redox proteomics. Int J Cosmet Sci 2018 Dec;40(6):536-548. [doi: 10.1111/ics.12495] [Medline: 30229956]
37. Richena M, Rezende CA. Morphological degradation of human hair cuticle due to simulated sunlight irradiation and washing. J Photochem Photobiol B 2016 Aug;161:430-440. [doi: 10.1016/j.jphotobiol.2016.06.002] [Medline: 27318298]
38. Csuka DA, Csuka EA, Juhász MLW, Sharma AN, Mesinkovska NA. A systematic review on the lipid composition of human hair. Int J Dermatol 2023 Mar;62(3):404-415. [doi: 10.1111/ijd.16109] [Medline: 35218566]
39. Selleri S, Seltmann H, Gariboldi S, et al. Doxorubicin-induced alopecia is associated with sebaceous gland degeneration. J Invest Dermatol 2006 Apr;126(4):711-720. [doi: 10.1038/sj.jid.5700175] [Medline: 16470179]
40. Pappas A. Epidermal surface lipids. Dermatoendocrinol 2009 Mar;1(2):72-76. [doi: 10.4161/derm.1.2.7811] [Medline: 20224687]
41. Thiele JJ, Weber SU, Packer L. Sebaceous gland secretion is a major physiologic route of vitamin E delivery to skin. J Invest Dermatol 1999 Dec;113(6):1006-1010. [doi: 10.1046/j.1523-1747.1999.00794.x] [Medline: 10594744]
42. Picardo M, Ottaviani M, Camera E, Mastrofrancesco A. Sebaceous gland lipids. Dermatoendocrinol 2009 Mar;1(2):68-71. [doi: 10.4161/derm.1.2.8472] [Medline: 20224686]
43. Md Jaffri J. Reactive oxygen species and antioxidant system in selected skin disorders. Malays J Med Sci 2023 Feb;30(1):7-20. [doi: 10.21315/mjms2023.30.1.2] [Medline: 36875194]
44. Ekwall EM, Nygren LML, Gustafsson AO, Sorbe BG. Determination of the most effective cooling temperature for the prevention of chemotherapy-induced alopecia. Mol Clin Oncol 2013 Nov;1(6):1065-1071. [doi: 10.3892/mco.2013.178] [Medline: 24649294]

XSL•FO
RenderX

45. van den Hurk CJG, Mols F, Vingerhoets AJJM, Breed WPM. Impact of alopecia and scalp cooling on the well-being of breast cancer patients. Psychooncology 2010 Jul;19(7):701-709. [doi: 10.1002/pon.1615] [Medline: 19637250]
46. van den Hurk CJG, van de Poll-Franse LV, Breed WPM, Coebergh JWW, Nortier JWR. Scalp cooling to prevent alopecia after chemotherapy can be considered safe in patients with breast cancer. Breast 2013 Oct;22(5):1001-1004. [doi: 10.1016/j.breast.2013.07.039] [Medline: 23973267]
47. Zhang Y, Jimenez JJ. Mild oxidative stress protects against chemotherapy-induced hair loss. Front Oncol 2022;12:1078916. [doi: 10.3389/fonc.2022.1078916] [Medline: 36703797]
48. Chen Y, Zhou L, Ding Y, et al. Kartogenin regulates hair growth and hair cycling transition. Int J Med Sci 2022;19(3):537-545. [doi: 10.7150/ijms.68434] [Medline: 35370470]
49. Soma T, Tsuji Y, Hibino T. Involvement of transforming growth factor-beta2 in catagen induction during the human hair cycle. J Invest Dermatol 2002 Jun;118(6):993-997. [doi: 10.1046/j.1523-1747.2002.01746.x] [Medline: 12060393]
50. Lazarus RS, Folkman S. Stress, Appraisal, and Coping: Springer Publishing; 1984.
51. Bigatti SM, Steiner JL, Miller KD. Cognitive appraisals, coping and depressive symptoms in breast cancer patients. Stress Health 2012 Dec;28(5):355-361. [doi: 10.1002/smi.2444] [Medline: 22888085]
52. Kershaw TS, Mood DW, Newth G, et al. Longitudinal analysis of a model to predict quality of life in prostate cancer patients and their spouses. Ann Behav Med 2008 Oct;36(2):117-128. [doi: 10.1007/s12160-008-9058-3] [Medline: 18830672]
53. Almeida M, Griff MI, Brandão T. Coping and positive body image in young women with breast cancer: the buffering role of social support. Healthcare (Basel) 2025 Feb 6;13(3):346. [doi: 10.3390/healthcare13030346] [Medline: 39942534]
54. Folkman S, Lazarus RS, Gruen RJ, DeLongis A. Appraisal, coping, health status, and psychological symptoms. J Pers Soc Psychol 1986 Mar;50(3):571-579. [doi: 10.1037//0022-3514.50.3.571] [Medline: 3701593]
55. Pendergrass JC, Targum SD, Harrison JE. Cognitive impairment associated with cancer: a brief review. Innov Clin Neurosci 2018 Feb 1;15(1-2):36-44. [Medline: 29497579]
56. Liu Z, Doege D, Thong MSY, et al. Distress mediates the relationship between cognitive appraisal of medical care and benefit finding/posttraumatic growth in long-term cancer survivors. Cancer 2021 Oct;127(19):3680-3690. [doi: 10.1002/cncr.33684] [Medline: 34196976]
57. Tamai N, Minematsu T, Ikeda M, Mugita Y, Sanada H. Effects of improved hypoallergenic fabrics in medical wigs in patients with breast cancer with chemotherapy-induced alopecia: a randomised clinical trial. BMJ Support Palliat Care 2024 Dec 19;14(e3):e2563-e2570. [doi: 10.1136/bmjspcare-2020-002309] [Medline: 33452047]
58. Rugo HS, Melin SA, Voigt J. Scalp cooling with adjuvant/neoadjuvant chemotherapy for breast cancer and the risk of scalp metastases: systematic review and meta-analysis. Breast Cancer Res Treat 2017 Jun;163(2):199-205. [doi: 10.1007/s10549-017-4185-9] [Medline: 28275922]
59. Lemieux J, Amireault C, Provencher L, Maunsell E. Incidence of scalp metastases in breast cancer: a retrospective cohort study in women who were offered scalp cooling. Breast Cancer Res Treat 2009 Dec;118(3):547-552. [doi: 10.1007/s10549-009-0342-0] [Medline: 19241158]
60. Ross M, Fischer-Cartlidge E. Scalp cooling: a literature review of efficacy, safety, and tolerability for chemotherapy-induced alopecia. Clin J Oncol Nurs 2017 Apr 1;21(2):226-233. [doi: 10.1188/17.CJON.226-233] [Medline: 28315539]

Abbreviations

CIA: chemotherapy-induced alopecia
EORTC: European Organisation for Research and Treatment of Cancer
HMI: hair mass index
OR: odds ratio
PRO: patient-reported outcome
ROS: reactive oxygen species
SEM: scanning electron microscopy
TAC: simultaneously administered docetaxel, doxorubicin, and cyclophosphamide
TEM: transmission electron microscopy

Edited by M Balcarras; submitted 04.Mar.2026; peer-reviewed by M Chakit, T Fujii; revised version received 10.Jun.2026; accepted 10.Jun.2026; published 20.Jul.2026.

Please cite as:
Schaffrin-Nabe D, Josten-Nabe A, Heinze A, Tannapfel A, Schaffrin M, Voigtmann R
Scalp Cooling as a Biopsychosocial Intervention in Patients Receiving Highly Alopecia-Inducing Chemotherapy: Prospective Single-Arm Study
JMIR Cancer 2026;12:e94635
URL: https://cancer.jmir.org/2026/1/e94635
doi:10.2196/94635

XSL•FO
RenderX

© Dörthe Schaffrin-Nabe, Anke Josten-Nabe, Adrian Heinze, Andrea Tannapfel, Merle Schaffrin, Rudolf Voigtmann. Originally published in JMIR Cancer (https://cancer.jmir.org), 20.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Multimethod Evaluation to Assess Feasibility, Acceptability, and Preliminary Efficacy of HPVVaxFacts, a Tailored Mobile Web App, for Parents With Unvaccinated Children: Pilot 2-Arm Randomized Controlled Trial

Jennifer Cunningham-Erves[1], PhD; Pamela C Hull[2], PhD; Amanda F Dempsey[3], MD, PhD; Tatsuki Koyama[4], PhD; Lili Sun[4], PhD; Mariya Chalise[1], MPH; Elizabeth C Stewart[5], PhD; Jessica L Jones[6], MS; Douglas Landsittel[7], PhD; Meredith Smalls[6], MPH; Freneka F Minter[6], PhD; Lindsay S Mayberry[8], PhD; Lora Harnack[9], MSN; Michael J Hook[9], MBA; Janet Cates[9], BA; Tammy Van Maanen[10]; Taylor Burnett[11], MHS; Consuelo H Wilkins[8], MD, MSCI

[1]Department of Health Policy, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 700, Nashville, TN, United States

[10]Spring Creek Pediatrics, Chattanooga, TN, United States

[11]Department of Family Medicine, Meharry Medical College, Nashville, TN, United States

[2]Behavioral Science Department, University of Kentucky, Lexington, KY, United States

[3]MERCK, Rahway, NJ, United States

[4]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, United States

[5]Graduate School of Public Health Practice, Meharry Medical College, Nashville, TN, United States

[6]Meharry-Vanderbilt Alliance, Nashville, TN, United States

[7]Department of Biostatistics, The State University of New York, University at Buffalo, Buffalo, NY, United States

[8]Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, United States

[9]Cumberland Pediatric Foundation, Nashville, TN, United States

Corresponding Author:
Jennifer Cunningham-Erves, PhD
Department of Health Policy, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 700, Nashville, TN, United States

Abstract

Background: Mobile health (mHealth) interventions may improve provider-parent communication on human papillomavirus (HPV) vaccination to reduce concerns, and increase intention and uptake. HPVVaxFacts (233 Analytics) is a novel, mobile web app delivering tailored education based on the Health Belief Model and Theory of Reasoned Action, addressing parental concerns preclinic visit.

Objective: This study aimed to assess the feasibility, acceptability, and preliminary efficacy of HPVVaxFacts among parents of adolescents aged 9 - 17 years.

Methods: We conducted a pilot, randomized controlled trial in 2 urban Tennessee clinics from June to September 2023 comparing 2 groups: tailored education via HPVVaxFacts mobile web app (intervention, n=27), and nutrition education (attention control, n=30). Eligible parents had or were caregivers to a child aged 9 to 17 years unvaccinated against HPV, had a mobile phone, had an upcoming clinic visit, and spoke English. The recruitment strategy was patient intake software—Phreesia (Phreesia, Inc) and eClinicalWorks (eClinicalWorks). Although unblinded, parents could deduce their study arm assignment. Providers were blinded. Feasibility, acceptability, and preliminary efficacy (HPV vaccine knowledge, concerns, intentions, and vaccination rates) were assessed using multimethod evaluation. Parents were assessed at baseline and immediately post intervention via surveys. Vaccination rates were assessed at 12 months post intervention via electronic health records. Nineteen parent interviews were conducted up to 9 months post intervention. A clinic staff consultation (n=6) was 1 month post intervention.

Results: Of 57 enrolled parents, most were female (52/57, 91%), non-Hispanic White (44/57, 77%), had ≤US $80,000 household income (32/57, 56%), and had some college or less (27/57, 47%). In total, 81% (29/36) of parents viewed HPVVaxFacts. Post intervention, HPV vaccine initiation was higher in the intervention group compared to the attention control group (48% vs 17%; difference 0.24; 95% CI 0.03-0.46; *P*=.01). Parents in the HPVVaxFacts arm demonstrated a greater reduction in knowledge (ie, knowledge increase; mean change: −0.6 vs 0.1) and concern scores (mean change: −3.4 vs −1.4) than those in the nutrition education arm. However, between-arm differences were not statistically significant (*P*=.13 and *P*=.14, respectively). The majority found the study protocol and HPVVaxFacts acceptable. Benefits of HPVVaxFacts include confirming their decision to vaccinate,

XSL•FO
RenderX

supporting parent-child discussion on the vaccine, and answering questions preclinic visit or offering questions for the provider. Study protocol delivery and mobile web app instructions were suggested areas for improvement. Barriers for HPVVaxFacts use include content in English only and digital format.

Conclusions: Our study suggests HPVVaxFacts was feasible and acceptable among parents to provide previsit, tailored information on HPV vaccination. Outcomes offer a positive trajectory but need more exploration. Next steps include a well-powered efficacy trial to determine the impact of HPVVaxFacts on initiation vaccine rates and parental hesitancy factors, as well as to explore an interaction, effect modification, and mediation among different variables.

Trial Registration: ClinicalTrials.gov NCT04380623; https://clinicaltrials.gov/study/NCT04380623

(JMIR Cancer 2026;12:e78910) doi:10.2196/78910

KEYWORDS

human papillomavirus; vaccine uptake; digital intervention; mHealth; children and adolescents; parents; HPV; mobile health

Introduction

Parental vaccine hesitancy remains a critical barrier to suboptimal rates of human papillomavirus (HPV) vaccination. Studies show up to 63% of US parents are hesitant [1,2] (ie, reluctant or refuse to vaccinate) despite its availability and proven effectiveness [3]. This contributes to stalled HPV vaccination rates among adolescents aged 13 - 17 years, 63% in the United States and 64% in Tennessee in 2025 [4]. Unvaccinated adolescents are at risk for HPV infection, which is implicated in almost all cervical cancers and 63% - 90% of penile, vulvar, vaginal, anal, and oropharyngeal cancers [5]. Many of these HPV-related cancers are rising in incidence [6]. Based on several US studies, parental predictors for HPV vaccine hesitancy are complex and multifactorial [7-9]. Common reasons include concern about safety or side effects, lack of knowledge on HPV and associated cancers, and lack of or poor-quality provider recommendations and communication [10]. Hesitant parents often report dissatisfaction with patient-provider communication, further undermining their confidence and decision-making to initiate and complete HPV vaccination [11-14]. There is an urgent call to identify evidence-based practices (EBPs) to address vaccine hesitancy globally and nationally [4,13,15].

Few interventions to date have been found impactful in addressing parental HPV vaccine hesitancy. Mobile health (mHealth) interventions show promise in enhancing traditional medical care. With provider-based strategies (eg, a strong recommendation and motivational interviewing), mHealth interventions could positively impact parental concerns by countering misinformation or addressing lack of information about HPV vaccination [16-22]. Some data exist on mHealth in promoting HPV vaccination, but important limitations remain. A systematic review from January 2017 to July 2022 identified 24 digital health interventions for HPV vaccination, with only 4 using theory-based mHealth interventions [23-26]. Limitations of these interventions are (1) a focus on general HPV vaccine promotion and not unique concerns of vaccine-hesitant parents, (2) targeted only at girls or younger adolescents, and (3) limited testing across diverse sociodemographics and geography. Two of these studies tailored information using reminders or by individual vaccine status [23,26-29], but these approaches were limited in impact. There remains an unmet need to develop an efficacious mHealth program that provides parents succinct, resonating messages individually tailored to their concerns on the HPV vaccine for use by providers to supplement parent-provider in-clinic discussion [30,31].

Past studies suggested that many parents wanted HPV vaccine information before a clinic visit with their child's provider. Information could be delivered prior to the clinic visit using mHealth [16-22]. The purpose of receiving this information was to get the answers to the questions they may have about HPV and the vaccine prior to the visit. Furthermore, parents used this information to engage in parent-provider communication about the vaccination and support shared decision-making (J Cunningham-Erves, unpublished data, June 2022).

This study's objective was to conduct a pilot test of a mobile web app intervention, HPVVaxFacts (233 Analytics), designed to provide tailored, previsit education to improve HPV vaccination rates. We hypothesized that : (1) the HPVVaxFacts intervention arm would achieve acceptability and feasibility rates ≥70%; and (2) parents in the intervention arm would show greater reductions in vaccine concerns and greater increases in knowledge, intentions, and vaccination rates compared to the attention control arm. To prepare for a future full-scale randomized controlled trial (RCT) to determine efficacy, our aims were as follows:

- Aim 1: To assess feasibility of the HPVVaxFacts intervention and overall study protocol.
- Aim 2: To assess acceptability of the HPVVaxFacts intervention within 2 urban pediatric clinics.
- Aim 3: To explore preliminary efficacy of the HPVVaxFacts intervention on the parent's decision to get their adolescents the HPV vaccination.

Methods

Setting and Study Design

A multimethod approach was used to evaluate the feasibility, acceptability, and preliminary efficacy of a mobile web app, HPVVaxFacts, among parents of unvaccinated children aged 9 to 17 years [32]. A pilot, 2-arm, multisite RCT was conducted at 2 urban pediatric clinics in Tennessee to determine feasibility and acceptability. Clinic Site 1 is an urban pediatric clinic serving more than 8000 children with a broader demographic of families with varied socioeconomic backgrounds. There were

XSL•FO
RenderX

421 unvaccinated children at Site 1 from June to September 2023. In contrast, Clinic Site 2 is an urban pediatric clinic serving more than 5000 children primarily uninsured, underinsured, and marginalized children. A total of 23 unvaccinated children were at Site 2 for 1 week in August 2023. Clinic providers and staff stated they had higher vaccination rates, so they chose a 1-week back-to-school event where vaccinations were offered. This time was perceived to have the highest volume of unvaccinated patients. While operational workflows (ie, administrative protocols, patient volume, and technology adoption) and catchment areas differ, both had patient intake management platforms for recruitment and followed the standardized protocols to be discussed in detail below.

Parents in both arms who agreed to follow-up were interviewed to explore acceptability and feasibility of the protocol using a qualitative, descriptive study design. Intervention acceptability and parent-child discussion were examined among parents in the intervention arm only. The reporting of the pilot RCT follows the CONSORT (Consolidated Standards of Reporting Trials) 2010 guidelines to report a pilot or feasibility trial (Multimedia Appendix 1) and the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist to evaluate web-based apps (CONSORT-HEALTH checklist in Checklist 1) [33].

Ethical Considerations

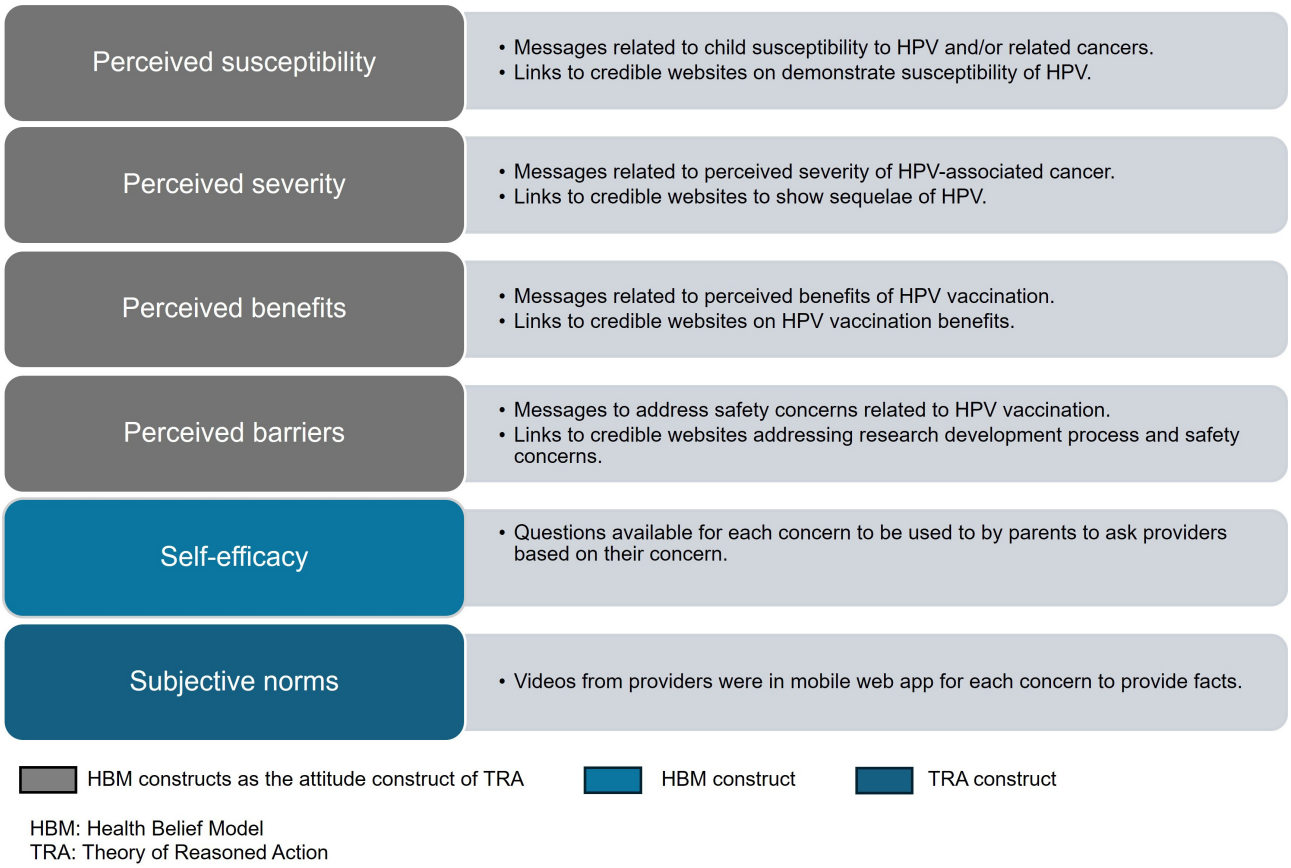
The Institutional Review Board (IRB) of Meharry Medical College (Protocol Number: 18-12-890) approved this study as expedited as it was considered no more than minimal risk. The study was registered at ClinicalTrials.gov (NCT04380623) and performed in line with the principles of the Declaration of Helsinki. Electronic consent was obtained from participants using REDCap (Vanderbilt University) [34,35] for the pilot RCT, and verbal consent for the semistructured interviews. To protect participant information, survey and interview data were stored via REDCap and on a password-protected server. Data were deidentified prior to provision to statisticians and research team members for data analysis. Dissemination of results was aggregate and anonymous. Participants were compensated with a US $30 gift card for pilot study participation, and a US $20 gift card for completing a follow-up interview.

Intervention

HPVVaxFacts provides individually tailored education on HPV and the vaccine to parents [36]. Education and video content for each concern was developed using the Health Belief Model and Theory of Reasoned Action (Figure 1). To tailor the education in the app, each parent answered a 10-item baseline survey adapted from the Vaccination Confidence Scale [37] to identify parental concerns. The survey responses are mapped to an algorithm that selects the top 3 parent concerns. Parents can change the selection if they disagree. Parents then review the content (ie, educational information, provider's video, questions for the provider, and links to other credible sources verifying content) associated with each concern. An adolescent corner is available for parents to share with and/or have an optional discussion with their child about each concern. Finally, parents could submit concerns or additional needs via the provision of feedback. Refer to Cunningham-Erves et al [36] for full description of HPVVaxFacts development.

XSL•FO
RenderX

Figure 1. Examples of theory application in HPVVaxFacts development. HPV: human papillomavirus.



Attention Control

Parents in the attention control arm received a link to a website providing nutrition education for parents of children and/or teens [38], a program unrelated to the study outcome. This approach served two purposes: (1) ensuring all parents receive a benefit regardless of study arm assignment, and (2) balancing time and effort between groups [39]. Because parents' perceptions of HPV vaccination significantly influence vaccination outcomes [40], an attention control group allowed us to account for potential placebo effects and isolate the specific effects of HPVVaxFacts. Furthermore, the study arm activities were comparable to the intervention activities. Parents completed a 5-item brief quiz with immediate feedback on correct or incorrect responses and explanations.

Study Participants

This study was conducted with parents who were caregivers for children in 2 urban pediatric clinics in Tennessee who were representative of the continuum of vaccine hesitancy. Eligible parents had a child aged 9 to 17 years who was unvaccinated against HPV, had a mobile phone, had an upcoming clinic visit, and spoke English. When multiple children per eligible caregiver existed, random selection determined which child would participate. Random numbers were first generated in Stata version 17 (StataCorp LLC). Children were ordered from youngest to oldest (or alphabetized by first names if twins or triplets). The child with the lowest corresponding random number was selected for the parent to be sent a study invitation. Eligibility criteria were chosen based on participating providers and staff input. They stated it was not feasible to enroll only hesitant parents. Instead, they perceived it was equally beneficial to recruit parents representing the continuum of hesitancy. Exclusion criteria were parents who did not consent, or their child had initiated or completed the HPV vaccination series.

Multimethod Evaluation

Quantitative Phase

Training and Preparation

The clinic sites had different recruitment approaches. Clinic Site 1 recruitment was conducted by 2 trained clinic staff. Clinic Site 2 recruitment was done by 2 members of the research team. Recruitment and enrollment procedures were standardized through 2-hour training. Training included a review of the study protocol and documents, integration with clinic workflow, and potential recruitment and retention challenges. Post training, a study protocol run-through was conducted at each clinic site to familiarize the study team with the protocol and identify necessary adjustments prior to implementation.

Site-Specific Recruitment Methods

At Clinic Site 1, parents were first introduced to the study via a post to the clinic's Facebook (Meta Platforms, Inc) page to increase study awareness in their clinic. At Clinic Site 2, flyers were posted in the clinic. Eligibility was determined via electronic health record (EHR) data by 2 trained clinic staff at Clinic Site 1 and 2 trained research staff at Clinic Site 2. The research team was trained by clinic staff to use EHR and eClinicalWorks (version 12.0.2; eClinicalWorks).

XSL•FO
RenderX

Enrollment Communication

Eligible parents received an IRB-approved text message explaining the study and inviting study participation, with emphasis on the HPV vaccination. If parents text "yes," a screener was sent text to parents to confirm eligibility. Trained clinic staff at Clinic Site 1 contacted parents via Phreesia, a HIPAA (Health Insurance Portability and Accountability Act)-compliant patient intake management platform that automates the check-in process as well as supports patient-provider communication [41]. At Clinic Site 2, parents were contacted via eClinicalWorks by study team members. Parents received up to 3 text messages on their mobile phones requesting participation.

Sample Size

GPower 3.1 program (version 3.1.9.7; Heinrich-Heine-Universität Düsseldorf) was used to conduct power analysis to determine the sample size. The necessary sample size for the pilot trial was calculated using a 1-sided 80% CI and the expected effect size of the future RCT [42], based on the expected proportion in the attention control arm (0.20 - 0.30) and the clinically meaningful difference to be observed in the intervention arm in the future RCT (at least 0.10 higher) [42,43]. A total pilot sample size of 46 - 60 accounts for expected attrition and was deemed sufficient. Our study goal was to enroll 70 parents, expecting 10% - 15% attrition (eg, no-show for clinic visit), yielding a final sample of 60 - 63 parents (28 - 30 per arm).

Randomization and Blinding

All data were collected via the parents' mobile phones. A member of the research study (coauthors: JC-E, ECS, and MS), all trained and with social science backgrounds, enrolled parents in the study. The biostatistician (coauthor DL) generated 2 randomization lists prior to study implementation. Using blocked randomization, randomization was conducted with blocks of 10 (using n=10 from each group). Group labels (denoted as 1 or 2) were generated in Stata version 17. Although parents could deduce their study arm assignment, providers were blinded. We had 2 protocol deviations in which 1 participant who decided not to participate later decided to participate, which led to 2 parents being assigned to the HPV program at the same time. We decided to retain both participants for the analysis. The second deviation was the continued enrollment of 7 participants more than the planned sample size due to consent initiation. However, these individuals were not a part of the random allocation process and were excluded [44].

Procedures

After confirming participant eligibility, parents were sent a REDCap link to their mobile phones to provide electronic consent up to 48 hours prior to the clinic visit. They were assigned to either the intervention or attention control arm post consent based on the computer-generated randomization numbers listed by a trained research team member with a social science background (lead author: JC-E). Parents were asked to view HPVVaxFacts or nutrition programs at least once prior to their upcoming clinic visit. At baseline, parents took the presurvey assessing their sociodemographics and knowledge and concerns related to HPV and the vaccine, and intentions to initiate vaccination. If randomized to the intervention arm, parents were sent a personalized link to the presurvey to assess knowledge and intention and then went to the mobile web app HPVVaxFacts to assess, receive, and confirm or change their top 3 concerns. After viewing the assigned mobile web app, parents took a postsurvey before the clinic visit with similar presurvey questions except for the sociodemographics. Parents in the intervention arm were also asked questions related to program acceptability. If any participant did not complete a survey or view the mobile web app, up to 3 reminders were sent for completion. Post study, the trained clinic staff provided a dataset on the child's HPV vaccine status at 0 and 12 months.

Outcomes

Overview

The clinic provided the vaccination status of each child at 12 months post intervention using the EHR (Primary Outcome). The pre-post survey asked parents about their knowledge, concerns, and intentions for their child to initiate the HPV vaccination process. Intervention acceptability and intervention use for parent-child discussion were assessed on the postsurvey of parents in the intervention, "HPVVaxFacts" study arm only. Intervention feasibility was assessed via several feasibility indicators (Secondary Outcomes). Sociodemographic data were collected on the presurvey only (Covariates).

Primary Outcome: HPV Vaccination Status (1-Item)

HPV vaccination status post study was retrieved from the clinics' EHR system. The child's vaccination status post study was labeled as "yes" and "no." Vaccination status was assessed at 12 months post intervention.

Secondary Outcomes

Knowledge of HPV and the Vaccine (10-Item Scale)

A validated knowledge scale [45] assessed parental knowledge of HPV and HPV vaccination. Response options were true, false, or unsure, in which some answers were true and others false. For analysis, false responses were reverse-coded, false and unsure answer responses were combined, and mean scores were calculated for pre- and postassessments. The lower the score, the greater the degree of knowledge a parent had. When measuring internal consistency reliability, the Cronbach α for knowledge data on presurvey was 0.80 and 0.81 on the postsurvey.

Concerns (9-Item Scale)

An adapted attitudes scale, the Vaccine Confidence Scale [37], assessed parental concerns on HPV vaccination based on a 4-point Likert scale to determine agreement. The parents receiving nutrition education completed this measure in REDCap, while parents receiving HPVVaxFacts completed these items in the mobile web app. The HPVVaxFacts mobile web app prompted parents to take a survey, receive their top 3 concerns based on a preexisting algorithm, and confirm and/or change their selection. The final selected top concerns were tracked within the app. For inferential analysis, mean pre- and postscores were calculated for both groups with the greater score indicating more parental vaccine concerns. Internal consistency was strong: Cronbach α 0.90 (presurvey) and 0.91 (postsurvey).

XSL•FO
RenderX

HPV Vaccination Intention (1-Item)

One question asked, "How likely is it that you will get your child the HPV vaccine?" with the response options "Definitely not likely," "Not likely," "Somewhat likely," and "Very likely" [46].

HPV Vaccine Recommendations

Parents were asked, "Has your child's doctor offered the HPV vaccine?" with answer options yes or no.

HPVVaxFacts Feasibility

We followed guidelines of Teresi et al [47] to assess various aspects of feasibility in our pilot study. We describe each below.

Data Collection Protocols and Data Availability

We documented whether participants complied with assessment and data collection protocols. This included (1) completion rates of the pre-post survey, (2) mobile web app tracking of parent concerns and selected top 3 concerns by date and time, and (3) the ability to obtain the necessary data from clinic administrative records, particularly age and HPV vaccination status. Feasibility was measured as 70% of parents viewing the mobile web app.

Conceptual and Psychometric Adequacy of Measures

We assessed the internal reliability of the concerns, knowledge, and intervention acceptability scales. Participants were allowed to provide feedback on the HPVVaxFacts mobile web app to determine whether important concerns were missing from the app, which suggests missed questions on the concern scale.

Intervention Implementation

We assessed the number of participants recruited and retained for the study.

HPVVaxFacts Acceptability (4-Items)

Parents were asked their acceptability of HPVVaxFacts using a validated, 4-item scale. Response options were on a 5-point Likert scale based on agreement [48]. An example item is "The HPV Education Program meets my approval." Acceptability was measured as 70% or more of parents stating they agree or strongly agree with each statement. Internal consistency yielded the Cronbach α value of 0.91 on presurvey data in this study.

Parent-Child Discussion (1-Item)

Developed by the research team, parents in the intervention arm were asked whether HPVVaxFacts was helpful in discussing the HPV vaccine with their child post study. The answer choices were on a 5-point Likert scale based on agreement.

Covariates: Sociodemographics (12-Items, Presurvey Only)

These variables were in 2 categories—parent and child sociodemographics.

Parent Sociodemographics

Age was a continuous variable. Sex was categorized as male, female, and prefer not to disclose. Parent sexual orientation included categories of asexual, bisexual, straight, fluid, and gay. Categories were combined to straight (heterosexual), prefer not to disclose, and other. Parents' race or ethnic categories were non-Hispanic White, African American, Latino or Spanish origin, Native Hawaiian or Other Pacific Islander, and Prefer Not to Answer. Education categories were General Educational Development (GED) or high school diploma, associate's degree, some college, bachelor's degree, master's degree, and doctoral or professional degree. Annual household income categories were less than US $20,000, US $20,001 to US $40,000, US $40,001 to US $60,000, US $60,001 to US $80,000, and more than $80,000, or did not answer.

Child Sociodemographics

Age, a continuous variable, was categorized to 9 - 10, 11 - 12, and ≥13. Sex was categorized as male, female, and prefer not to disclose. The child's race or ethnic categories were non-Hispanic White, African American, Latino or Spanish origin, American Indian and Alaskan Native, Prefer Not to Answer, and Other. The child's sexual orientation had several categories including aromantic, asexual, bisexual, fluid, straight, and gay. These categories were combined to straight (heterosexual), questioning or unsure, prefer not to disclose, and other.

Data Analysis

Descriptive statistics characterized the sample: continuous variables are reported as means and SDs, and categorical variables as frequencies and percentages. Between-group comparisons were conducted using appropriate statistical tests: Fisher exact test or chi-square tests for categorical variables and Wilcoxon rank-sum tests or *t* tests for continuous variables. Comparison of pre-post scores of HPV vaccine knowledge, concerns, and intention by study arm was performed using exploratory visual analysis. Provider recommendation was not included as a covariate because (1) recommendations were expected to be similar across study arms and (2) are rarely systematically documented in clinical practice [49], making reliable data collection impractical. Parents took the survey in the HPVVaxFacts mobile web app and were allowed to select their top concerns. The research team categorized the top 3 concerns and summed the concerns by question. The summed scores were sorted from highest to lowest, with the top 3 values being the top concerns.

Qualitative Phase

Recruitment

Interviews were conducted up to 9 months post intervention based on parent availability. On the postsurvey, 15 in the intervention study arm and 18 in the nutrition attention control agreed to be contacted for follow-up interviews. Of these participants, 9 in the intervention study arm and 10 in the nutrition attention control completed an interview. Other parents either could not be reached (n=9) to complete or later decided not to complete the interview (n=5).

Procedures

Parents in both intervention and attention control groups who agreed to be contacted for the follow-up interview were called to schedule the 20-minute interview. The purpose of these interviews was to complement quantitative data. Interviews were audio-recorded and transcribed for data analysis.

XSL•FO
RenderX

Interview Protocol

The semistructured interview protocol for parents who were assigned to the HPVVaxFacts intervention arm included questions that explored intervention acceptability. Parents in both study arms were queried regarding their study experiences and ways to improve the study protocol and implementation procedures.

Data Analysis

Interview transcripts were analyzed using a modified version of Braun and Clarke's 6-step thematic methodology [50]. A trained qualitative researcher and analyst (coauthors JC-E and TB) conducted the data analysis led by lead author JC-E, both with backgrounds in social science. Using the first 3 interview transcripts, coauthor JC-E created a codebook of a priori codes using line-by-line coding. Both coders used the codebook to independently code each transcript, creating reliability in the system and having each coder serve as a referent. The analysts then met to review coding decisions and reach a consensus on whether to accept, modify, remove, or add codes until a percentage agreement of >90% was reached. Coding saturation was met when no new codes emerged. Patterns and themes were generated by combining related codes and applying constant comparison methods across transcripts. Rigor was established using thick, rich descriptions of interview data. Interview data were managed using Excel (version 26; Microsoft 365).

Clinic Staff Feedback From Consultation

One month post intervention (October 2023), we held a 1-hour consultation (ie, an engagement strategy) with 2 clinic providers and 4 staff members over lunch to gather informal feedback on study protocol and implementation. Trained clinic staff coordinated the date and time of the consultation. The consultation explored ways to improve the study protocol and implementation procedures, and perceived barriers and facilitators to implementing HPVVaxFacts in routine clinic practice. This session was not audio-taped nor formally analyzed as qualitative research; instead, notes were recorded by coauthors JC-E and ES.

Overall, providers perceived that the program offered clinic benefits. Offering education preclinic visits to their patients was a top benefit. This complemented their recommendations and lessened clinic time for education. Another perceived benefit was the increased vaccination rates in the clinic. Regarding barriers, providers and clinic staff highlighted the need to improve logistics for large-scale implementation. Because the intervention was manual and required efforts from clinic staff and the research team to aid in recruitment and retention, providers questioned the feasibility of this study without full integration into the EHR but still perceived offering preclinic information in general as feasible. However, providers and staff stated they understood this was a pilot study and hoped their suggestions improved future implementation.

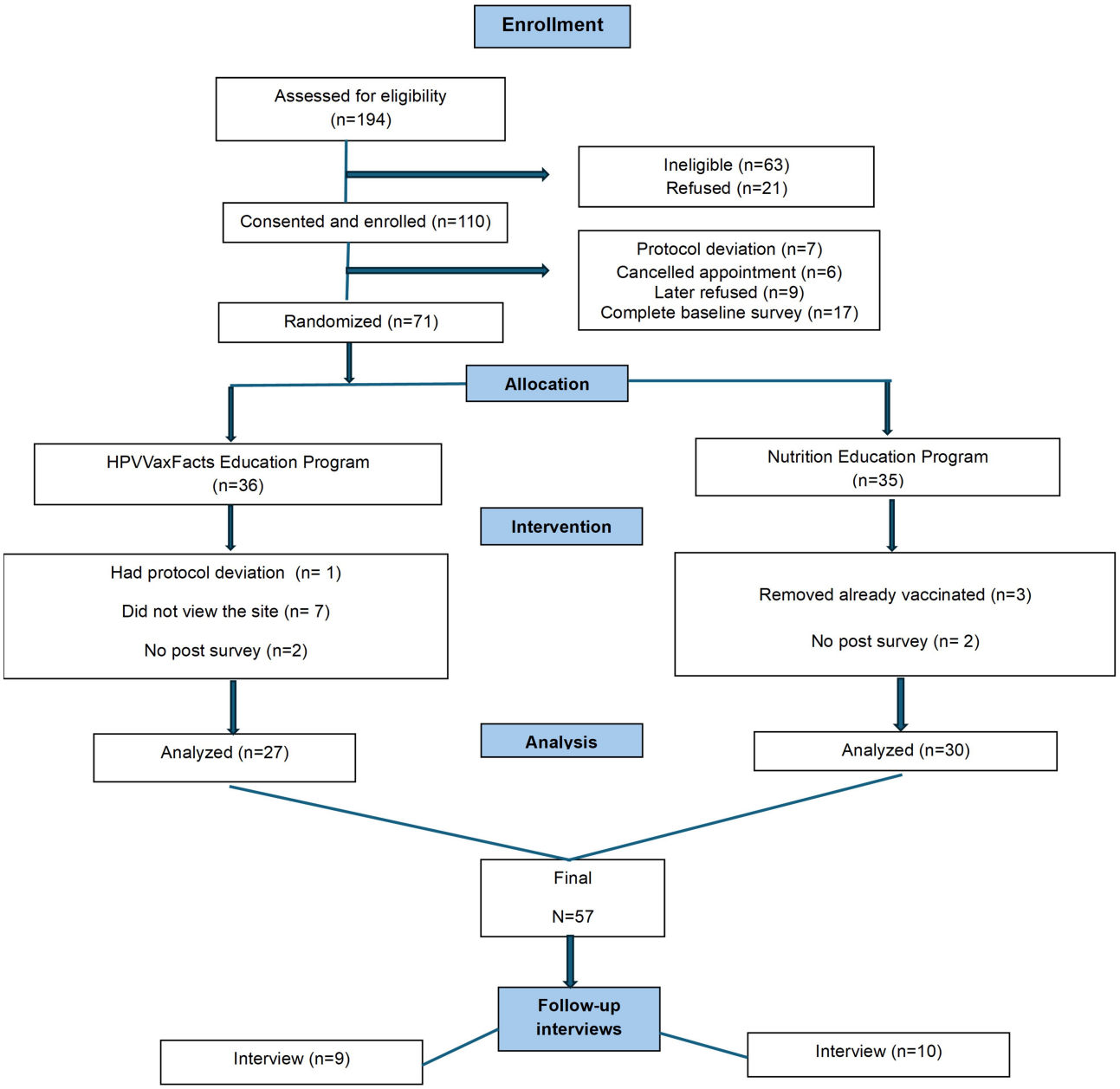
Results

Feasibility: Recruitment and Retention

We met our goal of training 2 clinic staff to aid in recruiting participants through Phreesia. For this study, the clinic staff sent potential parents an IRB-approved REDCap link with a screener. If they screened eligible, they received the presurvey, viewed the intervention, and took the postsurvey. Of the 194 assessed for eligibility, 63 were ineligible. A total of 21 (16%) declined to participate and 9 more later refused, increasing the refusal rate to 23%. We recruited 110 participants who consented and enrolled in the study (107 were recruited from Clinic Site 1 and 3 from Clinic Site 2) with an enrollment rate of 84%. Parents were enrolled 15 to 20 minutes up to 24 hours prior to the clinic visit. In total, 71 people were allocated to the study arms, a 35.5% reduction from enrollment to randomization. A total of 36 parents were randomly assigned to the HPVVaxFacts intervention arm, of which 7 did not go to the mobile web app. Of those who viewed the mobile web app (81%), 2 did not complete the postsurvey. These participants were excluded with a retention rate of 75%. A total of 35 people were randomly assigned to the attention control arm. We were unable to determine whether participants viewed the nutrition education website. However, 2 did not complete the postsurvey with a retention rate of 94% (Figure 2).

XSL•FO
RenderX

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.



Quantitative: Child and Parent Characteristics

The sociodemographic and clinical data by study arm assignment reported by parents demonstrated that most children belonged to the 11 - 12 age group with 44% and 50% in the HPVVaxFacts and nutrition study arms, respectively. Children in both groups were predominantly female, non-Hispanic White, and heterosexual. Among parents who participated in the survey, most of them were heterosexual. The mean age for participants in the HPVVaxFacts intervention arm was 41.8 (SD 14.6) years compared to participants in the nutrition attention control arm at 41.6 (SD 7.5) years. In terms of education level, a higher proportion of the parents in the HPVVaxFacts intervention arm had a bachelor's degree or higher (18/27, 66%) than those in the nutrition attention control arm (11/30, 40%). Additionally, the majority in both arms reported an annual household income of more than US $60,000. In the HPVVaxFacts intervention arm, 59% (16/27) of parents had been offered the HPV vaccine for their child by a provider compared to 70% (21/30) in the nutrition attention control arm. More than half in both arms were iPhone users (Table 1).

XSL•FO
RenderX

Table . Demographic and clinical information of parent and their child (N=57).

Characteristics	HPVVaxFacts program (n=27)	Nutrition program (n=30)
Child age (years), n (%)		
9 - 10	7 (26)	4 (13)
11 - 12	12 (44)	15 (50)
≥13	8 (30)	11 (37)
Child sex, n (%)		
Male	8 (30)	12 (40)
Female	19 (70)	18 (60)
Prefer not to answer	0 (0)	0 (0)
Child sex identification, n (%)		
Straight (heterosexual)	24 (89)	25 (83)
Questioning or unsure	0 (0)	1 (3)
Prefer not to disclose	3 (11)	3 (10)
Additional category or identity not listed	0 (0)	1 (3)
Child race, n (%)		
Non-Hispanic White	19 (70)	24 (80)
African American, African, or Afro Caribbean	5 (19)	3 (10)
Latino or Spanish Origin	1 (4)	0 (0)
American Indian, Alaskan Native, Native American	0 (0)	1 (3)
Native Hawaiian or Other Pacific Islander	0 (0)	0 (0)
Prefer not to answer	1 (4)	1 (3)
Other	1 (4)	1 (3)
Parent age (years), mean (SD)	41.8 (14.6)	41.6 (7.5)
Parent sex, n (%)		
Male	2 (7)	1 (3)
Female	23 (85)	29 (97)
Prefer not to answer	2 (7)	0 (0)
Parent sex identification, n (%)		
Straight (heterosexual)	25 (93)	29 (97)
Questioning or unsure	0 (0)	0 (0)
Prefer not to disclose	1 (4)	1 (3)
Additional category or identity not listed	1 (4)	0 (0)
Parent race, n (%)		
Non-Hispanic White	19 (70)	25 (83)
African American, African, or Afro Caribbean	5 (19)	2 (7)
American Indian or Alaskan Native	0 (0)	0 (0)
Latino or Spanish Origin	1 (4)	2 (7)
Native Hawaiian or Other Pacific Islander	1 (4)	0 (0)
Prefer not to answer	1 (4)	1 (3)
Parent education, n (%)		

XSL•FO
RenderX

Characteristics	HPVVaxFacts program (n=27)	Nutrition program (n=30)
GED[a] or High School Diploma	3 (11)	6 (20)
Associate's degree	5 (19)	8 (27)
Some college	1 (4)	4 (13)
Bachelor's degree	12 (44)	6 (20)
Master's degree	3 (11)	5 (17)
Doctoral or Professional degree	3 (11)	1 (3)
Household income (US $), n (%)		
<$20,000	2 (7)	1 (3)
$20,001 to $40,000	4 (15)	3 (10)
$40,001 to $60,000	4 (15)	3 (10)
$60,001 to $80,000	7 (26)	8 (27)
>$80,000	9 (33)	12 (40)
Did not want to answer	1 (4)	3 (10)
Phone type, n (%)		
Android	9 (33)	8 (27)
iPhone	18 (67)	22 (73)
Offered HPV[b] vaccination for child, n (%)		
Yes	16 (59)	21 (70)
No	11 (41)	9 (30)

[a]GED: General Educational Development.

[b]HPV: human papillomavirus.

HPV Vaccine Initiation Status Poststudy by Study Arm

HPV vaccine status was obtained for 57 children aged 9 - 17 years who were nonvaccinated for HPV prior to their parent participating in the clinical trial. HPV vaccine initiation (ie, those who received the first dose) rates of children whose parents participated in the HPVVaxFacts intervention arm were significantly higher than those of children whose parents were assigned to the nutrition attention control arm (48% vs 17%; Table 2).

Table . HPV[a] vaccine rates 12 months poststudy by study arm (N=57).

	HPVVaxFacts Intervention (n=27)	Nutrition Attention Control (n=30)	Total (N=57)	*P* value
Received >1 dose of HPV vaccine, n (%)				.01
Yes	13 (48)	5 (17)	18 (32)	
No	14 (52)	25 (83)	39 (68)	

[a]HPV: human papillomavirus.

Factors Associated With HPV Vaccination by Study Arm

We determined the impact of the HPVVaxFacts intervention and nutrition attention control on knowledge, concerns, and intention related to HPV vaccination. Analysis revealed that parents in the HPVVaxFacts intervention arm showed higher improvement in knowledge scores but remained largely unchanged in the nutrition attention control arm with no significant difference in knowledge before or after intervention. Similarly, parents in the HPVVaxFacts intervention arm had a larger reduction in concern toward HPV vaccination compared to those in the nutrition attention control arm. However, the difference in concern scores was not significant (Table 3).

XSL•FO
RenderX

Table . Knowledge and concerns by study arm assignment (N=57).

	Knowledge, mean (SD; 95% CI)			Concerns, mean (SD; 95% CI)		
	HPVVaxFacts program	Nutrition program	*P* value	HPVVaxFacts program	Nutrition program	*P* value
Pre	2.1 (0.5; 1.2 to 3.0)	2.1 (0.4; 1.3 to 2.9)	.97	26.6 (1.8; 23.1 to 30.1)	27.3 (1.4; 24.6 to 30.1)	.75
Post	1.4 (0.3; 0.8 to 2.1)	2.2 (0.5; 1.2 to 3.1)	.22	23.2 (1.7; 19.9 to 26.5)	25.9 (1.4; 23.3 to 28.6)	.21
Post-pre	–0.6 (0.4; –1.4 to 0.1)	0.1 (0.2; –0.4 to 0.5)	.13	−3.4 (1.2; −5.8 to −1.0)	−1.4 (0.5; –2.4 to –0.4)	.14

Analysis demonstrated parents had varying degrees of intent (ie, an imperfect surrogate for vaccine hesitancy) to vaccinate their child. Before the intervention, 74% (20/27) of the parents in the HPVVaxFacts intervention arm had the intention of "Very likely" or "Somewhat likely" to get their children vaccinated compared to 60% (18/30) in the nutrition attention control arm. Post intervention, parents who reported intentions (very likely and somewhat likely) were 70% (19/27) in the HPVVaxFacts intervention arm and 63% (19/30) in the nutrition attention control arm. However, when examining pre-post changes in parents' intentions, 3/27 (11%) parents in the HPVVaxFacts intervention arm and 1/30 (3%) parent in the nutrition attention control arm reported an increased likelihood of vaccinating their children. The pre-post intervention arm and between-arm differences were not statistically significant (Table 4).

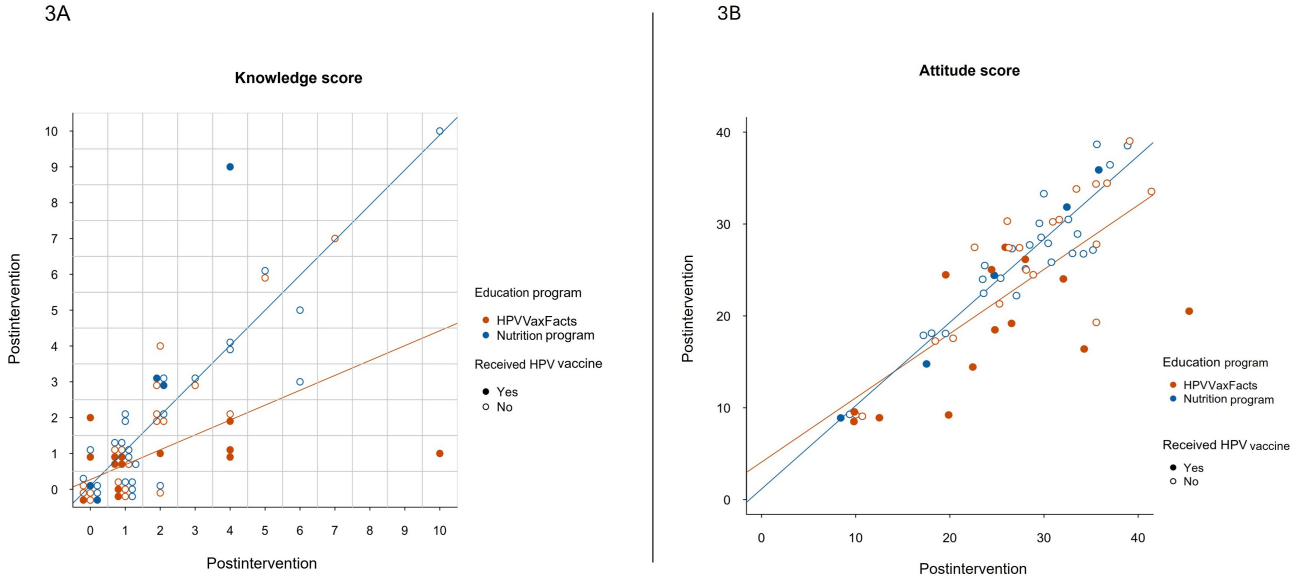
Table . Intention by study arm.

	HPVVaxFacts (n=27)	Nutrition program (n=30)	*P* value
Preintent: HPV[a] vaccine likelihood, n (%)			.56
Definitely not likely	4 (15)	5 (17)	
Not likely	3 (11)	7 (23)	
Somewhat likely	8 (30)	9 (30)	
Very likely	12 (44)	9 (30)	
Postintent: HPV vaccine likelihood, n (%)			.32
Definitely not likely	3 (11)	6 (20)	
Not likely	5 (19)	5 (17)	
Somewhat likely	6 (22)	11 (37)	
Very likely	13 (48)	8 (27)	
Change in intention, n (%)			.51
No change	22 (81)	27 (90)	
Less likely	2 (7)	2 (7)	
More likely	3 (11)	1 (3)	

[a]HPV: human papillomavirus.

XSL•FO
RenderX

Figure 3. Exploratory analyses comparing pre-post scores of human papillomavirus vaccination knowledge and concerns with human papillomavirus vaccination status based on study arm assignment. HPV: human papillomavirus.



Note: The lower the score, the greater the knowledge and lesser concerns related to the HPV vaccination though not significant.

Figures 3A and 3B show an exploratory analysis between pre- and postintervention knowledge and concern scores in both study arms. For knowledge in Figure 3A, the estimated slope was 0.76 (95% CI 0.48-1.04) for HPVVaxFacts intervention arm and 0.98 (95% CI 0.78-1.18) for the nutrition attention control arm, with no evidence that the association differed by study arm ($P_{interaction}$=.21). For concern in Figure 3B, the estimated slope was 0.70 (95% CI 0.52-0.89) for HPVVaxFacts intervention arm and 0.91 (95% CI 0.69-1.12) for the nutrition attention control arm, with no evidence of a differential association between study arms ($P_{interaction}$=.16; Figure 3). In summary, although participants in the HPVVaxFacts intervention arm showed lower postintervention knowledge and concern scores than those in the Nutrition attention control arm, consistent with greater knowledge and fewer concerns based on the scoring of these measures, the differences between the study arms were not significant.

Parental Top HPV Vaccination Concerns: PreSurvey and HPVVaxFacts Data Log

Only the HPVVaxFacts mobile web app allowed parents to take the quiz and confirm or change their top concerns produced via an existing algorithm. Parents in the HPVVaxFacts intervention arm were asked to identify and select their top 3 beliefs or concerns for HPV vaccination on the web app. We determined the frequency with which a concern was selected as a top 3 concern for parents. Based on the web app data, the majority of parents identified the vaccine could cause serious health problems (20/27, 74%), too many vaccines (17/27, 63%), and questioned HPV vaccine safety (12/27, 44%). Refer to Table 5 for a summary of the top 3 concerns selected for the HPVVaxFacts intervention arm participants.

Table . Top 3 selected concerns for HPVVaxFacts intervention arm participants.

Concern	WebApp (HPVVaxFacts[a] program), n (%)
Too many vaccines could cause harm to child	17 (63)
Natural Immunity against HPV[a] is better	7 (26)
HPV vaccine could cause serious health problems	20 (74)
Wonder if HPV vaccine is effective in preventing genital warts and cancer	8 (30)
Wonder if the HPV vaccine is safe	12 (44)
Wonder if child needs HPV vaccine to prevent cancer	6 (22)
Wonder if my child needs the HPV vaccine to prevent genital warts	1 (4)
I wonder if my child should wait until older	6 (22)
Worry that getting vaccine might make my child think it is okay to have sex	3 (11)

[a]HPV: human papillomavirus.

XSL•FO
RenderX

Feasibility: Study Procedures Implementation Into Clinic Workflow

Data Collection and Data Availability

At Clinic Site 1, quantitative data were collected between June and September 2023. At Clinic Site 2, data collection occurred 1 week in August 2023 for back-to-school immunization week.

The data needed for study evaluation were available with less than 5% missing data from the survey, EHR for child's age and vaccine status, and HPVVaxFacts data log tracking of parent concerns and selected top 3 HPV vaccine concerns. In total, 94% of parents completed the pre-post survey. One year post study, vaccination rates were manually pulled again by a clinic staff member to identify that 3 children in the active control group were already vaccinated when the parent enrolled in the study; these children had no record of HPV vaccination upon study enrollment, which reflects interoperability across EHRs. While these parents were excluded, it demonstrated they were interested in learning more about HPV vaccination.

Conceptual and Psychometric Adequacy of Measures

Psychometric measures for concerns, knowledge, and intervention acceptability were considered reliable. Cronbach α was 0.80 and higher, which is considered exceptionally good for internal consistency reliability. See measures in Methods for psychometric results.

Parental Acceptance of HPVVaxFacts

Parents who received the HPVVaxFacts (n=27) had consistently prominent levels of acceptability for the program across all the items. Over two-thirds agreed or strongly agreed that they liked the program (mean 4.1, SD 0.7), the program met their approval (mean 4.2, SD 0.8), and it was appealing (mean 4.2, SD 0.8), and they perceived that parents should use the tool for HPV vaccine education (mean 4.2, SD 0.7). Additionally, parents perceived that HPVVaxFacts was helpful and should be used to discuss the HPV vaccine with their child (mean 4.1, SD 0.7; Table 6).

Table . Intervention acceptability by parents (n=27).

Item	Mean (SD)
Meets approval	4.2 (0.8)
Appealing	4.2 (0.8)
Like program	4.1 (0.7)
Favor of tool being used by parents	4.2 (0.7)
Helpful to discuss HPV[a] vaccine with child	4.1 (0.7)

[a]HPV: human papillomavirus.

Qualitative

A total of 19 parents completed the interviews, 9 from the HPVVaxFacts intervention arm and 10 from the nutrition attention control arm. In the HPVVaxFacts intervention arm, the mean age of the parent was 38.9 (SD 5) years. All participants were female (9/9, 100%) and non-Hispanic White (9/9, 100%). For the nutrition attention control arm, all were female (10/10, 100%), with the majority being non-Hispanic White (7/10, 70%) and almost half had an education level of GED or High School Diploma (4/10, 40%; Table S1 on parent sociodemographics in Multimedia Appendix 2).

Themes

Overview

Based on participant responses, we identified 4 themes related to the study protocol. We identified one theme related to intervention acceptability using participants' responses in the HPVVaxFacts intervention arm only. Each theme is briefly described below.

Theme 1: Study Protocol Acceptability and Recommendations for Improvement

Acceptability

Parents across both study arms perceived the study protocol as acceptable and manageable. They commented on how the clinic staff were welcoming and provided clear study enrollment information. Mobile-based delivery of each study step—consent, enrollment, presurvey, view of nutrition or education program, postsurvey—was intuitive for participants, especially those in the nutrition education arm. One nutrition program parent stated,

> *So, with that said, it [steps of study protocol] was fairly easy to navigate, to read through, and to get it completed, and to gain that information in a short amount of time, um, despite what I was doing.*

However, receiving multiple texts for each study step posed a challenge, particularly for those in the HPVVaxFacts intervention arm. A few parents noted how clinic staff availability was helpful when questions arose. A parent in the HPVVaxFacts intervention arm stated,

> *The only hiccup I remember having was like, I feel like, to get from point A to point Z in the process, like it required the coordinator coming in and helping me get to where I needed to go.*

Finally, parents appreciated receiving education prior to their clinic visit, which allowed them to engage with study materials without the time constraints of an in-clinic visit appointment.

Parents in the HPVVaxFacts intervention arm did not use all the mobile web app features. Three stated they did not use the videos, citing personal preference. Four did not access the kid's corner, perceiving their child would not understand, was too

XSL•FO
RenderX

young, or lacked time to view or discuss with the child. A parent in the HPVVaxFacts intervention arm stated,

> *Maybe when they got a little bit older, I would, but my son would, it would go right over his head. Um, but, uh, my daughter, now that she's like 13, I hopefully would share it with her.*

Parents in the nutrition attention control arm found the nutrition webpage easy to navigate and comprehensible. One parent highlighted the survey was simple, and the webpage provided immediate feedback with explanations for both correct and incorrect answers.

Recommendations

To optimize the study protocol, parents from both study arms identified the need to improve logistics, particularly reducing the number of text messages containing links for each step. Regarding content delivery timing, parents across study arms acknowledged their busy schedule noted the night before or day of (at home or in clinic before the visit) is an ideal time to review materials. A parent in the HPVVaxFacts intervention arm stated:

> *I think the day before would be good just because then they have the information to make an informed decision at the actual doctor's visit.*

However, one HPVVaxFacts parent offered a different perspective, noting an optimal window of 2 weeks before the clinic visits to get more informed about the HPV vaccine. She stated,

> *I would say no more than two weeks before. That's kind of what we have found to be the sweet spot when we give deadlines for something because if you give them too much time, they're gonna forget about it because they're like, Oh, I've got plenty of time. And then if you don't give them enough time, I mean, you know, everybody's busy and it's like, well, I've got two days and I plan I can't get this done in two days.*

Other suggestions related to expanding the study protocol to address all vaccines and identify alternative content delivery mechanisms. A parent in the nutrition attention control arm stated the study protocol should address all vaccines.

> *My honest opinion would be that every vaccine needs to be treated this way, because a lot of emphasis is given into this one vaccine…If parents are given more information, I mean if a parent wants to educate themselves and they take the initiative, they educate themselves on all the vaccines.*

Another parent in the nutrition attention control arm highlighted,

> *I think maybe having the option of either you can get it beforehand on the phone or maybe you have it there for the provider to give to a parent to look over… I think that way you just kind of open it up to, you kind of cater to everyone that way because not everybody's tech savvy and wants it on their phone. Some people do still like to have a piece of paper to hold and look at.*

Theme 2: Acceptability of and Recommendations to Improve HPVVaxFacts

Acceptability

Parents who participated in the HPVVaxFacts intervention arm perceived the program as user-friendly, and the layout was easy to navigate. One parent stated,

> *I mean, I thought the whole thing was easy to use. I mean, I texted back our primary care. They got us set up with the links and everything was very easy.*

One even appreciated that the program asked targeted questions to narrow parental concerns and delivered tailored content. Parents found the content comprehensible and accessible, particularly on the mobile phone.

> *I thought it was good information. I liked that it was easily accessible and pretty quick to read through and understand.*

Another parent noted,

I think everything was pretty, pretty easy to understand and follow.

However, program completion was unclear for 2 parents. One parent stated,

> *A little unclear, like, I even messaged him and I was like, I think I finished everything, but I'm not sure that I did. And they were like, we think you did too, but we're not sure either.*

Recommendations

Parents suggested improving the web app and adding parent testimonials to demonstrate how others made vaccination decisions. One parent stated,

> *I think it's always good, especially for people who are, who have reservations to see. Stories of others who say, Hey, you know, I was really on the fence about this because I thought this, but, you know, I got this information and it may be understand that that wasn't the case because of whatever.*

Parents also requested more detailed information on "extreme" side effects and misinformation. One parent preferred links to research articles over the videos that discussed each concern.

Two parents highlighted the need for the app to be multilingual for greater reach. Parents wanted a reminder via text or email from providers to use the mobile web app to help prepare questions in advance. A few parents suggested increasing mobile web app availability for use by health care agencies, schools, and provider-led public service announcements (PSAs) to extend reach to parents. A parent stated,

> *I mean, my initial reaction is things like schools, but I know you have to be so careful with what kind of stuff you put out in schools nowadays.*

Another parent recommended,

> *If it became if it was, like, adopted by the state, the health department would be a really good channel,*

XSL•FO
RenderX

> *like, maybe the website through the health department.*

Two parents recommended that this program should be applied across all vaccines for parents to understand the benefits of vaccination. Specifically, the parent stated,

> *I think that if they had it for more vaccines that people would have a better understanding as to what it is that they are subjecting their child to by having the vaccinations, the benefits.*

Theme 3: Perceived Benefits of HPVVaxFacts

Parents described 4 benefits from participating in the HPVVaxFacts intervention arm. We summarize these benefits below.

Confirmed or Improved Knowledge for Decision-Making

Parents commonly reported that HPVVaxFacts improved or reaffirmed their knowledge of HPV and the vaccine, regardless of their baseline knowledge level. One parent stated,

> *It reaffirmed kind of what I already knew, and it did give me additional information that I might not necessarily have had or even have thought to ask the doctor about.*

Another parent described,

> *I thought it was very informative. A lot of the stuff I already knew, but I could see how, especially with new parents or parents who aren't as hypervigilant, could benefit from having that information if they didn't already know it.*

Some parents felt better equipped to make an informed decision after reviewing the program. A parent stated,

> *I feel like the information was clear and straightforward, and I think it was presented in a way to help parents more understand the vaccine and make an informed decision.*

Preclinic Information Improves Appointment Efficiency

All parents valued receiving HPV vaccination information preclinic visit and found it feasible. The majority of parents had adequate time to review the information and form a decision. A parent stated,

> *I liked having it ahead of time because then by the time I got in there, I had all the information that I needed in order to make the decision on whether or not we were going to do it.*

Several parents noted that information prior to the clinic visit allowed appointments to be more efficient with fewer questions. One parent emphasized,

> *I think that with the program, it gave me all the information that was needed to be able to just make the appointment smoother and just be able to like go in there and have our doctor's appointment and not have to worry so much about any of the questions. Because I felt like it was there for me to be able to know all the things before, I got there.*

Other parents noted HPVVaxFacts helped facilitate the discussion with the child's provider. A parent stated,

> *Honestly, it makes the appointment go faster. I mean, I know that the doctors are busy, so it's good to just have before so you just know what, what exactly it is before you, you know, you end up doing it if you do decide to do it or whatever.*

However, 2 parents wreported no provider recommendation or discussion on the HPV vaccination during the clinic visit, making previsit information particularly helpful in their decision-making.

Trust in Provider-Delivered Information

All parents appreciated that the program was delivered by their child's provider, who was deemed a trusted information source. A parent stated,

> *Yeah, I think it [HPVVaxFacts] was provided by his pediatrician. That's probably the best thing about it. In my opinion.*

Parents acknowledged widespread misinformation on vaccines and identified how this app could address those concerns. One parent stated,

> *Yeah, I definitely trust information provided through his pediatrician more so than just something you find online because I feel like it's easy to find supporting information for your bias ideas online. I can find something from the pediatrician. I'm going to trust that further than whatever I research on the Internet.*

Furthermore, they valued the credibility of provider-sourced information. One parent lauded,

> *There is way too much out there, and you don't know what you can trust and what you can't trust. Um, so definitely coming from a clinic and a trusted place is very comforting.*

Facilitates Parent-Child Discussion on HPV Vaccine

One parent used the kid corner to guide discussion with her child, while another found the information "lay" enough to discuss the HPV vaccine with her child. A parent shared,

> *I remember, thinking that it had good information for me to be able to explain it to my children in a way that they would understand better.*

Parents valued the program to facilitate parent-child conversations on vaccines such as HPV, with some preferring to use their older children who could benefit from these open discussions. A parent stated,

> *I think for older children, especially and where they are having conversations that are more open with their children about vaccines.*

Theme 4: Perceived Barriers to HPVVaxFacts Use

Most parents reported no barriers to using the program. A parent stated,

> *"I don't think so. I don't recall having anything that was, that I would consider a barrier."*

XSL•FO
RenderX

Three parents highlighted potential barriers that were not program-specific. One barrier was that some parents may lack access to the mobile web app. Another barrier older or not technologically "savvy" parents may have difficulty with web app navigation. A parent stated,

> *That [navigating a technology] would be a little bit more difficult, like later 50s, because I know I'm I had my daughter when I was 22, but I know a lot of parents aren't having them until like they're in their, you know, early to late 30s.*

Finally, one parent noted that some parents may not trust their health care provider and the research team or the information provided. It was further recommended to use alternative distribution channels (eg, media, social media, PSA) to reach these individuals.

Clinic Staff Consultation: Provider Feedback

Overall, providers perceived the program as beneficial for their clinic. Particularly, the top benefit was offering preclinic education visits to their patients. This complemented their recommendations and lessened clinic time for education. Another perceived benefit was the increased vaccination rates in the clinic because of the intervention. Regarding barriers, providers and clinic staff highlighted the need to improve logistics for large-scale implementation. Because the intervention was manual and required efforts from clinic staff and the research team to aid in recruitment and retention, providers questioned the feasibility of recruitment and retention procedures but still perceived offering preclinic information in general as feasible. However, providers and staff stated they understood this was a pilot study and hoped their suggestions improved future implementation.

Discussion

Principal Findings

Our study assessed feasibility, acceptability, and preliminary efficacy of the theory-informed, mobile web app HPVVaxFacts designed to decrease concerns and increase knowledge, intentions, and uptake of the vaccine among children aged 9 to 17 years. This study is one of the first to use a multimethod evaluation to assess mHealth interventions to increase HPV vaccination. More importantly, HPVVaxFacts is one of the first mobile web apps to identify and tailor information to parental concerns while supporting autonomy in both parental decision-making and educating their child using an optional kid's corner on the web app.

Feasibility

Our study reported feasibility across multiple metrics. The enrollment rate was 84%, and the random allocation achieved 100% recruitment. The retention rates in both study arms were 75% in the HPV program and 94% in the nutrition program. Notably, approximately 4 out of 5 (81%) parents in the intervention arm viewed HPVVaxFacts. These rates are considered good, especially with the sensitive context surrounding HPV vaccination and typical survey response rates. The engagement of clinic staff and parents in the development of the study protocol, a known factor in improving study outcomes (eg, logistics and behavior) [51], partly attributes to this success.

Additionally, the Cronbach α for concerns, knowledge, and intervention acceptability scales were considered reliable, further supporting study feasibility and potential deployment of these scales in a larger study. However, because one clinic had vaccination rates substantially increase prior to study implementation, this underscores the importance of conducting a feasibility assessment at multiple checkpoints before and during study implementation. Furthermore, both providers and participants highlighted the logistical barrier of study steps being manual, a factor which may have affected response rates at different steps in the research process. This supports the need to automate these steps for future efficacy trials to improve recruitment and retention.

Interestingly, parents identified factors that could influence use of HPVVaxFacts long-term in research and in clinical practice. For protocol application, improving the study logistics was commonly mentioned by both parents and providers, though not reflective of app functionality. This suggests that the tool alone could be applied in real-world settings with minimal human involvement, a major advantage for sustainability. Similar to past studies [52], parents identified limited access to technology (eg, mobile web app) as a potential barrier reflecting the "digital divide" in America [53]. Some parents, particularly older parents or non–English-speaking parents, may struggle to navigate the mobile web app, which has been identified as a barrier previously [54]. Some parents may mistrust a provider who delivered resources because of broader mistrust in health care. This demonstrates how mistrust can extend beyond the individual to tangibles perceived to be associated with the individual [55]. These findings suggest that feasibility of HPVVaxFacts can be improved using multilingual languages and addressing the digital divide (eg, physical infrastructure expansion, training, and tech support). Also, alternative strategies for providing health care information should be considered, including multiple dissemination channels and provider-led initiatives to build support and trust.

Acceptability

Acceptability was considered good with a rate of 74% or higher with room for improvement. Of interest, acceptability of HPVVaxFacts centered around the ability of the app to tailor to the parents' needs. Particularly, parents were able to access the app when available (previsit information), identify their top vaccine concerns, and view the modality that best served their need to get their concerns addressed. This suggests the importance of a multilevel, tailored approach that is cost-effective to target parent decision-making processes for vaccination, creating a unique opportunity for this tool to have broad reach. Other acceptable components of the mobile web app, though deemed unsurprising, were its comprehensible content, ability to navigate, and being time-friendly for use. Previsit information has been supported by past studies [26], but its use in the context of a mobile app has been underexplored. Parents lauded it for reasons such as lessened time in clinic or improved preparation for the visit, an acceptable

XSL•FO
RenderX

strategy we add to the literature. Our findings also support that increased acceptability of mobile web apps for parent-child discussion could be achieved via provision of parental autonomy to share select information or allow parents to use information to guide a discussion with the child without direct access. This should be further explored as mHealth continues to evolve to expand the reach of health information for caregivers and/or their children.

Not only did parents view HPVVaxFacts to improve feasibility through content being multilingual, but it was also viewed to increase acceptability. Particularly, the suggestion to expand this app content to different languages and its use in alternative settings that provide vaccination to increase acceptability reflects the altruism of these parents. Walker et al [56] found that parents with higher interpersonal altruism had higher odds of agreeing with HPV-specific state vaccine requirements, a confirmation of our qualitative findings. These findings suggest the need for inclusivity when developing strategies and tools that promote HPV vaccination among children. Exploring the app's use in alternative immunization settings supports the HPV National Roundtables' goal to increase HPV vaccination in these settings.

Preliminary Efficacy

Similar to other mHealth studies [26], HPVVaxFacts increased HPV vaccine initiation rates among children at significantly higher rates compared to the nutrition program. Interestingly, this increase occurred despite no significant differences in knowledge, concerns, or intentions between the study arms. Because knowledge and attitudes had a positive trajectory through the increase in effect size in the intervention arm compared to the control arm, this could reflect the pilot study being underpowered for these secondary outcomes. This is a common outcome, as many pilot studies are powered for testing feasibility and not proving an intervention's efficacy [57]. Also, it remains unclear whether, and to what extent, this effect was mediated through provider engagement or indirect engagement via app delivery through the EHR. Nevertheless, our qualitative findings support that parental willingness to get vaccinated was due to receipt of the information from the provider. Because parents' most trusted information source [55,58] is providers and their recommendation is critical for uptake [59], future studies should explore in-depth whether and how provider communication modalities can be leveraged to complement provider recommendations and communication in-clinic with a goal to increase HPV vaccination compliance. Finally, these findings could also reflect a ceiling effect in which a significant portion of participants in both the HPV and nutrition programs already had positive intentions to vaccinate [60]. Future studies could either recruit parents who are "not likely" or "somewhat likely" only or use baseline status to stratify analysis, both strategies to better capture intervention effects [61]. Finally, it was not surprising that the top 3 concerns identified by these parents centered around vaccine safety, the leading barrier to vaccine hesitancy and suboptimal rates [62].

Limitations

First, this study was conducted in 2 pediatric urban clinics in Tennessee in which the primary clinic for recruitment was predominantly female and non-Hispanic White. Collectively, this may limit generalizability to other populations, including those living in different geographical regions and settings along with racial and ethnic groups. Second, selection bias may have occurred if some parents with low mobile app literacy or who did not have a mobile app did not participate, which should be explored in future studies. However, we have begun to take steps to improve inclusion through the provision of HPVVaxFacts in Spanish. Third, this study had a small sample size, which limited our ability to detect a difference for some of the study outcomes (eg, concerns and intention) and whether those outcomes are a result of clinic demographics and protocol. Also, the small number of parents recruited at Clinic Site 2 made it impossible to assess whether clinic-level differences influenced results. Fourth, there were updates on the app during the feasibility study; however, study implementation metrics, including recruitment and retention, improved over time. Fifth, because parents agreed to be in the study and the invite mentioned HPV vaccination, parents may be more engaged or interested in receiving information. This should be explored in a larger study.

Sixth, some parents were lost due to their child already being vaccinated during study participation and study logistics, in which we could not recruit additional participants due to the study duration ending at both clinics. However, this reflects EHR interoperability, a commonly cited issue confirmed by clinic staff [63]. Seventh, we only studied the impact of HPVVaxFacts on initiation rates. The impact on completion rates (ie, 2 doses of HPV vaccine among children younger than 15 years old and 3 doses among those older than 15 years old) [64] will be explored in a future study, as we are exploring reasons for noncompliance and informational needs of parents to inform the update of messaging on the mobile web app to be more comprehensive. Eighth, for the interviews, there may be potential for recall bias, which increases due to the variable time between parent participation and the interview due to demanding study logistics. Ninth, some mobile web app components (ie, average session length) could not be tracked due to financial constraints. However, it is currently being updated to address this limitation. Tenth, small discrepancy between the HPVVaxFacts website content and knowledge scale could have led to a lack of significant improvement in knowledge. We were also unable to track participants' engagement within the website that may have varied from full review to partial or no interaction with the content. It is possible that the nutrition education within the attention control arm could have influenced parents to accept and/or obtain the first dose of the HPV vaccination for their child. Finally, the nutrition attention control arm did not select their top concerns as this was not an option on the app. Therefore, we are unable to compare parental top concerns by study arm.

Conclusion

Our pilot study suggests HPVVaxFacts is feasible and acceptable to provide tailored information to parents' preclinic visits to aid parent decision-making for HPV vaccination and ultimately increase uptake. This study's outcomes, both unclear yet offer a positive trajectory, lay the foundation for a full RCT to test the efficacy of HPVVaxFacts being offered preclinic visit to increase HPV vaccine initiation and completion rates.

XSL•FO
RenderX

Future studies can explore the mediating (eg, provider recommendation and communication) and moderating factors for successful intervention effects with a large sample size. Findings also reinforce that provider modalities such as mHealth are acceptable and can complement provider communication to address parental HPV vaccine concerns.

Acknowledgments

The authors would like to thank the participating clinics who offered their insights on how to enhance the study protocol and their participation in the study recruitment and retention. We also thank the parents who took part in the study, from either informing the study protocol or being a participant in the study. Collectively, this study's success can be attributed to these individuals. The authors would also like to thank 233 Analytics, experts who provide complex technical litigation and corporations, who developed and maintained the intervention. AI was not used to generate manuscript content.

Funding

This work was supported by the National Cancer Institute of the National Institutes of Health (NIH; award Number: K01CA237748-01), and the National Center for Advancing Translational Sciences (NCATS) Clinical Translational Science Award (CTSA; award number: 5UL1TR0002243). The funding sponsors had no role in the design and conduct of the study, collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript or decision to submit the manuscript for publication. The contents of this article are solely the responsibility of the authors and do not necessarily represent those of the National Institutes of Health.

Data Availability

Due to confidentiality agreements, supporting data cannot be made openly available.

Authors' Contributions

Conceptualization: JC-E, CW, PH, AD
Methodology: JC-E, CW, PH, AD
Funding acquisition: JC-E, CW, PH, AD
Supervision: JC-E
Validation: JC-E
Resources: JC-E, LM
Data curation: JC-E, ES, JJ, MS, FM, TVM, LH, MH, JC
Investigation: JC-E, ES, JJ, MS, FM, TVM, LH, MH, JC
Formal analysis: JC-E, DL, JJ, TK, LS, MC, TB
Visualization: JC-E, DL, JJ, TK, LS, MC, TB
Writing – original draft: JC-E, MC, TK, LS
Writing – review & editing: All authors

Conflicts of Interest

JC-E is an education consultant for MERCK not related to vaccination. PCH has received research support from Merck Foundation and Eli Lilly and Company that is not related to the HPV vaccine. The other authors have no relationships to declare. JC-E is the owner of the web app HPVVaxFacts.

Multimedia Appendix 1
HPVVaxFacts pilot, feasibility trial CONSORT.
[DOCX File, 34 KB - cancer_v12i1e78910_app1.docx]

Multimedia Appendix 2
Qualitative interviewee sociodemographics (n=19).
[DOCX File, 16 KB - cancer_v12i1e78910_app2.docx]

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 1193 KB - cancer_v12i1e78910_app3.pdf]

References

XSL•FO
RenderX

1. Szilagyi PG, Albertin CS, Gurfinkel D, et al. Prevalence and characteristics of HPV vaccine hesitancy among parents of adolescents across the US. Vaccine 2020 Aug 27;38(38):6027-6037. [doi: 10.1016/j.vaccine.2020.06.074] [Medline: 32758380]
2. Rositch AF, Liu T, Chao C, Moran M, Beavis AL. Levels of parental human papillomavirus vaccine hesitancy and their reasons for not intending to vaccinate: insights from the 2019 National Immunization Survey-Teen. J Adolesc Health 2022 Jul;71(1):39-46. [doi: 10.1016/j.jadohealth.2022.01.223] [Medline: 35279361]
3. Beavis AL, Meek K, Moran MB, Fleszar L, Adler S, Rositch AF. Exploring HPV vaccine hesitant parents' perspectives on decision-making and motivators for vaccination. Vaccine X 2022 Dec;12:100231. [doi: 10.1016/j.jvacx.2022.100231] [Medline: 36337836]
4. Pingali C, Yankey D, Elam-Evans LD, et al. Vaccination coverage among adolescents aged 13–17 years — National Immunization Survey–Teen, United States, 2022. MMWR Morb Mortal Wkly Rep 2023 Aug 25;72(34):912-919. [doi: 10.15585/mmwr.mm7234a3] [Medline: 37616185]
5. Della Fera AN, Warburton A, Coursey TL, Khurana S, McBride AA. Persistent human papillomavirus infection. Viruses 2021 Feb 20;13(2):321. [doi: 10.3390/v13020321] [Medline: 33672465]
6. Viens LJ, Henley SJ, Watson M, et al. Human papillomavirus-associated cancers - United States, 2008-2012. MMWR Morb Mortal Wkly Rep 2016 Jul 8;65(26):661-666. [doi: 10.15585/mmwr.mm6526a1] [Medline: 27387669]
7. Gilkey MB, Calo WA, Marciniak MW, Brewer NT. Parents who refuse or delay HPV vaccine: differences in vaccination behavior, beliefs, and clinical communication preferences. Hum Vaccin Immunother 2017 Mar 4;13(3):680-686. [doi: 10.1080/21645515.2016.1247134] [Medline: 27763818]
8. HPV and cancer. National Cancer Institute. 2019 Mar 1. URL: https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/hpv-and-cancer [accessed 2026-07-15]
9. Cooper S, Schmidt BM, Ryan J, et al. Factors that influence acceptance of human papillomavirus (HPV) vaccination for adolescents: a qualitative evidence synthesis. Cochrane Database Syst Rev 2019(9):9. [doi: 10.1002/14651858.CD013430]
10. Adjei Boakye E, Nair M, Abouelella DK, et al. Trends in reasons for human papillomavirus vaccine hesitancy: 2010-2020. Pediatrics 2023 Jun 1;151(6):e2022060410. [doi: 10.1542/peds.2022-060410] [Medline: 37218460]
11. Cunningham-Erves J, Smalls M, Stewart EC, et al. Communication interactions, needs, and preferences during clinical encounters of African American parent-child dyads. J Racial Ethn Health Disparities 2024 Oct;11(5):2958-2972. [doi: 10.1007/s40615-023-01754-8] [Medline: 37603225]
12. Hamlish T, Clarke L, Alexander KA. Barriers to HPV immunization for African American adolescent females. Vaccine 2012 Oct 5;30(45):6472-6476. [doi: 10.1016/j.vaccine.2012.07.085] [Medline: 22910288]
13. Morales-Campos DY, Markham CM, Peskin MF, Fernandez ME. Hispanic mothers' and high school girls' perceptions of cervical cancer, human papilloma virus, and the human papilloma virus vaccine. J Adolesc Health 2013 May;52(5 Suppl):S69-S75. [doi: 10.1016/j.jadohealth.2012.09.020] [Medline: 23601613]
14. Fishman J, Taylor L, Kooker P, Frank I. Parent and adolescent knowledge of HPV and subsequent vaccination. Pediatrics 2014 Oct;134(4):e1049-e1056. [doi: 10.1542/peds.2013-3454] [Medline: 25225141]
15. Ten threats to global health in 2019. WHO. 2024. URL: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [accessed 2024-02-01]
16. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283. [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
17. Aapro M, Bossi P, Dasari A, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct;28(10):4589-4612. [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
18. Gibson DG, Kagucia EW, Ochieng B, et al. The Mobile Solutions for Immunization (M-SIMU) Trial: a protocol for a cluster randomized controlled trial that assesses the impact of mobile phone delivered reminders and travel subsidies to improve childhood immunization coverage rates and timeliness in Western Kenya. JMIR Res Protoc 2016 May 17;5(2):e72. [doi: 10.2196/resprot.5030] [Medline: 27189422]
19. Fadda M, Galimberti E, Fiordelli M, Romanò L, Zanetti A, Schulz PJ. Effectiveness of a smartphone app to increase parents' knowledge and empowerment in the MMR vaccination decision: a randomized controlled trial. Hum Vaccin Immunother 2017 Nov 2;13(11):2512-2521. [doi: 10.1080/21645515.2017.1360456] [Medline: 29125783]
20. Wilson K, Atkinson KM, Westeinde J. Apps for immunization: leveraging mobile devices to place the individual at the center of care. Hum Vaccin Immunother 2015;11(10):2395-2399. [doi: 10.1080/21645515.2015.1057362] [Medline: 26110351]
21. Shourie S, Jackson C, Cheater FM, et al. A cluster randomised controlled trial of a web based decision aid to support parents' decisions about their child's Measles Mumps and Rubella (MMR) vaccination. Vaccine 2013 Dec 5;31(50):6003-6010. [doi: 10.1016/j.vaccine.2013.10.025] [Medline: 24148574]
22. Tubeuf S, Edlin R, Shourie S, Cheater FM, Bekker H, Jackson C. Cost effectiveness of a web-based decision aid for parents deciding about MMR vaccination: a three-arm cluster randomised controlled trial in primary care. Br J Gen Pract 2014 Aug;64(625):e493-e499. [doi: 10.3399/bjgp14X680977] [Medline: 25071062]
23. Choi J, Tamí-Maury I, Cuccaro P, Kim S, Markham C. Digital health interventions to improve adolescent HPV vaccination: a systematic review. Vaccines (Basel) 2023 Jan 22;11(2):249. [doi: 10.3390/vaccines11020249] [Medline: 36851127]

XSL•FO
RenderX

24. Langford AT, Solid CA, Scott E, et al. Mobile phone ownership, health apps, and tablet use in US adults with a self-reported history of hypertension: cross-sectional study. JMIR Mhealth Uhealth 2019 Jan 14;7(1):e12228. [doi: 10.2196/12228] [Medline: 31344667]
25. Demographics of mobile device ownership and adoption in the united states. Pew Research Center. 2024. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2024-02-01]
26. Woodall WG, Zimet G, Kong A, et al. Vacteens.org: a mobile web app to improve HPV vaccine uptake. Front Digit Health 2021;3:693688. [doi: 10.3389/fdgth.2021.693688] [Medline: 34713171]
27. Ilozumba O, Schmidt P, Ket JCF, Jaspers M. Can mHealth interventions contribute to increased HPV vaccination uptake? A systematic review. Prev Med Rep 2021 Mar;21:101289. [doi: 10.1016/j.pmedr.2020.101289] [Medline: 33425667]
28. Teitelman AM, Gregory EF, Jayasinghe J, et al. Vaccipack, a mobile app to promote human papillomavirus vaccine uptake among adolescents aged 11 to 14 years: development and usability study. JMIR Nurs 2020;3(1):e19503. [doi: 10.2196/19503] [Medline: 34345789]
29. Shegog R, Savas LS, Healy CM, et al. AVPCancerFree: impact of a digital behavior change intervention on parental HPV vaccine –related perceptions and behaviors. Hum Vaccin Immunother 2022 Nov 30;18(5):2087430. [doi: 10.1080/21645515.2022.2087430] [Medline: 35699953]
30. Sonawane K, Zhu Y, Montealegre JR, et al. Parental intent to initiate and complete the human papillomavirus vaccine series in the USA: a nationwide, cross-sectional survey. Lancet Public Health 2020 Sep;5(9):e484-e492. [doi: 10.1016/S2468-2667(20)30139-0] [Medline: 32707126]
31. Fontenot HB, White BP, Rosenberger JG, et al. Mobile app strategy to facilitate human papillomavirus vaccination among young men who have sex with men: pilot intervention study. J Med Internet Res 2020 Nov 4;22(11):e22878. [doi: 10.2196/22878] [Medline: 33146621]
32. MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: Definition, scope and determinants. Vaccine 2015 Aug 14;33(34):4161-4164. [doi: 10.1016/j.vaccine.2015.04.036] [Medline: 25896383]
33. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg 2012;10(1):28-55. [doi: 10.1016/j.ijsu.2011.10.001] [Medline: 22036893]
34. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
35. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
36. Cunningham-Erves J, Wilkins CH, Dempsey AF, et al. Development of a tailored mobile phone-based intervention to facilitate parent-child communication and build human papillomavirus vaccine confidence: formative qualitative study. JMIR Form Res 2023 Apr 4;7:e43041. [doi: 10.2196/43041] [Medline: 37014680]
37. Gilkey MB, Magnus BE, Reiter PL, McRee AL, Dempsey AF, Brewer NT. The Vaccination Confidence Scale: a brief measure of parents' vaccination beliefs. Vaccine 2014 Oct 29;32(47):6259-6265. [doi: 10.1016/j.vaccine.2014.09.007] [Medline: 25258098]
38. Nutrition quiz. Advantage4Parents. URL: https://advantage4parents.com/quizzes/nutrition-quiz/ [accessed 2024-12-26]
39. Aycock DM, Hayat MJ, Helvig A, Dunbar SB, Clark PC. Essential considerations in developing attention control groups in behavioral research. Res Nurs Health 2018 Jun;41(3):320-328. [doi: 10.1002/nur.21870] [Medline: 29906317]
40. Angriman F, Masse MH, Adhikari NKJ. Defining standard of practice: pros and cons of the usual care arm. Curr Opin Crit Care 2019 Oct;25(5):498-504. [doi: 10.1097/MCC.0000000000000642] [Medline: 31335382]
41. Phreesia. URL: https://www.phreesia.com [accessed 2025-12-09]
42. Cocks K, Torgerson DJ. Sample size calculations for pilot randomized trials: a confidence interval approach. J Clin Epidemiol 2013 Feb;66(2):197-201. [doi: 10.1016/j.jclinepi.2012.09.002] [Medline: 23195919]
43. Moore CG, Carter RE, Nietert PJ, Stewart PW. Recommendations for planning pilot studies in clinical and translational research. Clin Transl Sci 2011 Oct;4(5):332-337. [doi: 10.1111/j.1752-8062.2011.00347.x] [Medline: 22029804]
44. Rehman AM, Ferrand R, Allen E, Simms V, McHugh G, Weiss HA. Exclusion of enrolled participants in randomised controlled trials: what to do with ineligible participants? BMJ Open 2020 Dec 2;10(12):e039546. [doi: 10.1136/bmjopen-2020-039546] [Medline: 33268410]
45. Cunningham-Erves J, Talbott LL, O'Neal MR, Ivankova NV, Wallston KA. Development of a theory-based, sociocultural instrument to assess Black maternal intentions to vaccinate their daughters aged 9 to 12 against HPV. J Canc Educ 2016 Sep;31(3):514-521. [doi: 10.1007/s13187-015-0867-3]
46. Reno JE, Sevick C, Maertens J, Dempsey AF. Is tailored messaging more effective? An analysis of a digital health intervention to promote HPV vaccination intent among Latinx. J Behav Med 2023 Apr;46(1-2):335-345. [doi: 10.1007/s10865-022-00340-3] [Medline: 35927547]
47. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022 Jan 1;60(1):95-103. [doi: 10.1097/MLR.0000000000001664] [Medline: 34812790]
48. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]

XSL•FO
RenderX

49. Chang RS, Shing JZ, Erves JC, et al. Measurement of provider fidelity to immunization guidelines: a mixed-methods study on the feasibility of documenting patient refusals of the human papillomavirus vaccine. BMC Med Inform Decis Mak 2022 Dec 22;22(1):339. [doi: 10.1186/s12911-022-02083-2] [Medline: 36550466]
50. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
51. Cohn E, McCloskey DJ, Berman L, et al. Principles of Community Engagement, 3rd edition: Agency for Toxic Substances and Disease Registry (U.S); 2025. URL: https://stacks.cdc.gov/view/cdc/251390 [accessed 2026-03-02]
52. Lavallee DC, Lee JR, Austin E, et al. mHealth and patient generated health data: stakeholder perspectives on opportunities and barriers for transforming healthcare. Mhealth 2020;6:8. [doi: 10.21037/mhealth.2019.09.17] [Medline: 32190619]
53. Sanders CK, Scanlon E. The digital divide is a human rights issue: advancing social inclusion through social work advocacy. J Hum Rights Soc Work 2021;6(2):130-143. [doi: 10.1007/s41134-020-00147-9] [Medline: 33758780]
54. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 4;8:1687. [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
55. Harrington N, Chen Y, O'Reilly AM, Fang CY. The role of trust in HPV vaccine uptake among racial and ethnic minorities in the United States: a narrative review. AIMS Public Health 2021;8(2):352-368. [doi: 10.3934/publichealth.2021027] [Medline: 34017897]
56. Walker KK, Kasting ML, Head KJ, Sturm L, Zimet GD. Indiana parental perceptions of the acceptability of general and HPV-specific state vaccine requirements. J Community Health 2023 Jun;48(3):528-538. [doi: 10.1007/s10900-023-01198-5] [Medline: 36745356]
57. Donald G. A brief summary of pilot and feasibility studies: exploring terminology, aims, and methods. Eur J Integr Med 2018 Dec;24:65-70. [doi: 10.1016/j.eujim.2018.10.017]
58. Fu LY, Zimet GD, Latkin CA, Joseph JG. Associations of trust and healthcare provider advice with HPV vaccine acceptance among African American parents. Vaccine 2017 Feb 1;35(5):802-807. [doi: 10.1016/j.vaccine.2016.12.045] [Medline: 28063706]
59. Ylitalo KR, Lee H, Mehta NK. Health care provider recommendation, human papillomavirus vaccination, and race/ethnicity in the US National Immunization Survey. Am J Public Health 2013 Jan;103(1):164-169. [doi: 10.2105/AJPH.2011.300600] [Medline: 22698055]
60. Andrade C. The ceiling effect, the floor effect, and the importance of active and placebo control arms in randomized controlled trials of an investigational drug. Indian J Psychol Med 2021 Jul;43(4):360-361. [doi: 10.1177/02537176211021280] [Medline: 34385732]
61. Tetmeyer L, Farkas N, Steinbrink GM, et al. Feasibility, acceptability, and preliminary efficacy of a remotely-delivered health coaching intervention for young and middle-aged cancer survivors: a phase II pilot study. J Cancer Educ 2025 Nov 6. [doi: 10.1007/s13187-025-02770-4] [Medline: 41196510]
62. Sonawane K, Lin YY, Damgacioglu H, et al. Trends in human papillomavirus vaccine safety concerns and adverse event reporting in the United States. JAMA Netw Open 2021 Sep 1;4(9):e2124502. [doi: 10.1001/jamanetworkopen.2021.24502] [Medline: 34533574]
63. Lutz O, Feld S, Winters A. Examining health care professionals' views on using electronic health records and immunization information systems to increase HPV vaccination. : National HPV Vaccination Roundtable; 2017 URL: https://hpvroundtable.org/wp-content/uploads/2017/09/EHR.IIS_.National.HPV_.Vax_.Roundtable.Report.Final_.pdf [accessed 2026-07-15]
64. HPV vaccine recommendations human. CDC. 2025 Jan 21. URL: https://www.cdc.gov/hpv/hcp/vaccination-considerations/index.html [accessed 2025-04-16]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
EBP: evidence-based practice
EHR: electronic health record
GED: General Educational Development
HIPAA: Health Insurance Portability and Accountability Act
HPV: human papilloma virus
IRB: Institutional Review Board
mHealth: mobile health
PSA: public service announcement
RCT: randomized control trial

XSL•FO
RenderX

Edited by M Balcarras; submitted 16.Jun.2025; peer-reviewed by A Chen, K Culhane-Pera; revised version received 16.Jun.2026; accepted 16.Jun.2026; published 09.Sep.2026.

Please cite as:
Cunningham-Erves J, Hull PC, Dempsey AF, Koyama T, Sun L, Chalise M, Stewart EC, Jones JL, Landsittel D, Smalls M, Minter FF, Mayberry LS, Harnack L, Hook MJ, Cates J, Van Maanen T, Burnett T, Wilkins CH
A Multimethod Evaluation to Assess Feasibility, Acceptability, and Preliminary Efficacy of HPVVaxFacts, a Tailored Mobile Web App, for Parents With Unvaccinated Children: Pilot 2-Arm Randomized Controlled Trial
JMIR Cancer 2026;12:e78910
URL: https://cancer.jmir.org/2026/1/e78910
doi:10.2196/78910

© Jennifer Cunningham-Erves, Pamela C Hull, Amanda F Dempsey, Tatsuki Koyama, Lili Sun, Mariya Chalise, Elizabeth C Stewart, Jessica L Jones, Douglas Landsittel, Meredith Smalls, Freneka F Minter, Lindsay S Mayberry, Lora Harnack, Michael J Hook, Janet Cates, Tammy Van Maanen, Taylor Burnett, Consuelo H Wilkins. Originally published in JMIR Cancer (https://cancer.jmir.org), 9.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Influence of Social Media Messaging on Human Papillomavirus Vaccine Attitudes and Confidence Among Adolescent Males: Group Discussion Study

Neha Trivedi[1], MPH, PhD; Rachel Van Vleet[1], MPH; Czesia Eid[1], MPH; Amelia Burke-Garcia[1], MA, PhD; Amy E Leader[2], MPH, DrPH

[1]NORC at the University of Chicago, Washington, DC, DC, United States

[2]Thomas Jefferson University, Philadephia, PA, United States

Corresponding Author:
Neha Trivedi, MPH, PhD
NORC at the University of Chicago
1828 L St NW 9th floor,
Washington, DC, DC, 20036
United States
Phone: 1 7322075756
Email: trivedi-neha@norc.org

Abstract

Background: Despite the human papillomavirus (HPV) vaccine being available to males for the past 12 years, adolescent males continue to lag in HPV vaccine uptake due to a variety of factors. With the ubiquitous nature of social media use among this population, further research is needed to improve HPV vaccine confidence among young adolescent males using these platforms.

Objective: Using the Elaboration Likelihood Model (ELM), this study sought to better understand knowledge, attitudes, and beliefs about HPV and HPV vaccination among male adolescents aged 14-17 years and examine how social media messaging influences vaccine perceptions and explore the characteristics of persuasion and trustworthiness of digital content narratives across age and vaccination status.

Methods: The study team recruited 18 adolescent males to participate in a series of online focus groups and interviews stratified by age (14-15 years and 16-17 years) and HPV vaccination status. The study team worked with our in-house national probability survey panel, AmeriSpeak Teen Panel, to recruit for the discussion groups and interviews. All discussions were recorded, transcribed, and coded using NVivo 14 (Lumivero). ELM was used to develop discussion guides, codebook, and analysis. Emergent themes were identified, and the full team participated in double-coding and analysis. Data were analyzed using rigorous thematic analysis to identify patterns across groups.

Results: Knowledge of HPV varied by age and vaccination status. Older, vaccinated adolescents were more likely to understand that HPV affects both genders and is sexually transmitted, whereas younger, unvaccinated adolescents often had little to no prior awareness. Parents—especially mothers—were the primary decision-makers for vaccinated adolescents, with older adolescents more likely to be engaged in that decision. Trust in social media health messages was low overall, with participants relying on perceived source credibility over content format or style. Trusted sources included parents, physicians, and well-known health organizations. Younger adolescents were more influenced by personal anecdotes, whereas older adolescents preferred statistics, facts, and reliable sources.

Conclusions: Our findings highlight the importance of tailoring HPV vaccine messaging by age and developmental stage. Trusted messengers, clear factual content, and platform appropriateness are critical for engaging adolescent males. ELM provides a useful lens for interpreting these findings: younger adolescents were more influenced by peripheral cues, personal stories, or visual appeal, whereas older adolescents engaged in more central processing, evaluating the credibility of the source and the factual content of the message. This underscores the need to match message strategies with adolescents' cognitive and motivational readiness. Future interventions should consider dual strategies targeting both youth and parents, especially for younger adolescents. As the digital landscape continues to evolve, further research should also examine how generative artificial intelligence (AI) may serve as a trusted source or tool for communicating accurate, engaging health information to adolescent audiences.

(JMIR Cancer 2026;12:e82210) doi:10.2196/82210

XSL•FO
RenderX

KEYWORDS

adolescent males; HPV vaccination; HPV; human papillomavirus; qualitative research; social media influence

Introduction

Human Papillomavirus Vaccine Uptake Among Adolescents and Slow Adoption of HPV Vaccination Among US Adolescent Males

Human papillomavirus (HPV) remains the most common sexually transmitted infection in the United States despite vaccine availability to prevent many of the infection's strains in both young women and men [1]. Recent data from the CDC (Centers for Disease Control and Prevention) indicate that HPV vaccine uptake among US adolescents remains suboptimal, with notable gender disparities. In 2023, approximately 72% of adolescents aged 13-17 years initiated the HPV vaccine series, but only 49% completed it [2]. Although the CDC recommends routine vaccination at ages 11-12 years, initiation and completion rates remain consistently higher among females than males. For instance, from 2011 to 2020, the vaccination rate among males increased from 8% to 36%, whereas among females, it rose from 38% to 49% [3]. Studies suggest factors contributing to higher uptake among females include greater health awareness, proactive health care engagement, and stronger recommendations from health care providers [4]. These findings underscore the need for better understanding of the facilitators and barriers to vaccination to ensure equitable access to HPV prevention measures for adolescent males.

While efforts to vaccinate adolescent females for HPV have made strides in reducing HPV infections among this group, HPV vaccination delay, hesitation, and rejection continue to be prevalent among adolescent males [5]. Research suggests that if vaccination levels among adolescent males aged 14-17 years continue to dwindle, rates of HPV-related diseases will rise, continuing to widen vaccination disparities among this population [6]. Numerous interventions to promote HPV vaccination uptake among females exist [7-9], as well as interventions for parents and providers to promote HPV vaccination among younger males [10,11]; however, additional catch-up opportunities are needed for adolescent males aged 14-17 years. Various factors may contribute to the low uptake of HPV vaccination among adolescent males. For example, some health care providers have expressed reluctance to recommend the vaccine to males, citing perceived low risk and concerns about cost-effectiveness [12]. Parents may also be less aware of the vaccine's benefits for males due to an overall lack of communication about the HPV virus, therefore perceiving HPV-related diseases as less of a concern for their sons [13]. Furthermore, misconceptions that the vaccine promotes promiscuity and general vaccine hesitancy contribute to lower uptake [13].

Adolescents' HPV Vaccine Confidence and Health-Related Decision-Making Capabilities

Adolescents aged 14-17 years are developing greater autonomy and self-efficacy, which shape not only their health-related decision-making but also how they engage with health information on social media. Research shows that better knowledge of vaccine-preventable diseases such as HPV, higher confidence in vaccines, and active involvement in the decision-making process may explain a positive relationship with adolescents' vaccine uptake [14]. Similarly, adolescent maturity and self-efficacy play a role in HPV vaccination decision-making and overall vaccine confidence [15]. Theoretical models of health behavior suggest a role for self-efficacy (one's belief in one's ability to perform a certain behavior) in explaining why individuals do or do not engage in health behaviors, and these models have been applied in the context of HPV vaccination [16]. Social media platforms support self-efficacy by allowing adolescents to seek information independently, observe peer norms, and form attitudes outside of parental or clinical settings [17]. Recognizing that this age group prefers to be involved in decision-making, additional information is needed to understand what factors influence communication and messaging about health-related decision-making and vaccine confidence.

Effective communication strategies, particularly those initiated by health care providers, have been shown to increase vaccine uptake. However, given adolescents' high engagement with digital platforms, social media now plays a critical role in shaping health beliefs and behaviors. Platforms like TikTok, Instagram, and YouTube serve as both information sources and social influence networks. Through these platforms, digital storytelling has emerged as a promising approach; studies have shown the importance of tailored, evidence-based messaging in promoting HPV vaccination among adolescent males and how the evolving digital landscape can help engage this vulnerable age group [18]. This study will add critical information to understanding adolescent males' motivation and confidence in vaccination decision-making and the role social media plays in vaccine decision-making.

Social Media Messaging and Influence on US Adolescent Males

Social media plays a pivotal role in shaping adolescent males' health decisions, including HPV vaccination knowledge, attitudes, and beliefs. Male adolescents often seek health information on platforms like Instagram, TikTok, and YouTube, where content is frequently shared by peers, influencers, or health care professionals [19]. Studies have shown that social media channels can act as positive conduits of message delivery for adolescents if messages are considered interesting, their privacy is protected, and the source is credible [20].

The credibility of these messages is influenced by factors such as the perceived trustworthiness of the source, message framing, and platform-specific features. A systematic review identified 3 key domains affecting adolescents' trust in health information on social media: trust in the platform, trust in other users, and trust in the content itself [21]. Adolescents tend to trust content that aligns with their social identity, is presented in an engaging manner, and comes from sources they perceive as authentic or relatable [21]. For example, messages delivered by peers or influencers with whom adolescents identify may be more

XSL•FO
RenderX

persuasive than those from traditional health authorities. However, other studies suggest that messages developed and delivered by health care organizations or professionals provide reliable and factual information, providing higher believability and trustworthiness of a social media post [22,23].

The rapid evolution of platforms and the constant influx of new content can create an environment in which adolescents are continuously exposed to health-related messages. Studies have shown that exposure to vaccine misinformation on social media is associated with increased vaccine hesitancy, particularly when individuals rely solely on social media for health information without consulting trusted sources such as health care providers [24]. Additionally, the persuasive power of social media is heightened when messages are tailored to resonate with the audience's values and concerns. For instance, content that emphasizes peer experiences or addresses specific concerns about HPV vaccination may be more effective in influencing adolescent males' attitudes and behaviors. Therefore, understanding the factors that contribute to trust and engagement on social media is crucial for developing effective health communication strategies aimed at increasing HPV vaccination confidence among adolescent males. Research has yet to show how message characteristics and believability of social media messages targeting adolescent males impact vaccine confidence.

Theoretical Framework

This study used a qualitative descriptive design grounded in the Elaboration Likelihood Model (ELM), which is well suited for examining how adolescents perceive, interpret, and make meaning of health-related messaging in real-world contexts and social media. ELM posits that when an individual encounters various forms of communication, they process the communication with varying levels of elaboration ranging from low to high [25]. There are multiple factors that contribute to the elaboration of the messaging, requiring the individual to take 2 different routes, peripheral or central. Engaging in lower elaboration, or the peripheral route, results in the individual having lesser ability to process the message and having little or no interest in the subject. By contrast, engaging in higher elaboration, or the central route, results in higher motivation and ability to think about the message and its topic and to react.

More recently, ELM has been adopted to understand attitudes toward COVID-19 vaccination and intention to obtain the vaccine [26]. One study applied ELM to explore the characteristics of COVID-19 vaccine messages that may appeal to Twitter users, with findings suggesting that content-related messaging played a role in shaping decisions regarding whether to retweet antivaccine messages [27]. ELM has also been previously used to understand HPV vaccine messaging on social media to examine intentions and decision-making among parents. One study found that the central route, represented by information quality, affected parents' perceptions of HPV severity and susceptibility; the peripheral route, represented by source credibility, influenced their perceptions of HPV severity, HPV susceptibility, and vaccine response efficacy [28]. Another study with young girls found that recommended HPV vaccine messages should include focusing on cancer prevention rather than sexual transmission, routinization of the vaccine, and highlighting risks and costs of getting HPV [29]. Lastly, one study found that among college students, the use of hybrid messages containing both statistical and narrative descriptions of HPV resulted in greater perceived risk of getting HPV than either message type alone [30]. However, there has yet to be a study assessing believability of messaging impacting adolescent males aged 14-17 years in their decision-making to obtain HPV vaccines.

Study Design and Aims

This is an observational study that used qualitative methods to explore the knowledge, attitudes, and beliefs, along with the believability, persuasion, and influence of social media messaging on adolescent males' HPV vaccine confidence among those aged 14-17 years. Using the ELM framework [31,32], a dual-process communication and psychological framework, as our foundation, we examined how adolescent males process social media messaging (exploring the characteristics and persuasion of various narratives) and how such messages influence believability and adolescent males' confidence in HPV vaccine uptake.

The purpose of this study was to better understand the characteristics of social media narratives and their impact on adolescent males' perceptions about and confidence in receiving the HPV vaccine among those aged 14-17 years. ELM guided our protocol development and analysis, focusing on what this group believes and ways they are persuaded, based on social media influence. The study had two aims: (1) to understand knowledge, attitudes, and beliefs about HPV vaccination among male adolescents aged 14-17 years; and (2) to explore the characteristics and persuasive nature of social media (Facebook, TikTok, Twitch, YouTube, etc) narratives to which adolescent males are exposed and whether messaging plays an active role in impacting vaccine confidence among adolescent males.

Methods

Study Overview

A qualitative descriptive approach was selected to capture participants' perspectives in their own words and to generate findings with direct relevance for public health communication practice. The original study design included conducting 4 online focus groups with adolescent males aged 14-17 years to address the 2 key objectives. Using AmeriSpeak (Nation Opinion Research Center [NORC]'s the nationally representative probability panel of more than 30,000 US households) and AmeriSpeak Teen Panel (NORC's nationally representative sample of US adolescents aged 13-17 years), a group of male adolescents were recruited into the study [33]. The groups were stratified by age and HPV vaccination status: vaccinated males aged 14-15 years, unvaccinated males aged 16-17 years, vaccinated males aged 16-17 years, and unvaccinated males aged 14-15 years. However, due to recruitment challenges with participants aged 14-15 years, the study design was adapted to include a combination of focus groups and individual interviews. This revised approach maintained stratification by age and vaccination status and ultimately improved recruitment outcomes, particularly for the younger age group, through successful scheduling of individual interviews.

XSL•FO
RenderX

Study Eligibility and Recruitment

To be eligible for this study, parents must have (1) had a male child or male children aged 14-17 years and (2) not expressed strong antivaccine views to the following three questions: (1) "Overall, how hesitant about childhood shots would you consider yourself to be?"; (2) "To what extent do you agree with the following statement: I believe my son should be vaccinated against HPV?"; and (3) "Have you ever spoken with your son about the HPV vaccine?" Additional parental demographics were also collected, which were not part of the eligibility criteria.

Eligibility for adolescent males included (1) being aged 14-17 years and (2) HPV vaccination status ("Yes, I have received both HPV shots and I am fully vaccinated," "Yes, I have received one HPV shot but I have not received my second dose yet," "Yes, but I'm not sure if I received both doses," and "No"). Additional items such as demographics and social media usage were also captured, not part of eligibility criteria. Recruitment and data collection spanned from October 2024 to January 2025.

The panel invited 464 parents, and 326 consented for their adolescent to be offered the study opportunity. Among the 326 male adolescents aged 14-17 years invited to the recruitment screener survey, 111 adolescents were eligible and assented to be contacted for scheduling into a focus group or interview. To ensure homogeneity within each focus group or interview, we stratified participants based on the adolescents' age and vaccination status.

Study Procedures

The research team was composed of investigators with training and experience in public health, health communication, qualitative research methods, and vaccine confidence research. Several team members have prior experience designing and evaluating adolescent-focused health communication interventions, which informed the development of the discussion guide and the selection of social media stimulus materials.

The research team developed the focus group discussion guide using ELM to ask questions about knowledge, attitudes, and beliefs around HPV vaccination and to assess the believability of social media messaging on adolescent males' decision-making and vaccine confidence through message testing. We examined how adolescent males process social media messaging (exploring the characteristics and persuasion of various narratives) and how such messages influence believability and adolescent males' confidence in HPV vaccine uptake. The discussion guide (Multimedia Appendix 1) was constructed with specific topics (Textbox 1).

Textbox 1. Discussion group section topics.

Section topic
• Knowledge of the human papillomavirus (HPV) and HPV vaccine
• Attitudes and beliefs toward HPV vaccination
• Barriers and motivators to HPV vaccine uptake
• Social media platform use
• Trust/distrust in messages/sources on social media
• Believability of vaccine messages on social media

We also visually presented 4 social media posts on HPV vaccination for discussion. The research team created a social media HPV vaccine media scan protocol to gather a sample of media on HPV vaccination to which adolescent boys may be exposed or that they may actively seek online. Queries were created based on platforms (eg, TikTok, Reddit, X [formerly known as Twitter], Twitch, Facebook, Instagram), time period (eg, past year), content, narrative format, and messenger. TalkWalker [34] and CrowdTangle [35] (before it was obsolete) were used to pull and analyze media posts. Posts were analyzed and discussed as a team for inclusion based on the factors within the query. A matrix was created based on the query terms, and the research team discussed which posts were available and relevant to the study. We compared posts selected from both individuals and organizations, platforms, and messages, including facts compared with narratives (Multimedia Appendix 2). One member of the study team moderated the discussion, while another member was the notetaker. Each focus group lasted approximately 1 hour, whereas interviews lasted approximately 30 minutes.

Analysis

Multiple strategies were used to ensure the trustworthiness and rigor of the qualitative findings. Credibility was enhanced through the use of verbatim transcripts, double coding of a subset of interviews and focus groups, and regular analytic debriefings among the research team to discuss emerging themes and resolve discrepancies. Upon completion of the discussion groups and interviews, the transcripts were reviewed, updated by inserting moderator questions for clarity, and deidentified. A codebook was developed using a mix of a priori codes based on the discussion guide along with themes that emerged from each discussion group and interview that were relevant to the research questions [36]. Two independent coders from the research team double-coded 4 transcripts (based on the stratification of the groups by vaccination status and age) using NVivo 14 (Lumivero) [37] to conduct qualitative analysis. Following the first round of coding, discrepancies were discussed and addressed, and the codebook was updated where needed. Data saturation was assessed iteratively during analysis, with no new themes emerging across later focus groups and interviews, suggesting that thematic sufficiency was achieved.

XSL•FO
RenderX

In addition to double coding by the research team, regular team debriefings, and the use of a structured codebook helped ensure analytic consistency and credibility of findings. Interrater reliability [38] was calculated in MAXQDA using the Cohen κ metric. Using the 4 transcripts, double coding yielded an interrater reliability of 0.74. Codes were then applied to the remainder of the transcripts. Once coding was complete, we structured the presentation of results by themes.

Ethical Considerations

This study was reviewed and approved by the NORC at the University of Chicago Institutional Review Board (IRB; protocol number 24-07-1895). NORC operates its own IRB, which is registered with the US Department of Health and Human Services Office for Human Research Protections. For projects that are federally funded, the research protocol and accompanying documents have been approved under the expedited review procedure in accordance with HHS Regulations at 45 CFR 46.110 expedited review categories, by the NORC IRB (IRB00000967), under its Federalwide Assurance Number FWA00000142 (FWA 00000142). NORC's IRB follows a formal process for reviewing all research projects to assure human participants' protections and minimize respondent burden.

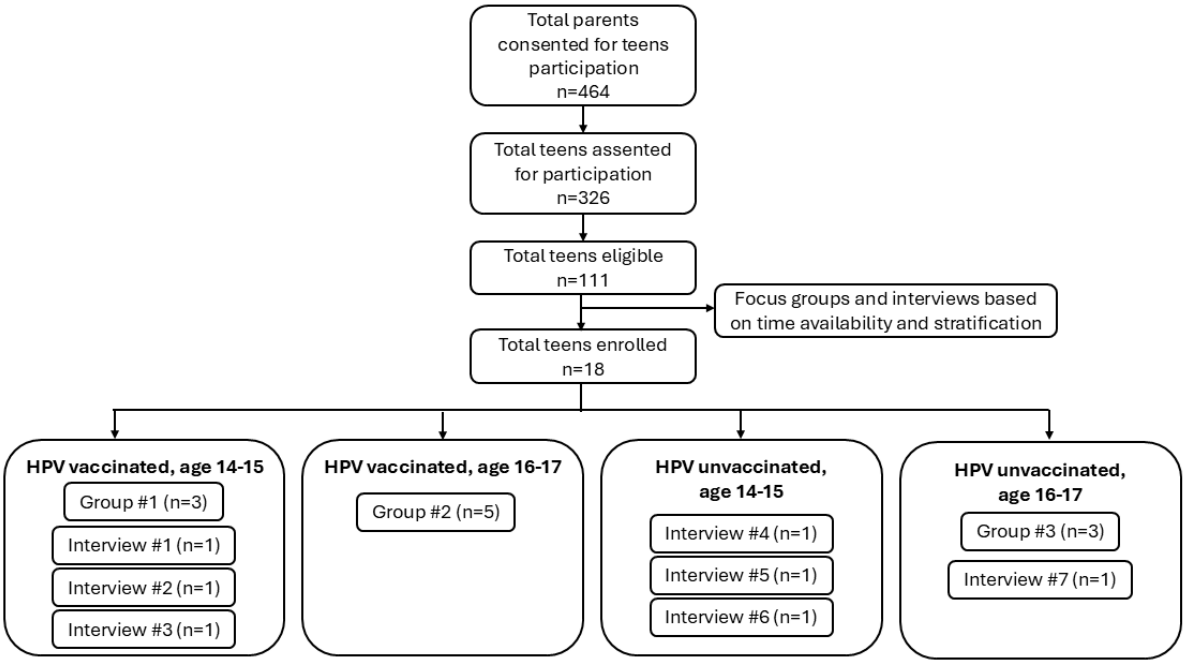
Due to the sensitive nature of the study topic, parents of eligible adolescents provided consent for their adolescent to participate. Both the parent consent language and the adolescent recruitment screener and assent were offered only in English and were self-administered online by the respondent. Consent and assent materials were provided in English and were self-administered online via the AmeriSpeak and AmeriSpeak Teen Panel. No waivers of consent or assent were requested or granted. All discussions were audio-recorded and transcribed. Transcripts were deidentified before analysis (eg, removal of names and other direct identifiers), with moderator questions inserted for clarity without introducing identifying information. Access to deidentified data was restricted to the research team, and results are reported in aggregate to further reduce reidentification risk. Participants received AmeriPoints in the amount of US $50 for participating in the focus group discussion and US $20 for an interview.

Results

Study Flow and Participant Distribution

Figure 1 provides the study flow and final distribution of participants across data collection methods.

Figure 1. Study flow diagram and recruitment.



Sample Characteristics

A total of 18 male adolescents aged 14-17 years participated in either a focus group discussion or an interview.

Table 1 presents the overall characteristics of the adolescent male participants. A total of 9 males were aged between 14-15 years, while 9 males were aged 16-17 years old. Participants were primarily White, non-Hispanic (n=10), and varied in terms of state of residence. The majority (n=11) reported receiving the HPV vaccination, whether 1 dose or both. YouTube (n=15) and Instagram (n=13) were the most used social media platforms.

XSL•FO
RenderX

Table 1. Participant characteristics.

Participant characteristics	Frequency, n (%)
Age (years)	
14	5 (28)
15	4 (22)
16	3 (17)
17	6 (33)
Vaccination status	
Yes, I have received both HPV shots, and I am fully vaccinated	8 (44)
Yes, I have received one HPV shot, but I have not received my second dose yet	0 (0)
Yes, but I am not sure if I received both doses	3 (17)
Not vaccinated	7 (39)
Race/ethnicity	
Black, non-Hispanic	3 (17)
White, non-Hispanic	10 (56)
Asian-Pacific Islander	1 (6)
Hispanic	2 (11)
2+, non-Hispanic	2 (11)
Household income (US $)	
<5000	1 (6)
15,000-29,999	3 (17)
30,000-39,999	3 (17)
40,000-59,999	3 (17)
85,000-124,999	3 (17)
125,000-149,999	3 (17)
175,000-199,999	2 (11)
State of residence	
Arizona	1 (6)
California	1 (6)
Florida	3 (17)
Kansas	1 (6)
Michigan	1 (6)
Minnesota	1 (6)
Missouri	1 (6)
New York	1 (6)
North Carolina	1 (6)
Ohio	1 (6)
Oregon	1 (6)
South Carolina	2 (11)
Tennessee	1 (6)
Wisconsin	2 (11)
Social media platform use[a]	
Instagram	13 (72)
Facebook	9 (50)

XSL•FO
RenderX

Participant characteristics	Frequency, n (%)
Reddit	5 (28)
Snapchat	11 (61)
TikTok	10 (56)
Twitch	5 (28)
YouTube	15 (83)
X (formerly known as Twitter)	1 (6)

[a]Respondents reported using multiple social media platforms; total will not equal N=18.

Thematic Analysis

Findings from this study map clearly onto the core constructs of ELM, which posits that individuals process persuasive information along either the central or peripheral route depending on their motivation and ability to engage with the message. During analysis, transcripts were coded using ELM. ELM codes included central route processing (ability to process, motivation to process, emotional cues, personal relevance, motivation to process, and message quality) and peripheral route processing (familiarity or visual cues, heuristic processing, and source credibility). Social media post text vs source reliability were analyzed to determine how different message features influenced adolescents' engagement and vaccine-related attitudes. This approach allowed us to explicitly link participants' reactions to theoretically defined ELM pathways and assess how message characteristics may differentially shape HPV vaccine decision-making.

Table 2 presents results from the thematic analysis and is organized by knowledge, attitudes, and beliefs about HPV and HPV vaccine, barriers and motivators to HPV vaccine uptake, social media use, and trust and believability of vaccine messages on social media. Sample quotes are provided to support findings, based on ELM, and are identified by age and vaccination status of the focus group participant or interviewee (Table 2).

XSL•FO
RenderX

Table 2. Thematic analysis of focus group and interview data.

Theme and subtheme (if applicable)	Related ELM[a] construct	Illustrative quote
Knowledge, attitudes, and beliefs about HPV[b] and HPV vaccine	• Ability to process • Motivation to process	• I knew boys should get the vaccine from school, so my mom took me to get it, and my aunt had cervical cancer so knew it was something that affects men and women. [FG participant, 16yo, vaccinated]
Barriers and motivators to HPV vaccine uptake	• Emotional cues • Personal relevance	• I'm just cautious of what I put in my body, like any medication. I'm always concerned with side effects. My parents are cautious too. What are the fertility implications? [FG participant, 16yo, unvaccinated]
Social media use	• Motivation to process • Central route processing • Message quality	• I don't really use [social media] for anything research related, I only really use it for talking to my friends or communication because I would rather take some more time using Google or some other website rather than looking on a social media app, because misinformation on there spreads like wildfire and I don't want to be misinformed about something that I'm trying to learn. So I don't really use it much for things that are political or medical. [FG, 15yo, vaccinated]
Social media use for health information	• Motivation and ability to process • Central route processing	• We are interested to learn about the risks and the contents of it and eligibility. [FG participant, 17yo, vaccinated] • I would like to see more of the health concerns. And maybe the risk, because they usually have that really quick. So people don't hear about it very much. People need to talk about how you can get it. [FG, 14yo, vaccinated] • I'd like it to tell the risks about if you got the vaccine and if you didn't, like the pros and cons to each. And what is HPV, and more about the risks of that. [IDI, 14yo, unvaccinated]
Trust and believability of vaccine messages on social media	• Peripheral route processing (familiarity/visual cues), • Heuristic processing • Source credibility	• It looks relatively believable because it shows a picture of Robert F. Kennedy Jr. The 5thlawsuit he has filed against Merck, challenging the company's dangerous and defective HPV vaccine for causing severe and life-changing injuries. Now, obviously, in the second picture you can see that he is not in good health. Like he might be smiling there but he does not look happy. Yeah, I'm pretty sure it's real, but since there is a picture of when he was young and a picture of when he is older sometimes those pictures can be not always right, could be a different person, but I'm pretty sure this is real. [FG, 15yo, vaccinated] • Some of these social media messages are old, this one [RFK] is even older, a lot of things can change in a vaccine, modern medicine advances.I don't know if it's true, anyone can get hospitalized – that's a lot of likes but don't know if that matters that much. [IDI, 14yo, vaccinated] • Since Elon bought it [Twitter], it's filled with misinformation- blue check marks are bot accounts to farm likes for cash, account doesn't seem like genuine, would be paid bot account. [FG, 17yo, vaccinated] • The TikTok video is less believable because anyone can dance with scrubs on. I have no idea if she's a real doctor. [IDI, 14yo, vaccinated]

[a]ELM: Elaboration Likelihood Model.

[a]HPV: Human papillomavirus.

Knowledge, Attitudes, and Beliefs About HPV and HPV Vaccine

Adolescents' knowledge, attitudes, and beliefs about HPV and the vaccine varied notably by age and vaccination status. All vaccinated adolescents aged 16-17 years (n=5) demonstrated the most comprehensive understanding, often recognizing that HPV affects both males and females and that the vaccine is recommended for all adolescents regardless of gender. In contrast, two-thirds of unvaccinated adolescents aged 14-15 years (2/3) held the misconception that the HPV vaccine is only for girls, and several reported hearing about it for the first time during the study.

All vaccinated adolescents aged 14-15 years (n=6) were aware that they had received the vaccine but lacked clarity about what HPV is, how it is transmitted, and what symptoms it causes. They described vaccination as something that happened during a routine physician visit, without detailed explanation. Approximately 80% of the vaccinated and unvaccinated adolescents aged 16-17 years (7/9) generally knew that HPV is a sexually transmitted infection and could name potential consequences of infection, including warts and cancer. These older participants had also heard about methods of prevention beyond vaccination, such as using protection during sex. Overall, knowledge was limited among younger adolescents, particularly those who were unvaccinated, whereas older participants—regardless of vaccination status—were more

XSL•FO
RenderX

informed and engaged in thinking critically about HPV and its risks.

Barriers and Motivators to HPV Vaccine Uptake

Among all vaccinated adolescent males, parental influence—notably that of their mothers—was a consistent motivator for HPV vaccine uptake. A total of 80% (4/5) of the vaccinated adolescents aged 16-17 years described shared decision-making with their parents before receiving the vaccine, during which they discussed the benefits of the vaccine, asked questions, and ultimately agreed to get vaccinated. Some recalled learning about HPV from their parents or pediatricians, reinforcing their confidence in the decision. In contrast, 80% (5/6) of vaccinated adolescents aged 14-15 years did not have conversations with their parents or physicians before receiving the vaccine. Despite limited involvement in the decision-making process, most expressed trust in their parents' choices regarding their health.

Unvaccinated adolescents identified several barriers to vaccination. Across age groups, concerns about adverse effects and parental skepticism were prominent. All unvaccinated adolescents aged 14-15 years (n=3) believed that vaccination was not necessary because they were not engaging in sexual activity through which HPV could be contracted, reflecting a perception that the vaccine is only needed once someone becomes sexually active. Many felt that "if you're safe, you will be fine." All unvaccinated adolescents aged 16-17 years (n=4) were more cautious about vaccine uptake because of adverse effects commonly associated with vaccine, which was often influenced by parental concerns. Despite hesitancy, many unvaccinated adolescents from both age groups expressed interest in learning more and emphasized that they would want to conduct their own research before deciding whether to be vaccinated.

Social Media Use

Overview

Across vaccination status and age groups, adolescent males noted that they were on social media daily, primarily for entertainment and to chat with friends. Instagram was reported as the most popular platform for everyday social media use. TikTok, YouTube, and X (formerly known as Twitter) were also noted as platforms used across age groups. Among older adolescents aged 16-17 years, platforms such as Snapchat and Discord were especially prevalent and were often used for direct messaging, gaming-related conversations, and school-related group chats. Reddit was more commonly mentioned by younger adolescents aged 14-15 years, who used it to follow specific interests and read anonymous discussions. While social media was part of everyday life, few participants reported using it intentionally for health information.

Social Media Use for Health Information

Across vaccination status and age group, adolescent males indicated that they do not seek health information on social media because of concerns regarding the reliability of information circulating online. When asked what types of information they would want HPV social media content to address, many noted that they would like to understand the risks of contracting HPV, how the vaccine would mitigate those risks, and where to obtain the vaccine. Similarly, many adolescents noted that they wanted to understand the risks associated with vaccination and how vaccination might affect their future health if they were to receive the vaccine (eg, fertility issues in the future).

Trust and Believability of Vaccine Messages on Social Media

Across all age groups and vaccination statuses, adolescent males noted that social media is not their first source for accurate, reliable health information. Most participants said they would not trust HPV vaccine content on social media, with some saying they would trust it if the content was from a trustworthy source. A total of 67% (12/18) of participants said they would do more research to understand which social media posts are accurate and trustworthy and would seek multiple sources for health information given the spread of misinformation on social media.

When asked about trusted sources, older male adolescents noted that they would turn to their parents or personal physician for reliable and accurate information about HPV. Approximately 60% (3/5) of vaccinated older adolescents also said they trust health information from organizations such as the World Health Organization, Mayo Clinic, or another major hospital. About 40% (2/5) of older adolescents also said they would trust content that their peers, classmates, friends, or family members shared on social media.

When shown the 4 social media narratives, adolescent males, regardless of age, noted that the Maryland Department of Health post was trustworthy because it provided facts and was from a reliable source. There were mixed results, leaning to not trustworthy, about the InfoWars post because the source cited its own article, and many adolescents did not know the source. All adolescents noted that the TikTok narrative was not believable because messaging with singing and dancing was deemed unserious. A little over 80% (15/18) of participants noted that the narrative from Robert F Kennedy Jr would warrant a conversation with their physician or parent about the adverse effects of the HPV vaccine. However, two-thirds of unvaccinated adolescents aged 14-15 years (n=3), believed that narrative because they knew Robert F Kennedy Jr and had heard about risks associated with vaccination.

Participants found that the persuasiveness of a message was influenced by 3 key factors: content format, the substance of the message, and the source's credibility. Most participants found the dancing doctor video on TikTok to be "unprofessional," saying that it does not convey the seriousness of the issue and the content did not seem reliable as presented. On the other hand, the images from the Robert F Kennedy Jr post made it compelling through emotional appeal. Unvaccinated adolescents aged 14-15 years were more likely to believe social media messaging that included personal anecdotes, whereas the adolescents aged 16-17 years were more persuaded by statistics and facts.

Older adolescents also emphasized the importance of examining the account that posted the content. Source credibility emerged

XSL•FO
RenderX

as a salient factor for adolescents aged 16-17 years, highlighting that respondents were more likely to evaluate the reliability and validity of health information based on the authority, expertise, and affiliations of the poster. For example, the Maryland Department of Health post was widely considered credible because of its factual tone and clear institutional affiliation. The majority of respondents had heard of the Maryland Department of Health and found that this source would be credible for health information. In contrast, the InfoWars post was largely dismissed, with participants citing unfamiliarity with the source and concerns about its legitimacy.

Some participants said that the platform on which the narrative was posted impacted believability, suggesting that Instagram and Facebook are more trustworthy than TikTok and X. Participants also noted that other post characteristics, such as the date of the post and engagement metrics (eg, likes and reposts), impacted believability.

Discussion

Overview

This study underscores the significant role of social media in shaping adolescent males' knowledge and attitudes toward HPV vaccination. Consistent with the study aims, we found that adolescent males' engagement with HPV vaccine messaging on social media varies by age and vaccination status. Notably, unvaccinated and younger adolescent males exhibited lower awareness of HPV and the vaccine, rendering them more susceptible to misinformation prevalent on social media platforms. This aligns with findings from a systematic review indicating that adolescents often distrust online health information but continue to use it, using various heuristics to assess credibility, such as the source's perceived trustworthiness and the content's presentation [39]. The preference for information from familiar or relatable sources, including personal physicians and mothers, highlights the need for targeted health communication strategies that resonate with this demographic.

Furthermore, our research revealed that the believability of social media messages was predominantly influenced by the source rather than the content or format of the post. Older adolescent males demonstrated a higher propensity to trust messages from credible sources, such as health care providers and reputable organizations, and were more likely to engage with authoritative websites such as the CDC and Mayo Clinic for health information. Conversely, younger, unvaccinated males were more inclined to believe anecdotal messages featuring personal stories, which can be compelling but may not always present accurate or comprehensive information. This finding is consistent with studies showing that adolescents' trust in health information on social media is complex and influenced by factors such as platform trustworthiness, user identity, and content credibility [21]. Additionally, while parents and personal physicians were universally trusted sources of information for all adolescent males, social media posts featuring unrecognizable physicians were found to be unpersuasive, underscoring the importance of source credibility in health communication. These insights contribute to the growing body of research on the influence of social media on health-related decision-making and highlight the necessity for public health interventions to navigate the evolving digital landscape effectively.

Alignment With ELM

Our findings align with ELM, which posits that individuals process persuasive messages through either the central or peripheral route depending on motivation and ability to engage with the content. Older adolescent males aged 16-17 years, especially those vaccinated, appeared to engage more often through the central route, processing information thoughtfully, evaluating arguments, and preferring content from trusted, authoritative sources. Their critical evaluation of message credibility, preference for fact-based content, and desire to consult external sources (such as physicians or reputable websites) are all hallmarks of central-route processing.

In contrast, younger adolescents were more likely to engage through the peripheral route, relying on emotional cues, personal anecdotes, and familiarity rather than systematic analysis. They tended to be more receptive to personal narratives from public figures or peers, relying on heuristic processing. These distinctions underscore the importance of tailoring message design not only by age but also by presumed level of cognitive engagement, which ELM helps conceptualize effectively. This theoretical grounding offers a valuable lens for public health practitioners aiming to design persuasive health messages based on audience readiness, message processing style, source credibility, and story-telling content.

Strengths and Limitations

This study has many strengths. First the research team's ability to engage directly with adolescent males through a nationally representative probability panel was a major strength. This group is often a difficult-to-reach population in health communication research; however, we were able to explore their nuanced perspectives on vaccine messaging and social media use through innovative recruitment efforts. Second, the use of a nationally representative, probability-based recruitment panel enhances the generalizability of findings to US adolescent males by capturing perspectives across diverse geographic regions and vaccination statuses rather than relying on convenience samples. Lastly, by using real-world examples of social media narratives, we captured insights grounded in participants' lived experiences, increasing the validity of our findings.

However, this study is not without limitations. First, despite using a nationally representative panel for recruitment, the final participant pool was relatively homogenous in terms of race/ethnicity and socioeconomic background, which may limit generalizability. Our sample was predominantly White and non-Hispanic. This homogeneity in our adolescent sample may have resulted in greater trust in clinicians and well-known public health agencies than would be observed among individuals with differing racial/ethnic and socioeconomic backgrounds. Similarly, participants described low reliance on social media for health decisions and a preference for "official" sources, noting that their information-seeking behaviors may not be the norm. Evidence suggests that Black, Hispanic/Latine, immigrant,

XSL•FO
RenderX

and economically disadvantaged communities can experience higher medical mistrust toward government, pharmaceutical companies, and, at times health care systems, rooted in historical and ongoing inequities; such mistrust is associated with lower HPV vaccine uptake and different messenger preferences (eg, family networks, faith leaders, and community-based organizations).

Second, while the original study design emphasized focus groups to encourage peer discussion, recruitment challenges, particularly among adolescents aged 14-15 years, necessitated a pivot to individual interviews. As a result, we were not able to fully explore group dynamics or peer influence, which may have provided additional regarding social norms and vaccine attitudes. It is important to note that this transition may have limited the ability to capture peer-to-peer influences that are particularly salient during adolescence. Group discussions may have revealed additional insights into how social norms, shared beliefs, and peer reinforcement shape perceptions of vaccine-related content. While individual interviews allowed for a safe space to discuss this sensitive topic and were effective in eliciting personal reasoning and decision-making processes, the absence of group interaction may have obscured how peer discussions and perceived norms influence older adolescents' confidence in, or resistance to, vaccine messaging.

Lastly, although the study leveraged the AmeriSpeak parent–teen panel to reduce recruitment bias, reliance on parental consent may still introduce some degree of selection bias. Parents and adolescents were consented into the AmeriSpeak panel before learning about individual study topics, which helps mitigate overrepresentation of parents with particularly strong provaccine views. Additionally, both parents and adolescents completed vaccine-related screener items that allowed us to capture a range of vaccine attitudes.

Future research should explore alternative strategies to better engage younger adolescents, including school-based recruitment partnerships, collaboration with pediatric practices, and integration of research activities into youth-centered digital environments such as moderated online communities or gamified platforms. In addition, incorporating parent-adolescent dyadic designs may also enhance participation while capturing the shared decision-making dynamics that are central to HPV vaccination at this age.

Implications for Future Communication Efforts

The distinct differences in how younger and older adolescent males engage with and interpret vaccine messaging on social media suggest that a one-size-fits-all approach would be ineffective. Instead, interventions should be tailored by age group, vaccination status, and other key factors such as health information seeking behaviors and use of digital tools (social media, generative artificial intelligence [AI], etc). For younger adolescents, messaging should likely target both the adolescent and their parents, given the significant influence of parental decision-making in this age group and the limited health literacy among adolescents aged 14-15 years. Messaging for this audience should use emotionally engaging content that includes personal stories and encouragement for younger adolescents to have conversations with their parents and physicians when engaging with social media health information posts that need further context. A testable hypothesis for future research while using the principles of ELM, could examine whether, among younger adolescents with lower baseline knowledge, brief narrative posts paired with visible peripheral cues of credibility (eg, verified badges and institutional logos) would increase perceived relevance and discussion with parents more than statistics-heavy posts.

Messaging for older adolescents should be grounded in facts and delivered by highly credible sources, such as pediatricians or well-known public health agencies. Delivery platforms that are familiar and accessible to both youth and parents, such as Instagram and Facebook, may be more effective than TikTok or X. For older adolescents, who are increasingly autonomous in their health decisions, messaging should include more fact-based content with direct appeals to logic, data, and the consequences and risks of not vaccinating. This audience showed a higher likelihood of engaging in their own research with reputable websites or posts from verified sources, indicating a preference for content that satisfies a need for accurate, actionable information. A testable hypothesis for future research, while using the principles of ELM, could examine whether older adolescents with higher motivation who receive fact-forward messages from identifiable health care professionals would demonstrate higher message credibility and intent to vaccinate compared with entertainment-style narratives without evidence-based content. These findings suggest future HPV vaccine campaigns should develop parallel message strategies for different developmental stages: one focusing on emotional resonance and parental influence for younger adolescents and the other emphasizing autonomy, scientific rationale, and source credibility for older adolescents.

Conclusion

Summary of Key Findings

This study sought to explore how adolescent males engage with and interpret HPV vaccine-related messaging on social media, with a focus on differences by age and vaccination status. Findings indicate that younger and unvaccinated adolescents demonstrate lower HPV awareness and greater susceptibility to misinformation, whereas older adolescents are more likely to engage critically with health information and seek content from credible, authoritative sources. Social media plays a central yet complex role in shaping perceptions, with message credibility often hinging on the source rather than the content or format. Trust in personal physicians and parents remains strong, and different age groups exhibit distinct preferences for how and from whom they receive health information.

Public Health Implications and Practical Applications

These findings have important implications for both HPV vaccination efforts and the broader field of adolescent health communication. Public health practitioners should consider age-segmented strategies tailored to developmental stages and align message framing, emotional tone, and source credibility with adolescents' decision-making autonomy when designing campaigns. For younger adolescents, practical applications include developing social media content that encourages

XSL•FO
RenderX

parent-adolescent conversations and uses emotionally resonant narratives while clearly directing families to trusted health care providers for follow-up information. For older adolescents, campaigns may be more effective when they emphasize factual content, transparent risk-benefit information, and messages delivered by identifiable, reputable health professionals or organizations.

Health care providers and public health agencies can also apply these findings by strategically integrating social media into vaccine counseling, using trusted platforms to reinforce in-clinic conversations rather than replace them. Training clinicians and public health practitioners to recognize how youth interpret social media narratives may enhance their ability to address misinformation and tailor recommendations during clinical encounters. Additionally, vaccine campaigns should prioritize platforms familiar to both adolescents and parents to maximize reach and credibility.

Future Research Considerations

As the digital ecosystem continues to evolve rapidly, it is also important to explore emerging technologies, such as generative AI, as potential tools for delivering reliable, factual vaccine content to youth. Understanding how adolescent males interact with AI-generated information, alongside traditional social media, will be critical for designing future interventions that keep pace with their digital footprint. Studies incorporating parent-adolescent dyads, peer-based interactions, and more diverse populations are also needed to better understand shared decision-making processes and social influences on HPV vaccine uptake. Longitudinal designs may further clarify how adolescents' message processing and trust in digital health information change over time.

Acknowledgments

Supported in part by a research grant from Investigator-Initiated Studies Program of Merck Sharp & Dohme LLC. The opinions expressed in this study are those of the authors and do not necessarily represent those of Merck Sharp & Dohme LLC.

Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

Funding

This research protocol was funded by Merck Sharp & Dohme LLC as part of their Investigator-Initiated Studies Program. The study was funded for a 1-year period beginning in June 2024. Total study funding was US $97,691.69.

Authors' Contributions

Conceptualization: NT (lead)
Data curation: NT (lead), RVV (supporting), CE (supporting)
Formal analysis: NT (lead), RVV (supporting), CE (supporting)
Funding acquisition: NT (lead), ABG (supporting)
Investigation: NT (lead), RVV (supporting), CE (supporting)
Methodology: NT (lead), RVV (supporting), CE (supporting)
Project administration: NT (lead), RVV (supporting), CE (supporting)
Resources: NT (lead), RVV (supporting), CE (supporting)
Supervision: NT (lead), ABG (supporting), AEL (supporting)
Visualization: NT (lead), RVV (supporting), CE (supporting)
Writing – original draft: NT (lead), RVV (supporting), AEL (supporting)
Writing – review & editing: NT (lead), RVV (supporting), CE (supporting), ABG (supporting), AEL (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Moderator guide for exploring the role of social media narratives and hpv vaccine confidence among adolescent males.
[DOCX File , 21 KB - cancer_v12i1e82210_app1.docx]

Multimedia Appendix 2
Social media posts.
[DOCX File , 959 KB - cancer_v12i1e82210_app2.docx]

References

XSL•FO
RenderX

1. CDC. Sexually transmitted infections (STIs). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/sti/about/about-genital-hpv-infection.html [accessed 2025-05-29]
2. Pingali C, Yankey D, Chen M, Elam-Evans LD, Markowitz LE, DeSisto CL. National vaccination coverage among adolescents aged 13-17 years—National Immunization Survey-Teen, United States, 2023. MMWR Morb Mortal Wkly Rep 2024;73(33):708-714 [FREE Full text] [doi: 10.15585/mmwr.mm7333a1] [Medline: 39173168]
3. Johnson CK. HHPV vaccines prevent cancer in men as well as women, new research suggests. AP News. URL: https://apnews.com/article/hpv-vaccine-oral-cancer-prevention-males-749bf59c10d8a5a86c71c89c50eda0cb [accessed 2025-05-29]
4. Efua Sackey M, Markey K, Grealish A. Healthcare professionals' promotional strategies in improving human papillomavirus (HPV) vaccination uptake in adolescents: a systematic review. Vaccine 2022;40(19):2656-2666 [FREE Full text] [doi: 10.1016/j.vaccine.2022.03.054] [Medline: 35367068]
5. Kester LM, Shedd-Steele RB, Dotson-Roberts CA, Smith J, Zimet GD. The effects of a brief educational intervention on human papillomavirus knowledge and intention to initiate HPV vaccination in 18-26 year old young adults. Gynecol Oncol 2014;132 Suppl 1:S9-S12. [doi: 10.1016/j.ygyno.2013.12.033] [Medline: 24384459]
6. Hirth J. Disparities in HPV vaccination rates and HPV prevalence in the United States: a review of the literature. Hum Vaccin Immunother 2019;15(1):146-155 [FREE Full text] [doi: 10.1080/21645515.2018.1512453] [Medline: 30148974]
7. Rebolj M, Pesola F, Mathews C, Mesher D, Soldan K, Kitchener H. The impact of catch-up bivalent human papillomavirus vaccination on cervical screening outcomes: an observational study from the English HPV primary screening pilot. Br J Cancer 2022;127(2):278-287 [FREE Full text] [doi: 10.1038/s41416-022-01791-w] [Medline: 35347326]
8. Paskett ED, Krok-Schoen JL, Pennell ML, Tatum CM, Reiter PL, Peng J. Results of a multilevel intervention trial to increase human papillomavirus (HPV) vaccine uptake among adolescent girls. Cancer Epidemiol Biomarkers Prev 2016 Apr;25(4):593-602 [FREE Full text] [doi: 10.1158/1055-9965.EPI-15-1243] [Medline: 27196093]
9. Hou Z, Wu Z, Qu Z, Gong L, Peng H, Jit M. A vaccine chatbot intervention for parents to improve HPV vaccination uptake among middle school girls: a cluster randomized trial. Nat Med 2025;31(6):1855-1862. [doi: 10.1038/s41591-025-03618-6] [Medline: 40195450]
10. Cates JR, Diehl SJ, Crandell JL, Coyne-Beasley T. Intervention effects from a social marketing campaign to promote HPV vaccination in preteen boys. Vaccine 2014;32(33):4171-4178 [FREE Full text] [doi: 10.1016/j.vaccine.2014.05.044] [Medline: 24886960]
11. Perkins RB, Zisblatt L, Legler A, Trucks E, Hanchate A, Gorin SS. Effectiveness of a provider-focused intervention to improve HPV vaccination rates in boys and girls. Vaccine 2015;33(9):1223-1229. [doi: 10.1016/j.vaccine.2014.11.021] [Medline: 25448095]
12. Holman DM, Benard V, Roland KB, Watson M, Liddon N, Stokley S. Barriers to human papillomavirus vaccination among US adolescents: a systematic review of the literature. JAMA Pediatr 2014;168(1):76-82 [FREE Full text] [doi: 10.1001/jamapediatrics.2013.2752] [Medline: 24276343]
13. Dibble KE, Maksut JL, Siembida EJ, Hutchison M, Bellizzi KM. A systematic literature review of HPV vaccination barriers among adolescent and young adult males. J Adolesc Young Adult Oncol 2019;8(5):495-511 [FREE Full text] [doi: 10.1089/jayao.2019.0004] [Medline: 31090474]
14. Cadeddu C, Castagna C, Sapienza M, Lanza TE, Messina R, Chiavarini M, et al. Understanding the determinants of vaccine hesitancy and vaccine confidence among adolescents: a systematic review. Hum Vaccin Immunother 2021;17(11):4470-4486 [FREE Full text] [doi: 10.1080/21645515.2021.1961466] [Medline: 34473589]
15. Karafillakis E, Peretti-Watel P, Verger P, Chantler T, Larson H. The role of maturity in adolescent decision-making around HPV vaccination in France. Vaccine 2021;39(40):5741-5747 [FREE Full text] [doi: 10.1016/j.vaccine.2021.08.096] [Medline: 34483023]
16. Batista Ferrer H, Audrey S, Trotter C, Hickman M. An appraisal of theoretical approaches to examining behaviours in relation to Human Papillomavirus (HPV) vaccination of young women. Prev Med 2015;81:122-131 [FREE Full text] [doi: 10.1016/j.ypmed.2015.08.004] [Medline: 26314783]
17. Hocevar KP, Flanagin AJ, Metzger MJ. Social media self-efficacy and information evaluation online. Comput Human Behav 2014;39:254-262. [doi: 10.1016/j.chb.2014.07.020]
18. Digital storytelling shows promise as a method to increase HPV vaccination rates among college-aged men. School of Communication and Information. URL: https://comminfo.rutgers.edu/news/digital-storytelling-shows-promise-method-increase-hpv-vaccination-rates-among-college-aged-men [accessed 2025-05-29]
19. Vogels E, Gelles-Watnick R, Massarat N. Teens, social media and technology 2022. Pew Research Center. 2022. URL: https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/ [accessed 2026-04-28]
20. Ortiz RR, Shafer A, Cates J, Coyne-Beasley T. Development and evaluation of a social media health intervention to improve adolescents' knowledge about and vaccination against the human papillomavirus. Glob Pediatr Health 2018;5:2333794X18777918 [FREE Full text] [doi: 10.1177/2333794X18777918] [Medline: 29872667]
21. Freeman JL, Caldwell PHY, Scott KM. How adolescents trust health information on social media: a systematic review. Acad Pediatr 2023;23(4):703-719 [FREE Full text] [doi: 10.1016/j.acap.2022.12.011] [Medline: 36581098]

XSL•FO
RenderX

22. Trivedi N, Krakow M, Hyatt Hawkins K, Peterson EB, Chou WS. “Well, the message is from the institute of something”: exploring source trust of cancer-related messages on simulated Facebook posts. Front Commun 2020;5:12. [doi: 10.3389/fcomm.2020.00012]
23. Bautista JR, Zhang Y, Gwizdka J. Healthcare professionals' acts of correcting health misinformation on social media. Int J Med Inform 2021;148:104375. [doi: 10.1016/j.ijmedinf.2021.104375] [Medline: 33461008]
24. Al-Uqdah L, Franklin FA, Chiu C, Boyd BN. Associations between social media engagement and vaccine hesitancy. J Community Health 2022;47(4):577-587 [FREE Full text] [doi: 10.1007/s10900-022-01081-9] [Medline: 35332393]
25. Petty RE, Brinol P. The elaboration likelihood model. In: Handbook of Theories of Social Psychology. London: Sage Publication Ltd; 2011.
26. Ahmad Rizal AR, Nordin SM, Ahmad WFW, Ahmad Khiri MJ, Hussin SH. How does social media influence people to get vaccinated? The elaboration likelihood model of a person's attitude and intention to get COVID-19 vaccines. Int J Environ Res Public Health 2022;19(4):2378 [FREE Full text] [doi: 10.3390/ijerph19042378] [Medline: 35206563]
27. Saini V, Liang L, Yang Y, Le HM, Wu C. The association between dissemination and characteristics of pro-/anti-COVID-19 vaccine messages on Twitter: application of the elaboration likelihood model. JMIR Infodemiology 2022;2(1):e37077 [FREE Full text] [doi: 10.2196/37077] [Medline: 35783451]
28. Wang Q, Zhou F, Zhang W, Tang C. A study of parental decision-making over the vaccination of girls, based on the protection motivation theory and the elaboration likelihood model. Front Public Health 2022;10:1024399 [FREE Full text] [doi: 10.3389/fpubh.2022.1024399] [Medline: 36438257]
29. Cartmell KB, Mzik CR, Sundstrom BL, Luque JS, White A, Young-Pierce J. HPV vaccination communication messages, messengers, and messaging strategies. J Cancer Educ 2019;34(5):1014-1023 [FREE Full text] [doi: 10.1007/s13187-018-1405-x] [Medline: 30054900]
30. Nan X, Dahlstrom MF, Richards A, Rangarajan S. Influence of evidence type and narrative type on HPV risk perception and intention to obtain the HPV vaccine. Health Commun 2015;30(3):301-308. [doi: 10.1080/10410236.2014.888629] [Medline: 25061715]
31. Petty R, Cacioppo J. Communication and Persuasion: Central and Peripheral Routes to Attitude Change. New York: Springer Science & Business Media; 2012.
32. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In: Communication and Persuasion. New York: Springer; 1986:1-24.
33. AmeriSpeak Teen.: NORC at the University of Chicago URL: https://amerispeak.norc.org/us/en/amerispeak/our-capabilities/amerispeak-teen.html [accessed 2025-06-04]
34. Grunwald B, Nyarko J, Rappaport J. Police agencies on Facebook overreport on Black suspects. Proc Natl Acad Sci U S A 2022;119(45):e2203089119 [FREE Full text] [doi: 10.1073/pnas.2203089119] [Medline: 36322743]
35. Punziano G, Marrazzo F, Acampa S. An application of content analysis to CrowdTangle data: the 2020 Constitutional Referendum campaign on Facebook. Curr Polit Econ Eur 2021;32(4):371 [FREE Full text]
36. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
37. NVivo. Lumivero. URL: https://lumivero.com/products/nvivo/ [accessed 2026-04-28]
38. Cole R. Inter-rater reliability methods in qualitative case study research. Sociol Methods Res 2024;53(4):1944-1975 [FREE Full text] [doi: 10.1177/00491241231156971]
39. Freeman JL, Caldwell PH, Scott KM. The role of trust when adolescents search for and appraise online health information. J Pediatr 2020;221:215-223. [doi: 10.1016/j.jpeds.2020.02.074] [Medline: 32446485]

Abbreviations

AI: artificial intelligence
CDC: Centers for Disease Control and Prevention
ELM: Elaboration Likelihood Model
HPV: human papillomavirus
IRB: institutional review board
NORC: National Opinion Research Center

XSL•FO
RenderX

Edited by M Balcarras; submitted 12.Aug.2025; peer-reviewed by L Shrestha, J Guo, F Adjei, Y Oyende; comments to author 05.Nov.2025; revised version received 18.Feb.2026; accepted 24.Feb.2026; published 03.Jun.2026.

Please cite as:
Trivedi N, Van Vleet R, Eid C, Burke-Garcia A, Leader AE
The Influence of Social Media Messaging on Human Papillomavirus Vaccine Attitudes and Confidence Among Adolescent Males: Group Discussion Study
JMIR Cancer 2026;12:e82210
URL: https://cancer.jmir.org/2026/1/e82210
doi:10.2196/82210
PMID:

©Neha Trivedi, Rachel Van Vleet, Czesia Eid, Amelia Burke-Garcia, Amy E Leader. Originally published in JMIR Cancer (https://cancer.jmir.org), 03.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI-Driven Patient Screening for Clinical Trials in Pancreatic Cancer: The PANCR-AI Pilot Retrospective Comparative Study

Arthur Claessens[1,2], MD, MPH; Alizée Simon[2], MD; Agathe Manchart[2], MD; Laura Desmonts[2], MD; Cassandre Leguay[2], MD; Jean-Christophe Faivre[3], MD; Sébastien Cambier[4], MSc; Julia Salleron[4], MSc; Aurélien Lambert[2,5], MD, PhD

[1]Université de Lorraine, Nancy, France

[2]Medical Oncology Department, Institut de Cancérologie de Lorraine, 6 Avenue de Bourgogne, Vandoeuvre-Lès-Nancy, France

[3]Radiotherapy Department, Institut de Cancérologie de Lorraine, Vandoeuvre-Lès-Nancy, France

[4]Biostatistics Department, Institut de Cancérologie de Lorraine, Vandoeuvre-Lès-Nancy, France

[5]Interdisciplinarité en Santé Publique, Interventions & Instruments de mesure complexes – Région Est, INSPIIRE UMR 1319, INSERM, Université de Lorraine, Nancy, France

Corresponding Author:
Arthur Claessens, MD, MPH
Université de Lorraine, Nancy, France

Abstract

Background: Screening for clinical trials is challenging for clinicians due to its time-consuming and repetitive nature. The rise of artificial intelligence (AI) offers an opportunity to improve screening productivity and reproducibility. Pancreatic cancer is characterized by increasing incidence, poor survival outcomes, and an urgent need for improved management strategies.

Objective: This study aimed to assess the performance of AI in evaluating clinical trial inclusion and exclusion criteria, compared to a double-blind human gold standard, using a retrospective cohort.

Methods: In the PANCR-AI (Pancreatic Cancer Retrospective Screening with Artificial Intelligence) pilot study, we retrospectively reviewed cases from our institutional database of patients with advanced pancreatic cancer presented at tumor board meetings between January 2018 and December 2023. Each patient was screened for clinical trials open for inclusion at the time of the multidisciplinary meeting. Manual screening of eligibility criteria for each patient-trial pair was performed by 2 blinded oncologists to determine potential eligibility (gold standard), with a third oncologist resolving discrepancies. Potential eligibility was also assessed using 3 large language models (ie, GPT-4.5, Claude 3.7 Sonnet, and Mistral-7B-Instruct v0.3). Their performance was compared to the human gold standard using standard evaluation metrics (eg, sensitivity, specificity, precision, recall, and F_1-score). Correlations between the risk of failure and the number of words and characters in the criteria were analyzed. The time required to complete the screening was recorded for both human and AI assessments. The number of trials open for enrollment at the time of the tumor board meeting was also recorded as a variable for analysis.

Results: Across 341 patient-trial pairs, the AI models demonstrated high sensitivity, ranging from 83.3% to 92.2%. Analysis of the criteria showed a correlation between the risk of failure and the number of words and the number of characters in the criteria. Overall screening time for manual assessment was significantly longer for the human gold standard (44.70 hours) assessment than for AI (2.53-3.15 hours). Patients were more likely to have been included in a clinical trial if the number of trials open for enrollment was higher at the time of the tumor board meeting (P=.02).

Conclusions: Our study highlights the promising performance of AI in clinical trial screening. Future work should explore integration with structured clinical data, such as laboratory values or radiological findings, to improve multimodal comprehension. Expanding the evaluation to a broader range of tumor types and multicenter datasets would improve generalizability. Finally, real-time prospective validation and workflow integration with electronic health records will be critical to assess the feasibility and clinical impact of large language model–assisted screening in daily oncology practice. Addressing these challenges will be essential to move from proof of concept to scalable clinical implementation.

(JMIR Cancer 2026;12:e80268) doi:10.2196/80268

KEYWORDS

pancreatic neoplasms; artificial intelligence; AI; reproducibility of results; clinical trials as topic; pilot project

XSL•FO
RenderX

Introduction

Pancreatic cancer is a leading cause of cancer-related mortality, and its incidence continues to rise [1]. In advanced and metastatic cancer, therapeutic options are often limited, with poor results for overall survival in comparison with other malignancies [2,3]. There is a critical need to gather knowledge about the different subtypes of the disease and to adopt new approaches to improve patient outcomes. Clinical trials are the preferred option to ensure access to innovation. Screening patients for clinical trials is a complex and time-consuming procedure. Screening electronic patient records requires reviewing all inclusion and exclusion criteria for each potential clinical trial for which a patient might be eligible. This effort becomes obsolete as soon as new elements are added to the patient medical record. This assessment at a given moment does not predict short- or medium-term clinical evolution, particularly in the case of patients with locally advanced and metastatic pancreatic adenocarcinomas. The multidisciplinary tumor board meeting is held at a critical moment in the therapeutic decision-making process and represents an opportunity to suggest the inclusion of a patient in a clinical trial before treatment begins.

Promising artificial intelligence (AI) tools, such as assistant agents, are increasingly being incorporated into hospital information systems, and their applications now cover a wide range of needs expressed by health care professionals [4]. The emergence of generative AI could help physicians to screen patients by performing automated, real-time analysis of their electronic medical records to determine whether the inclusion and exclusion criteria for a clinical trial are met. This real-time screening would increase the chances of offering the patient a relevant clinical trial.

Screening for clinical trials at the time of the pretherapeutic tumor board meetings is of major interest for the prognosis of pancreatic malignancies. Current efforts in this area are hampered by a lack of human resources. Among the possible factors leading to the failure of clinical trials, the small number of patients included is one of the main causes of trial discontinuation [5]. Only an estimated 2% to 4% of all patients with cancer participate in clinical trials. Although multiple factors explain this low enrollment rate, a significant benefit has been observed with the employment of dedicated medical resources to screen patients, leading to an increase in enrollment. However, this possibility is limited by the availability of human resources [6]. Manual review of patient records for clinical trial eligibility represents a significant burden on research teams and medical oncologists. The manual extraction of this information is time consuming, labor-intensive, expensive, and error prone.

The information contained in electronic medical records is heterogeneous and complex. Extracting meaningful information from medical records presents significant challenges owing to the unstructured nature of these data, which limits their immediate applicability for clinical research [7]. As with other malignancies, information about patients with pancreatic cancer is primarily available in free-text form, which requires sophisticated processing. Rule-based extraction techniques are limited by their rigidity and the need for postprocessing. Machine learning and deep learning techniques show promise for recognizing word relationships; however, they require postprocessing for complex data and do not understand implicit or semantic information [8].

To address these limitations and challenges, we propose the use of a strategy for medical text mining with the help of large language models (LLMs) to convert text into structured computerized data using informatic extraction technologies. LLMs are high-performance technologies for classifying, filtering, and clustering information, enabling the direct extraction of information without additional postprocessing. These technologies can extract textual data from medical records and reorganize them for use in oncology research [9,10]. These technologies could also improve access to clinical trials in secondary centers or pediatric settings, with the expectation of improving health care democracy [11].

In PANCR-AI (Pancreatic Cancer Retrospective Screening with Artificial Intelligence) study, we assessed the performance of 3 different LLMs (GPT-4.5, OpenAI; Claude 3.7 Sonnet, Anthropic; and Mistral-7b-Instruct v0.3, Mistral) for clinical trial screening using a retrospective cohort of patients with metastatic or locally advanced pancreatic cancer [12-14].

Methods

Study Context and Population

The Institut de Cancérologie de Lorraine Comprehensive Cancer Center prospectively maintains a database of pancreatic malignancies. This database was queried to identify patients with newly diagnosed locally advanced, unresectable, or metastatic pancreatic adenocarcinoma. All patients with a pretherapeutic tumor board meeting report available in their electronic medical records between January 2018 and December 2023 were included in the study. Patients with other subtypes of pancreatic malignancies, such as intraductal papillary mucinous neoplasms or neuroendocrine tumors, were excluded. Each patient had a pretherapeutic tumor board meeting report summarizing their clinical and pathological status. For each patient, “candidate trials” were defined as all clinical trials open for enrollment in advanced or metastatic pancreatic cancer at the time of the tumor board meeting. Each possible match between a patient and a candidate trial constituted a patient-trial pairing, representing one evaluation of eligibility according to the trial’s inclusion and exclusion criteria. Clinical and pathological variables were extracted from the database and analyzed, including demographic characteristics, tumor pathology features, treatments received, and follow-up information. The study was registered with the French Health Data Hub on February 3, 2025 (22286214).

Study Design and Data Source

This retrospective single-center analysis was based on pretherapeutic tumor board meeting reports, which provided a structured summary of each patient’s clinical condition and characteristics as documented in their medical records. These reports served as a common database for both oncologists and LLMs.

XSL•FO
RenderX

The inclusion and exclusion criteria were extracted from the sponsor's trial protocol and reformatted into standardized textual items for analysis. Criteria related to biological samples, life expectancy, contraceptive use, or ethical consent were removed from the screening process due to the absence of corresponding data in the pretherapeutic tumor board meeting reports. The remaining criteria were independently reviewed and categorized according to their type (eg, clinical, pathological, or demographic) and analyzed for their textual characteristics, including word and character counts. The number of inclusion criteria ranged from 1 to 7, and the number of exclusion criteria ranged from 2 to 19. The characteristics of the 12 clinical trials are summarized in Table 1. Each criterion was evaluated for its number of characters, nature, and category of variables.

Table . Characteristics of the 12 candidate clinical trials open for recruitment (January 2018 to December 2023) and associated patient-trial pairs (N=341).

Clinical trials	ClinicalTrials.gov registry number	Phase	Sponsor type	Patients historically included in the trial (N)	Candidate patients for the trial (N)	Inclusion criteria analyzed, n (%)	Exclusion criteria analyzed, n (%)
Actuate 1801 - PDAC	NCT03678883	II	Industry	0	2	7 (58)	9 (75)
ALIX	NCT03974854	II	Health center	1	40	4 (57)	16 (84)
APACaP D-13	NCT02184663	N/A[a]	Medical association	1	43	5 (56)	6 (75)
AVENGERS 500 (PANC003)	NCT03504423	III	Industry	9	22	6 (50)	19 (76)
MAZEPPA GERCOR D19-02 PRODIGE-72	NCT04348045	II	Medical association	1	12	6 (35)	14 (88)
ONCOSNIPE PANCREAS	NCT04548960	II	Industry	0	12	1 (100)	2 (100)
OPTIMIZE-01	NCT04888312	Ib/II	Industry	3	21	6 (46)	10 (53)
PANDAS PRODIGE-44	NCT02676349	II	Health center	1	38	5 (42)	7 (78)
STEMNESS-PANC	NCT03721744	II/III	Industry	4	19	7 (47)	12 (92)
TEDOPaM study	NCT03806309	II	Medical association	7	55	4 (31)	20 (95)
D17-01 - PRODIGE 63							
URGENCE PANCREAS	NCT02979483	N/A	Medical association	5	44	3 (38)	3 (100)
EPIC	NCT02853474	III	Health center	0	33	5 (63)	2 (50)

[a]N/A: not applicable.

The inclusion and exclusion criteria for each candidate trial were evaluated on a criterion-by-criterion basis using information extracted from the pretherapeutic tumor board meeting reports. For each patient-trial pairing, fulfillment of every inclusion and exclusion criterion was assessed independently by the human gold standard and by each LLM. When the available clinical information in the reports was insufficient to assess a given criterion, or when it was not possible to reach a conclusion due to insufficiently consistent or contradictory information based on the tumor board meeting report, that criterion was considered to be met if it was an inclusion criterion and to be unmet if it was an exclusion criterion. This conservative assumption ensured that such patients were classified as eligible, reflecting a sensitivity-oriented approach consistent with the intended practical use of LLM-derived outputs. At the patient-trial level, the individual criterion assessments were then aggregated to determine the overall eligibility decision for each pairing.

Performance Evaluation of LLMs

Three LLMs (ie, GPT-4.5, Claude 3.7 Sonnet, and Mistral-7B-Instruct v0.3) were evaluated for their ability to assess patient eligibility for each inclusion and exclusion criterion of all candidate trials based on the information contained in the pretherapeutic tumor board meeting reports. LLMs were queried through the online interfaces and respective application programming interfaces of OpenAI, Anthropic, and Mistral. The LLMs' unique prompt design was developed according to European Medicines Agency guidelines and contextual prompt engineering methodology (role assignment, step-by-step reasoning, and constrained outputs) to ensure the reproducibility of the output for the same input. The input prompt is available in Multimedia Appendix 1 [15,16]. All

XSL•FO
RenderX

LLMs were given the same instructions and data to be analyzed as the medical oncologists (for the human gold standard). Results are displayed in a dedicated radar plot. The screening time for each LLM was determined between the time each tumor board meeting was submitted and the time of completion of the output for each trial-patient assessment, measured using digital tools.

Gold Standard Assessment

AC, AS, AM, LD, and CL are medical oncologists who participated in establishing the gold standard and volunteered to collect the data. Participation was proposed to a panel of physicians who were not involved in the care of the patients whose records were analyzed in this study to minimize potential bias in their assessments. The information provided for analysis included the inclusion and exclusion criteria for clinical trials and the deidentified tumor board meeting reports. The data were collected in a deidentified database stored on a secure internal server at the Comprehensive Cancer Center. All reports were reviewed by 2 medical oncologists to ensure reproducibility of the human assessment. Each oncologist was blinded to the evaluation of the other. In case of disagreement, a third medical oncologist independently reviewed the case to reach a consensus. Interrater reliability was assessed using Cohen κ coefficient. The screening time was reported by each medical oncologist for the determination of the gold standard and measured using digital tools. This duration was compiled and allocated to each corresponding candidate trial.

Statistical Analysis

Data analyses were performed using SAS (version 9.4; SAS Institute Inc). $P<.05$ was considered statistically significant. Given the exploratory nature of the study, no adjustment for multiple testing was applied.

Quantitative data are reported as median values unless otherwise specified. Comparisons between 2 groups were performed using the Mann-Whitney test, and comparisons among more than 2 groups were performed using the Kruskal-Wallis test. Categorical variables were compared using the chi-square test. Survival was estimated using the Kaplan-Meier method, and survival curves were compared with the log-rank test.

For each inclusion and exclusion criterion, and subsequently at the patient-trial level, the LLM outputs were compared with the human gold standard. The results were classified as true positives when both methods identified eligibility; true negatives when both identified noneligibility; false positives when the LLM indicated eligibility, whereas the human gold standard did not; and false negatives when the LLM indicated noneligibility, whereas the human gold standard indicated eligibility. Cases in which the LLM and the human reviewers provided divergent results were considered as discordant.

LLM performance was evaluated using standard metrics (eg, recall, sensitivity, specificity, precision, and F_1-score) to assess their ability to determine patient eligibility both at the level of individual inclusion criteria and across all clinical trials screened; sensitivity and specificity were computed with 95% CIs. Pairwise comparisons of LLM sensitivities (or specificities) were performed using McNemar test among cases identified as eligible (or ineligible) by the human gold standard.

Statistical analyses were conducted at the criterion level to assess factors associated with LLM-human discordance. Variables with $P<0.1$ in bivariate analyses were included in the multivariate logistic regression model.

Ethical Considerations

The PANCR-AI study was conducted at the Institut de Cancérologie de Lorraine (Vandœuvre-lès-Nancy, France). It complies with the French reference methodology MR-004 of the French Data Protection Authority (Commission Nationale de l'Informatique et des Libertés) for research involving human data. According to the MR-004 regulatory framework, ethics committee approval and individual informed consent were not required, as the study involved retrospective analyses of deidentified clinical data and had no impact on patient care or management. For trial involvement, patients were evaluated using the standard manual screening process at the Institut Cancérologie de Lorraine. The reports were deidentified by retaining only age and gender, modifying dates to preserve the temporal relationship with the tumor board meeting, and implementing measures to ensure patient anonymity. The physicians who volunteered to collect the data were medical oncologists practicing at a comprehensive cancer center, and thus consent was implied.

Results

Study Population

Among the 141 patients initially identified from the Institut de Cancérologie de Lorraine Comprehensive Cancer Center database, 85 (60%) patients with locally advanced or metastatic pancreatic adenocarcinoma were included in this analysis; 46 (32.6%) patients were excluded due to unavailable pretherapeutic data in the tumor board meeting report, and 10 (7%) patients were excluded from the analysis based on histological type (all neuroendocrine tumors). The flowchart of eligibility is available in Multimedia Appendix 2. Overall, 30 (35.2%) patients of this cohort were historically included in a clinical trial. The median age was 68 (IQR 32 - 89) years, and 49% (42/85) of patients were male. At the time of the tumor board meeting, 66 (78%) patients presented with metastatic disease, while 19 (22%) patients had locally advanced unresectable tumors. Patients were more likely to have been included in a clinical trial if the number of trials open for enrollment was higher at the time of the tumor board meeting ($P=.02$). Overall survival did not differ according to patients' participation in clinical trials ($P=.34$). No significant association was observed between clinical trial inclusion and age ($P=.09$) or personal history of previous malignancy ($P=.25$). The baseline characteristics of the study population are summarized in Table 2.

XSL•FO
RenderX

Table . Characteristics of patients from the cohort of locally advanced and metastatic pancreatic cancer, according to inclusion or noninclusion in a clinical trial (N=341).

Characteristics, baseline, and demographics	All patients (n=85)	Patient historically included in the clinical trial (n=30)	Patient not included in the clinical trial (n=55)	*P* value
Age (years), median (IQR)	68 (60 - 73)	62 (56 - 67.25)	70 (61.25 - 77.5)	.09
Female, n (%)	43 (51)	16 (53)	27 (49)	.82
Overall survival (months), median (95% CI)	9 (7-15)	14 (8 - 18)	7 (5-13)	.34
Locally advanced pancreatic cancer, n (%)	19 (22)	5 (16)	14 (25)	.49
Candidate trials per patient, median (IQR)	4 (3 - 5)	5 (4-6)	4 (2.5 - 5)	.02
ECOG[a], median (IQR)	1 (1 - 2)	1 (1 - 2)	1 (1 - 2)	.86
History of malignancies, n (%)	7 (0.08)	1 (0.03)	6 (0.11)	.25
Cardiovascular comorbidities, n (%)	40 (47)	16 (53)	24 (43)	.53
BMI at diagnosis (kg/m^2), median (IQR)	23.53 (20.96 - 25.88)	24.01 (20.53 - 27.12)	23.20 (21.36 - 25.29)	.58

[a]ECOG: Eastern Cooperative Oncology Group Performance Status Scale.

Gold Standard Assessment

Manual assessment of eligibility was performed as the gold standard for this analysis. Interrater reliability for manual assignment was substantial, with a Cohen κ value of 0.83 between reviewers (95% CI 0.81-0.84). Seven (0.06%) discordant cases required the intervention of a third reviewer to reach consensus.

Performance of LLMs

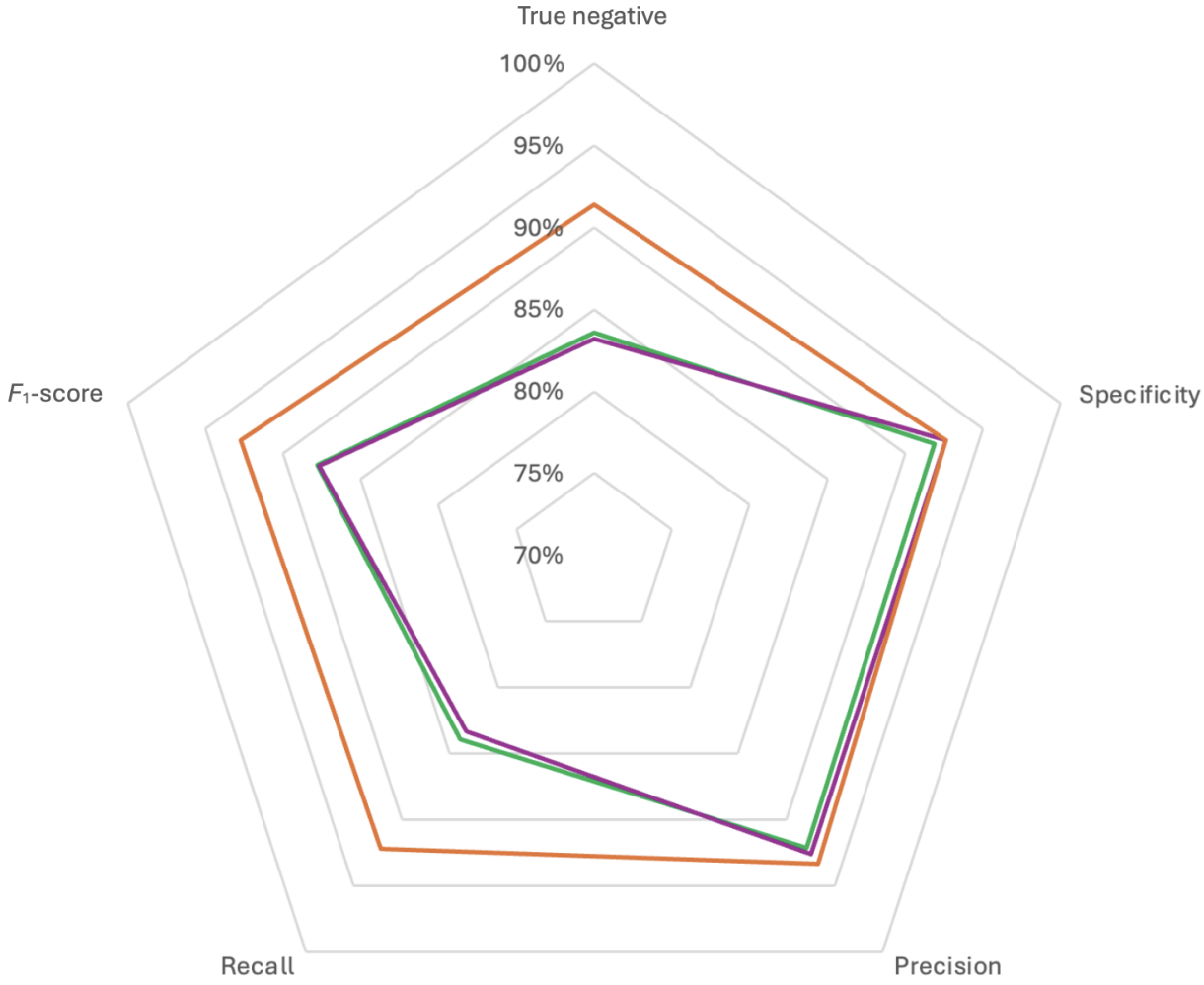
With a median of 4 (IQR 3-5) candidate trials per patient among the 85 patients analyzed, a total of 341 patient-trial pairs were identified. The overall performance of the 3 LLMs in evaluating patient eligibility for clinical trials is presented in Table 3. Detailed performance results of all LLMs for each trial are provided in Multimedia Appendix 3. Mistral-7b-Instruct v0.3 demonstrated the highest sensitivity, followed by GPT-4.5 and Claude 3.7 Sonnet (92.2%, 83.9%, and 83.3%, respectively), with a statistically significant difference between the Mistral-7b-Instruct v0.3 and the other 2 models (P<.01 for both comparisons). Overall performance metrics for each LLM are reported in Figure 1.

In clinical trials, analysis of the criteria revealed that the frequency of discordance between LLMs and human reviewers increased with the number of words and the number of characters in the criteria (P=.02 and P<.01, respectively). Both associations remained significant in a multivariable logistic regression model. The other characteristics of the criteria were not significantly associated with the risk of LLM error.

The total screening time for the gold standard manual assessment was 2682 minutes (ie, 44.70 hours). The total screening time using LLMs ranged from 152 minutes (2.53 hours) for Mistral-7B-Instruct v0.3 to 189 minutes (3.15 hours) for Claude 3.7 Sonnet. Screening time was significantly shorter for all LLMs compared to the gold standard assessment. Detailed screening time data are reported in Multimedia Appendix 4.

XSL•FO
RenderX

Figure 1. Radar plot of large language model performance compared to the human gold standard.



	True negative	Specificity	Precision	Recall	F_1-score
ChatGPT-4.5	83.6%	91.9%	92.1%	83.9%	87.8%
Mistral-7b-Instruct v0.3	91.4%	92.6%	93.3%	92.2%	92.7%
Claude-Sonnet-3.7	83.2%	92.6%	92.6%	83.3%	87.7%

Table . Overall performance of large language models in selecting patients for clinical trials compared to the human gold standard based on the patient-trial pairs (N=341).

All	Number of true positive pairs	Number of false negative pairs	Number of true negative pairs	Number of false positive pairs	Sensitivity (95% CI)	Specificity (95% CI)
GPT-4.5	151	13	148	29	83.9% (78.5% - 89.3%)	91.9% (87.7% - 96.1%)
Claude 3.7 Sonnet	150	12	149	30	83.3% (77.9% - 88.8%)	92.6% (88.5% - 96.6%)
Mistral-7b-Instruct v0.3	166	12	149	14	92.2%[a] (88.3% - 96.1%)	92.6% (88.5% - 96.6%)

[a]Statistically higher sensitivity.

XSL•FO
RenderX

Discussion

Main Findings

This study shows promising performance of LLMs for the prescreening of patients with locally advanced or metastatic pancreatic cancer based on pretherapeutic tumor board meeting reports. Screening for clinical trials assisted by LLMs would be appropriate in cases where there are a large number of inclusion and exclusion criteria or in peripheral centers with a view to referring patients to expert centers for a specific condition, thereby offering them the best chance of benefiting from therapeutic innovation.

The disparities in results between clinical trials using LLMs are difficult to interpret because of the small number of patients analyzed per trial, but overall, the risk of LLM error was correlated with the number of words or characters contained in a single inclusion or exclusion criterion. LLMs displayed poor performance in interpreting data, such as identifying clinically relevant past medical events or recent but resolved medical conditions. This aspect represents a major area for improvement in these technologies. No errors were observed in terms of timing between the discussion of the case in tumor board meetings and an event of particular interest. The identification of characteristics such as age, personal (rather than family) history of cancer, pregnancy status, or metastasis location was not a major cause of failure for LLMs. This is reassuring regarding the feasibility of using LLMs for the purpose of prescreening patients for clinical trials.

Mistral-7b-Instruct v0.3 is an LLM trained using multiple French-language databases and does not require translation when generating responses in French, unlike Anglo-Saxon models. Its optimization for Latin languages enables it to understand complex syntactic structures and semantic nuances. This characteristic may help explain the superior results obtained by this LLM, which outperformed the other 2 solutions tested.

Comparison to Prior Work

Recent developments in AI and the rise of LLMs have offered promising opportunities to address these challenges. The ability of these models to process large quantities of complex patient data quickly and reliably has already demonstrated benefits in the field of oncology research [17]. Studies have shown that AI-based approaches significantly reduce the time required for screening eligibility, without compromising the consistency of trial-patient identification [18,19]. By integrating AI into clinical workflows, physicians could spend more time on direct patient interactions, complex medical decision-making, and personalized care, thereby enhancing overall health care quality and potentially improving patient outcomes [20,21]. Our study concurs with previous research, demonstrating the clear relevance of these tools in this setting [22]. A recent review focusing on the application of AI to clinical trial recruitment demonstrated several positive outcomes, including increased efficiency and improved recruitment or accuracy. It also emphasized various technical and ethical issues, including privacy, data security, transparency, and selection bias [23]. The first recent multicenter study using AI to streamline screening processes in patient medical records, conducted on a cohort of patients with lung cancer, found that LLMs achieved promising performance across 28 points of interest, including metastatic status and comorbidities. This study also highlighted that LLMs trained for screening make fewer errors than humans, although the contribution of humans to complement automated screening in cases where LLMs returned results with lower levels of confidence reduced the risk of error. The use of AI would also reduce variability between centers in identifying key features for clinical trial inclusion [24].

Strengths and Limitations

Although the reasoning capabilities of generative AI models for clinical trial screening have shown impressive results, their implementation in clinical practice remains a challenge. The heterogeneity of software, data formats, extraction methods, and ethical considerations poses challenges to the routine use of these tools by health care providers.

Furthermore, because the analysis is based on a snapshot of the patient's disease history, it is possible that some information may evolve over time, leading to misjudgment of certain inclusion or exclusion criteria for clinical trial screening. A significant number of tumor board meeting reports were missing due to the recent digitization of patient records. This study was also unable to use certain scanned documents of poor quality.

Given the limited number of patients and clinical trials, as well as the retrospective and single-center design of this study, the results should be interpreted with caution. However, they suggest a potentially valuable contribution of generative AI tools in screening patients with locally advanced or metastatic pancreatic cancer for their eligibility for clinical trials.

Future Directions

Refining prompt engineering and fine-tuning models with domain-specific datasets could further enhance the accuracy and specificity of LLMs. Future work should explore integration with structured clinical data, such as laboratory values or radiological findings, to improve multimodal comprehension. Expanding the evaluation to a broader range of tumor types and multicenter datasets would improve generalizability.

The next step is to ensure outputs with clinical-grade standard requirements. Establishing a robust data governance framework to safeguard patient privacy, ensure data security, and maintain traceability of LLM actions could help to enhance acceptance of LLMs by regulatory agencies for real-time screening within electronic health records systems.

Finally, real-time prospective validation and workflow integration with electronic health records will be critical to assess the feasibility and clinical impact of LLM-assisted screening in daily oncology practice. Addressing these challenges will be essential to move from proof of concept to scalable and ethically sound clinical implementation.

XSL•FO
RenderX

Acknowledgments

Generative artificial intelligence was not used in any portion of the manuscript conception or writing.

Funding

The authors declare that no financial support was received for the research, authorship, and/or publication of this article.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AC was involved in the conception of the work, acquisition and interpretation of data, and drafting of the manuscript. AS, AM, LD, and CL contributed to data acquisition and critically revised the manuscript for important intellectual content. SC and JS contributed to the methodology, performed the statistical analyses, and critically revised the manuscript for important intellectual content. JCF critically revised the manuscript for important intellectual content. AL contributed to the study conception and critically revised the manuscript for important intellectual content. All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work, ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Prompt used as a framework for analyzing clinical trial criteria for patients by both the human gold standard and large language models.
[PDF File, 30 KB - cancer_v12i1e80268_app1.pdf]

Multimedia Appendix 2
Flowchart of patient eligibility.
[PDF File, 19 KB - cancer_v12i1e80268_app2.pdf]

Multimedia Appendix 3
Overall and detailed performance of large language models for screening patients for each candidate trial.
[PDF File, 74 KB - cancer_v12i1e80268_app3.pdf]

Multimedia Appendix 4
Screening time per clinical trial (minutes).
[PDF File, 43 KB - cancer_v12i1e80268_app4.pdf]

References

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin 2025;75(1):10-45. [doi: 10.3322/caac.21871] [Medline: 39817679]
2. Conroy T, Desseigne F, Ychou M, et al. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. N Engl J Med 2011 May 12;364(19):1817-1825. [doi: 10.1056/NEJMoa1011923] [Medline: 21561347]
3. Von Hoff DD, Ervin T, Arena FP, et al. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. N Engl J Med 2013 Oct 31;369(18):1691-1703. [doi: 10.1056/NEJMoa1304369] [Medline: 24131140]
4. Maleki Varnosfaderani S, Forouzanfar M. The role of AI in hospitals and clinics: transforming healthcare in the 21st century. Bioengineering (Basel) 2024 Mar 29;11(4):337. [doi: 10.3390/bioengineering11040337] [Medline: 38671759]
5. Stensland KD, McBride RB, Latif A, et al. Adult cancer clinical trials that fail to complete: an epidemic? J Natl Cancer Inst 2014 Sep;106(9):dju229. [doi: 10.1093/jnci/dju229] [Medline: 25190726]
6. Chen L, Grant J, Cheung WY, Kennecke HF. Screening intervention to identify eligible patients and improve accrual to phase II-IV oncology clinical trials. J Oncol Pract 2013 Jul;9(4):e174-e181. [doi: 10.1200/JOP.2012.000763] [Medline: 23942936]
7. Landolsi MY, Hlaoua L, Ben Romdhane L. Information extraction from electronic medical documents: state of the art and future research directions. Knowl Inf Syst 2023;65(2):463-516. [doi: 10.1007/s10115-022-01779-1] [Medline: 36405956]
8. Durango MC, Torres-Silva EA, Orozco-Duque A. Named entity recognition in electronic health records: a methodological review. Healthc Inform Res 2023 Oct;29(4):286-300. [doi: 10.4258/hir.2023.29.4.286] [Medline: 37964451]

XSL•FO
RenderX

9. Wang L, Luo L, Wang Y, Wampfler J, Yang P, Liu H. Natural language processing for populating lung cancer clinical research data. BMC Med Inform Decis Mak 2019 Dec 5;19(Suppl 5):239. [doi: 10.1186/s12911-019-0931-8] [Medline: 31801515]
10. Cho H, Yoo S, Kim B, et al. Extracting lung cancer staging descriptors from pathology reports: a generative language model approach. J Biomed Inform 2024 Sep;157:104720. [doi: 10.1016/j.jbi.2024.104720] [Medline: 39233209]
11. Ni Y, Wright J, Perentesis J, et al. Increasing the efficiency of trial-patient matching: automated clinical trial eligibility pre-screening for pediatric oncology patients. BMC Med Inform Decis Mak 2015 Apr 14;15:28. [doi: 10.1186/s12911-015-0149-3] [Medline: 25881112]
12. ChatGPT (version 45). OpenAI. URL: https://chat.openai.com/ [accessed 2025-05-27]
13. Claude 37 sonnet. Anthropic. 2025. URL: https://www.anthropic.com/news/claude-3-7-sonnet [accessed 2025-05-27]
14. Jiang AQ, Sablayrolles A, Mensch A, et al. Mistral 7B. arXiv. Preprint posted online on Oct 10, 2023. [doi: 10.48550/arXiv.2310.06825]
15. Guiding principles on the use of large language models in regulatory science and for medicines regulatory activities. European Medicines Agency. 2024. URL: https://www.ema.europa.eu/en/documents/other/guiding-principles-use-large-language-models-regulatory-science-medicines-regulatory-activities_en.pdf [accessed 2025-10-21]
16. Giray L. Prompt engineering with ChatGPT: a guide for academic writers. Ann Biomed Eng 2023 Dec;51(12):2629-2633. [doi: 10.1007/s10439-023-03272-4] [Medline: 37284994]
17. Haddad T, Helgeson JM, Pomerleau KE, et al. Accuracy of an artificial intelligence system for cancer clinical trial eligibility screening: retrospective pilot study. JMIR Med Inform 2021 Mar 26;9(3):e27767. [doi: 10.2196/27767] [Medline: 33769304]
18. Beck JT, Rammage M, Jackson GP, et al. Artificial intelligence tool for optimizing eligibility screening for clinical trials in a large community cancer center. JCO Clin Cancer Inform 2020 Jan;4:50-59. [doi: 10.1200/CCI.19.00079] [Medline: 31977254]
19. Gédor M, Desandes E, Chesnel M, et al. Développement d'une solution d'intelligence artificielle pour améliorer le screening en recherche clinique. Bull Cancer 2024 May;111(5):473-482. [doi: 10.1016/j.bulcan.2024.01.010]
20. Alexander M, Solomon B, Ball DL, et al. Evaluation of an artificial intelligence clinical trial matching system in Australian lung cancer patients. JAMIA Open 2020 Jul;3(2):209-215. [doi: 10.1093/jamiaopen/ooaa002] [Medline: 32734161]
21. Calaprice-Whitty D, Galil K, Salloum W, Zariv A, Jimenez B. Improving clinical trial participant prescreening with artificial intelligence (AI): a comparison of the results of AI-assisted vs standard methods in 3 oncology trials. Ther Innov Regul Sci 2020 Jan;54(1):69-74. [doi: 10.1007/s43441-019-00030-4] [Medline: 32008227]
22. Jin Q, Wang Z, Floudas CS, et al. Matching patients to clinical trials with large language models. Nat Commun 2024 Nov 18;15(1):9074. [doi: 10.1038/s41467-024-53081-z] [Medline: 39557832]
23. Wang K, Cui H, Zhu Y, et al. Evaluation of an artificial intelligence-based clinical trial matching system in Chinese patients with hepatocellular carcinoma: a retrospective study. BMC Cancer 2024 Feb 22;24(1):246. [doi: 10.1186/s12885-024-11959-7] [Medline: 38388861]
24. Aldea M, Zullo L, Levrat V, et al. Next-generation multicenter studies: using artificial intelligence to automatically process unstructured health records of patients with lung cancer across multiple institutions. Ann Oncol 2025 Dec 15:S0923-7534(25)06320-3. [doi: 10.1016/j.annonc.2025.12.006] [Medline: 41407195]

Abbreviations

AI: artificial intelligence
LLM: large language model
PANCR-AI: Pancreatic Cancer Retrospective Screening with Artificial Intelligence

Edited by M Balcarras, N Cahill; submitted 09.Jul.2025; peer-reviewed by Á García-Barragán, W Fu, Z Li; revised version received 18.Dec.2025; accepted 19.Dec.2025; published 23.Feb.2026.

Please cite as:
Claessens A, Simon A, Manchart A, Desmonts L, Leguay C, Faivre JC, Cambier S, Salleron J, Lambert A
AI-Driven Patient Screening for Clinical Trials in Pancreatic Cancer: The PANCR-AI Pilot Retrospective Comparative Study
JMIR Cancer 2026;12:e80268
URL: https://cancer.jmir.org/2026/1/e80268
doi:10.2196/80268

© Arthur Claessens, Alizée Simon, Agathe Manchart, Laura Desmonts, Cassandre Leguay, Jean-Christophe Faivre, Sébastien Cambier, Julia Salleron, Aurélien Lambert. Originally published in JMIR Cancer (https://cancer.jmir.org), 23.Feb.2026. This is

XSL•FO
RenderX

an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Navigating the Complexity of Lung Cancer Surveillance Practices: Qualitative Pilot Study on Provider Perspectives

Jenny M Woo[1,2*], BA; Sydney Conover[3*], BS; Caroline Gray[3], PhD; Shipra Arya[1,4,5], MD, SM; Julie T Wu[2,3], MD, PhD

[1]Division of Vascular Surgery, Stanford School of Medicine, 780 Welch Rd CJ350, Palo Alto, CA, United States

[2]Department of Oncology, Stanford School of Medicine, Palo Alto, CA, United States

[3]Center for Innovation to Implementation, VA Palo Alto Health Care System, Menlo Park, CA, United States

[4]Surgical Services, VA Palo Alto Health Care System, Palo Alto, CA, United States

[5]Stanford- Surgery Policy Improvement Research and Education (S-SPIRE) Center, Palo Alto, CA, United States

[*]these authors contributed equally

Corresponding Author:
Jenny M Woo, BA
Division of Vascular Surgery, Stanford School of Medicine, 780 Welch Rd CJ350, Palo Alto, CA, United States

Abstract

Background: Surveillance is noted to be an important part of survivorship to detect recurrence and/or second primary lung cancer (SPLC) at a curable stage. However, current surveillance guidelines remain controversial, and the factors providers consider in clinical decision-making are neither well-defined nor consistently applied.

Objective: In order to inform the qualitative protocol for a larger national study, this pilot study aimed to understand the factors that influence lung cancer surveillance and how providers view risk stratification as a potential tool to inform surveillance practices.

Methods: Semistructured interviews were conducted between October 2023 and July 2024 with purposively sampled providers involved in treating and surveilling patients with lung cancer from the US-based Palo Alto Veterans Affairs Medical Center and Stanford Medicine and its affiliate clinics. Providers were recruited through both email outreach and in-person invitations. Interviews were transcribed by an external transcription service and analyzed through a qualitative inductive content analysis approach to identify themes.

Results: In total, 11 physicians and 2 advanced practice providers (N=13) participated in interviews. The majority were from medical specialties (n=8, 61.5%), and the average number of years of practice as a provider was 9 years. A total of 3 themes were identified that describe the clinicians' sentiments about current surveillance practices and how a risk stratification tool could be used in screening for recurrence and/or SPLC. Clinicians consider a variety of clinical and nonclinical factors (category 1: factors that influence clinical decision making) and highlighted limits of a risk stratification tool, including concerns about generalizability, accuracy, and validity (category 2: sentiments toward a hypothetical risk stratification tool). Finally, concerns were raised about how delivering risk stratification data might impact patient anxiety, misinterpretation, and adherence to surveillance plans (category 3: delivery of risk stratification data to patients).

Conclusions: This qualitative analysis highlights the complexity of lung cancer surveillance decision-making and provider concerns about tool accuracy and delivery. While risk stratification tools may support surveillance decisions, their further development must address data quality, accuracy across diverse clinical and nonclinical risk factors, and effective patient-level data delivery. Doing so will facilitate the practical implementation of risk stratification tools to improve surveillance of SPLC and recurrence.

(JMIR Cancer 2026;12:e80659) doi:10.2196/80659

KEYWORDS

surveillance; lung cancer; risk stratification; second primary lung cancer; patient-provider relationship; clinical decision-making

Introduction

Surveillance is widely accepted to be an important part of lung cancer survivorship to help early detection of local recurrence and second primary lung cancer (SPLC) at a more curable stage. Guidelines by the American Society of Clinical Oncology, the National Comprehensive Cancer Network, and the European Society of Medical Oncology recommend surveillance imaging every 6 months for the first 2 years and then annually after those 2 years [1-3]. Despite these universal guidelines around

XSL•FO
RenderX

surveillance, not all survivors receive surveillance imaging. One study found that two-thirds of 1288 patients did not receive recommended surveillance imaging [4], while another study reported that less than 1 in 5 patients out of 1537 at Veterans Health Administration received recommended surveillance imaging [5].

Current variability in imaging receipt may be due to a complex mix of factors, including uncertainty about the optimal frequency of surveillance, duration of surveillance, which findings warrant continued surveillance, or who would be a good candidate for surveillance, considering other comorbidities. Although surveillance imaging is a widely accepted and guideline-recommended practice, there is limited clinical trial evidence that increased surveillance improved survival, especially for patients with advanced non-small cell lung cancer (NSCLC) [6]. Moreover, patients' comorbidities and overall life expectancy compel providers to balance the potential benefits of surveillance against its harms, including complications from biopsies of false positives and reduced quality of life from anxiety around scans [7,8]. How clinicians navigate these competing considerations remains unclear and may vary inconsistently across patients.

Tailored approaches to surveillance may benefit low-risk patients by reducing the frequency of surveillance, while focusing surveillance efforts on high-risk patients [6]. Risk calculators, developed using factors considered in retrospective evidence, can guide clinical judgment when current guidelines lack clear recommendations. For example, risk stratification based on malignancy risk is already used in primary lung cancer screening to identify high-risk patients and improve adherence to screening [9]. Applying a similar strategy to surveillance decision-making in survivors of lung cancer could help reduce variability and promote more consistent clinical practices.

Our study objective was to explore clinicians' clinical decision-making practices around surveillance and to assess the acceptability of risk stratification tools to guide surveillance.

Methods

Study Design

This pilot qualitative interview study was designed to gather insights and perspectives from providers who treat lung cancer. The aim of this study was to identify the array of factors that providers consider for risk of recurrence and/or SPLC and to explore provider perceptions of a hypothetical risk stratification tool that would evaluate the risk of recurrence and/or SPLC.

To contextualize clinician feedback, the hypothetical risk stratification tool referenced in interviews was modeled after the validated Second Primary Lung Cancer–Risk Assessment Tool (SPLC-RAT) developed by Stanford Medicine [10]. SPLC-RAT generates a 5-year SPLC risk estimate using regression-based modeling of clinical and smoking history inputs (eg, age at diagnosis, cancer stage, histology, surgical resection, smoking status, pack-years, and time since cessation). In practice, such a tool could integrate with the electronic health record to automatically pull structured data (eg, age, gender, and race) and combine it with manually entered variables, similar to the integration of analogous tools for primary lung cancer screening [11]. While modeled after SPLC-RAT, the hypothetical risk stratification tool was unnamed to study participants to collect their broader perceptions of risk stratification tools.

Recruitment

Providers were purposively identified and recruited from the Palo Alto Veterans Affairs (VA) medical center and Stanford Medicine and its affiliate clinics, all located within the United States. Eligibility criteria included licensed providers who are involved in treating and surveilling patients with lung cancer at these 2 health care systems. Eligible clinicians were invited to participate in the study through both email and in-person outreach. Initially, 1 provider from a thoracic oncology department was contacted, who then introduced the researcher (JMW) to other oncology providers in that department who were invited to be interviewed in person. The remaining providers were introduced by the study principal investigator (PI) (a thoracic oncologist; JTW) and were invited to participate via email.

Interview Process

The interview guide was modeled on a previous study on machine learning prognostic algorithms and adapted to fit the research questions for this study [12]. The interview guide contained 9 questions surrounding provider surveillance practices in evaluating risk of recurrence and perspectives on a hypothetical risk stratification tool (Textbox 1). Mock interviews were conducted among the research team (JMW, JTW, and SC), and the interview guide was subsequently adjusted for clarity. Research questions asked to providers aimed to explore their clinical decision-making practices around lung cancer surveillance and to assess their perspectives on a hypothetical risk stratification tool to guide surveillance.

XSL•FO
RenderX

Textbox 1. Interview questions.

1. When considering stopping surveillance for a lung cancer survivor, some clinicians stop once a certain timeframe has passed (ie, five y), while others may prefer to continue surveillance indefinitely. When do you think clinicians should consider stopping surveillance indefinitely? When do you think clinicians should consider stopping surveillance imaging for a survivor?

a. If they do not mention, ask: What are the reasons behind your opinion?

b. Follow up: What events or circumstances prompt you to order a surveillance scan?

c. If prompted for clarification or if the subject only discusses the time of diagnosis, mention that some clinicians are prompted by certain events or circumstances, such as multiple negative surveillance scans or declining functional status due to non-cancer-related causes.

2. How do you define risk of recurrence? Numerical risk? Or context of relative risk?

a. Do you think about risk in concrete terms, as in quantitative numbers or relative terms (ie, this pt is higher risk than someone else)?

3. What factors do you use in assessing the risk of recurrence and/or second malignancy in the context of the patient's overall health?

a. I understand guidelines say to screen regularly, but in practice, there are other co-morbidities that can change your decision. Does this come to mind apart from the guidelines regarding quantitative risk assessment?

    i. Do you use clinical intuition to modify guidelines?

4. What information about risk of recurrence and/or second malignancy do you usually share with a patient? [Do you have a discussion with a lung cancer survivor about whether or not to continue surveillance imaging?]

a. If they do not describe this, ask: Some clinicians frame surveillance in terms of general risk, such as the high risk of recurrence for lung cancer. Others specifically tailor it to patient-specific factors, such as ongoing smoking and advanced stage at diagnosis. Where along that spectrum do you fall?

5. We are developing a computer program that can use a patient's medical record data to estimate the patient's risk of developing a second primary lung cancer. If you had access to a score that accurately estimated a 5-year risk of each of your patients -- for example, a 0.1% risk of cancer or a 6% risk of cancer -- how would you use this information in the care of the patient, if at all?

a. If prompted for clarification of specific scenarios, specify: For example, a 60y patient diagnosed with early stage resected lung adenocarcinoma with a history of prior cancer before lung cancer and a 20 pk-yr smoking history has ~6% 5-year risk of SPLC. If that same patient did not receive resection and does not have a prior history of cancer, the risk reduces to 1.3%.

b. If they do not mention, ask: Do you think it would prompt a discussion about surveillance practices?

6. Who on the care team should be the primary recipient of this information and how should it be incorporated into the clinical workflow? [Are you comfortable with the primary recipient of this information being a designated screening coordinator?]

a. If prompted for clarification, mention that lung cancer screening has dedicated coordinators who triage patients based on risk assessment of lung nodule findings and that surveillance follow-up could follow a similar model

7. No prediction model will ever be perfect. A model could have many false positives, which in this case would be patients who are erroneously flagged as having a high risk of cancer. Alternatively, a model could have many false negatives, which in this case would be patients who are high risk of cancer but not flagged by the model. With that in mind, would you prefer to have a model that has more false negatives or more false positives? Why?

a. If prompted to define "high risk of cancer," it is whatever the clinician would consider high risk.

8. What concerns, if any, do you have about using computer-generated cancer risk predictions in the clinical care of your patients?

9. Do you have any current or prior experience using computer-generated models or cancer risk prediction tools to inform clinical decision-making?

a. If the person answers "Yes," ask: How does your experience impact the way you view the utility of computer-generated cancer risk predictions in clinical care?

After the main interview questions, demographic information including their years practicing in their specialty, specialty, age, gender, and race were collected.

Semistructured interviews with providers were conducted on Microsoft Teams by 2 trained female researchers (JMW and SC) and the female study PI (JTW). At each interview, verbal consent was obtained before the audio recording. Some study participants knew the PI before interviews, given her role as a thoracic oncology clinician at the health care systems participants were based. Some participants also had previous exposure to a previously developed SPLC risk stratification tool (SPLC-RAT) from Stanford Medicine, including as coauthors on the model development publication, which may have informed their responses to the hypothetical tool described in the interview guide. Interviews were conducted between October 2023 and July 2024, and audio files were transcribed by an external transcription service. The interviewers reviewed all

XSL•FO
RenderX

transcriptions to verify accuracy. Participants were not contacted following the initial interview.

Our research team consists of oncology and surgical clinicians, along with health services researchers, including experts in qualitative methods and oncology. We value a comprehensive understanding of current health system landscapes and the importance of engaging with providers directly involved in patient care before developing future interventions and tools to enhance the quality of health services.

Data Collection and Analysis

The research team aimed to interview at least 12 providers to support preliminary pilot findings; this has similarly been found sufficient for reaching a general level of saturation, or repetition of basic themes or findings [13]. When creating our participant sample, we also considered the concept of information power, which is defined as the capacity of a study's sample size to effectively capture the unique experiences and insights of participants, particularly when the sample is specifically aligned with the research objectives [14]. Since the eligibility criteria created a smaller, well-defined sample of eligible participants, a smaller sample size was sufficient to provide exploratory pilot data that addressed our research questions. Moreover, since interview questions were highly specific, with multiple follow-up probes, and there was depth to the answers, a wide array of insights was captured despite a smaller sample.

Interview transcripts were uploaded to Microsoft Excel. Data analysis was completed through an inductive content analysis approach. First, familiarization with the data was conducted by 2 researchers (JMW and JTW), after which a list of initial codes was identified. During this process, the researchers met weekly to discuss and determine consensus, creating the initial codebook, which contained preliminary codes and definitions, and to adapt the codebook as needed. The finalized codebook was further reviewed externally by the qualitative methods methodologist (CG), providing expert validation and addressing potential biases that may have arisen from the internal coding team. Once the codebook was agreed upon, the rest of the interviews were coded (JMW and SC). Finally, codes were analyzed and organized into groups to identify overarching themes. Themes were defined, and supporting data that were already analyzed were summarized into these larger themes.

Ethical Considerations

The study was approved by the Stanford University Institutional Review Board (IRB #67922) as an expedited review with a waiver of documentation of informed consent. VA Local Union Notification was obtained for VA employees. Oral information about the study was given before verbal informed consent. This included details about the study's purpose, risk, and benefits of participation, confidentiality protections, and participants' rights to withdraw at any time. All participants provided verbal consent for participation before audio recording. All data were deidentified before transcription using a coding system. Only the interviewers (JMW, SC, and JTW) had access to the code key. Transcriptions and audio recordings were stored in a secure and access-restricted folder on a server behind the VA firewall, accessible only to the research team. Participants did not receive any compensation for their participation.

Results

Participants

Of the 29 eligible providers recruited to participate, 11 physicians and 2 advanced practice providers (N=13) who were specifically involved in treating and surveilling patients with lung cancer were interviewed. The interviews ranged from 14 to 30 minutes. The average number of years the participants have been practitioners was 9 (SD 5.52, range 3 - 20) years. The majority of participants were medical specialists (n=8, 61.5%), while the remaining 38.4% (n=5) were interventional specialists. Additional demographic information is provided in Table 1.

Table . Demographic characteristics of providers (N=13).

Characteristics	Value
Age (y), mean (SD; range)	42.5 (5.75; 34-50)
Sex, n (%)	
Female	9 (69.2)
Male	4 (30.8)
Race, n (%)	
White	6 (46.1)
African American	1 (7.8)
Asian	6 (46.1)
Ethnicity, n (%)	
Hispanic, Latino/a/x, or Spanish Origin	0 (0)
Specialty, n (%)	
Interventional	5 (38.4)
Medical	8 (61.5)
Average time as a practitioner (y), mean (SD; range)	9 (5.52; 3-20)

XSL•FO
RenderX

Themes

Findings clustered around three main themes: (1) factors that influence clinical decision making, (2) sentiments toward a hypothetical risk stratification tool, and (3) delivery of risk stratification data to patients. Themes and illustrative quotes are displayed in Table 2.

XSL•FO
RenderX

Table . Themes and representative quotes.

Overarching theme and subcategory	Definition	Representative quotes
Factors that influence clinical decision-making (for surveillance)		
Clinical risk factors	Clinical risk factors providers consider for surveillance, including cancer-specific and non–cancer-specific factors.	• "I would probably consider smoking history in addition to the patient's medical history. Maybe other medical illnesses we need to consider, different types of cancers and pathologies" [Interventional Specialty, practicing for 10 y] • "What the patient's overall life expectancy is in respective to the diagnosis" [Medical Specialty, practicing for 1.5 y] • "With older people over the age of 75, they have competing risks. Not just their lung cancer recurrence but also dying from other causes so we certainly take that in account and a patient's ability to continue coming in. It's really a shared decision making with the patient" [Medical Specialty, practicing for 12 y]
Nonclinical risk factors	Nonclinical risk factors that providers consider such as patients' social circumstances and sentiments toward surveillance. Some social circumstances include differing emotional capacities, financial situations, and insurance coverage.	• "[There's] still a lot of anxiety result[ing] from living as a cancer survivor" [Medical Specialty, practicing for 21 y]
Sentiments toward a hypothetical risk stratification tool		
Positive sentiments	Providers' positive sentiments toward a hypothetical risk stratification tool	• "I think these risk calculators could be very helpful because it gives us concrete idea of the risk level and if it tells us of a very low risk level, it can give us some confidence in doing less surveillance imaging" [Medical Specialty, practicing for 8 y]
Provider concerns	Concerns providers have toward a hypothetical risk stratification tool. This includes concerns about the tool's generalizability, accuracy, and validity.	• "I'd have to trust the model and have to trust the data that is going into it to make sure I would be convinced that the information would be helpful to me" [Medical Specialty, practicing for 15 y] • "I think a computer-generated model will certainly have more pitfalls in poor and less fortified areas. You're going to have a lot of data for the rich folks and not so much data on the poor folks. This might be a serious social justice element to that – who gets screened for cancer, people who can afford a house in Palo Alto? If you live in Nevada, maybe not so." [Interventional Specialty, practicing for 10 y]
Views on false positives versus false negatives	Provider views on which error (type I or type II) they would choose between in a model setting.	• "The repercussions of missed opportunity to diagnose recurrence earlier is worse than the risk of overtreating" [Interventional Specialty, practicing for 14 y] • "Could do some Markov modeling to see ideal threshold. There could be downsides to over biopsying benign things. I would err on the side of flagging people" [Medical Specialty, practicing for 4 y]

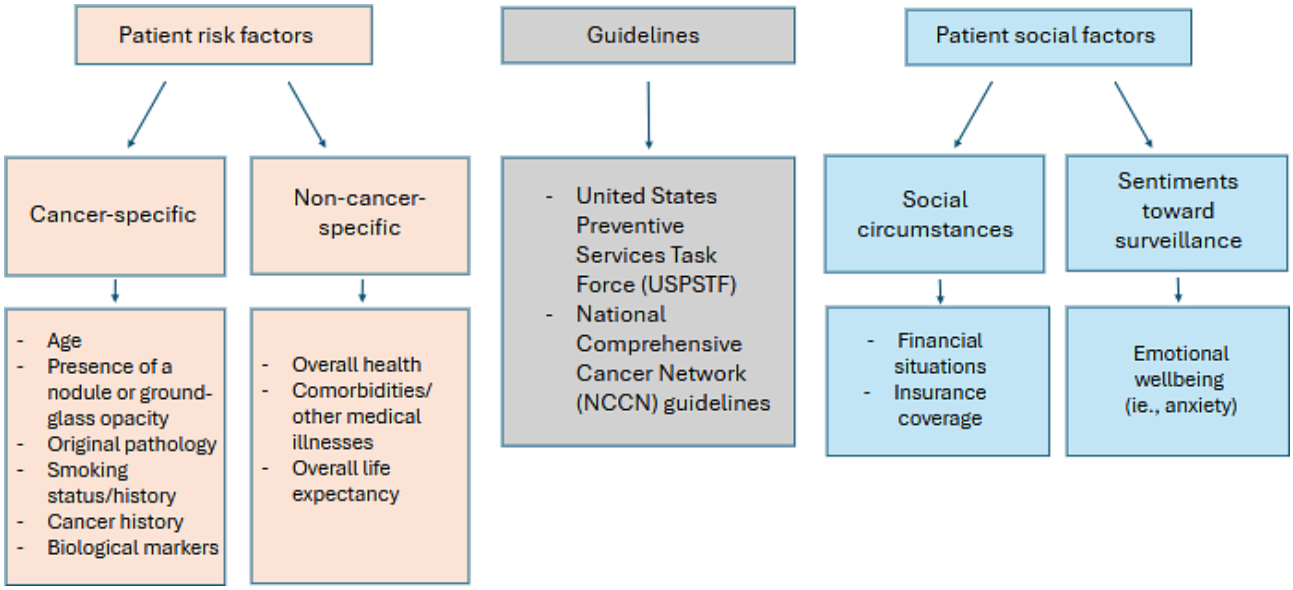
XSL•FO
RenderX

Overarching theme and subcategory	Definition	Representative quotes
Provider concerns about delivery of risk stratification data to patients	Providers' concerns about how patients would receive the risk stratification data.	• "Making sure that they understand the importance of coming back for their scans regularly" [Interventional Specialty, practicing for 5 y] • "I think it would be valuable to speak to that patient about the risk of radiation, risk of recurrence of cancer, and really put the numbers in front of them" [Interventional Specialty, practicing for 10 y]

Factors That Influence Clinical Decision Making

There are a multitude of factors that providers consider that can influence their clinical decision making, falling into 2 categories, clinical factors (patient risk factors and guidelines) and nonclinical factors (patients' social factors; Figure 1).

Figure 1. Factors that influence provider decisions on surveillance. This figure describes a summary of the factors mentioned that influence provider decisions on lung cancer surveillance. The first row of headings is different overarching groups of factors. Following each heading ("patient risk factors," "guidelines," and "patient social factors") are arrows pointing to further specific categories under the respective headings. For the "patient risk factors" and "patient social factors" sections, there is a second set of arrows that are pointing to specific examples of factors that were noted by the participants.



Clinical Factors (Patient Risk Factors and Guidelines)

When providers were asked about what factors they consider for surveillance, they first mentioned cancer-specific factors (patient's age, presence of a nodule or ground-glass opacity, original pathology, smoking status and history, and medical history). They also explained that they consider cancer history (stage of cancer, history of lung cancer, family history of cancer, and time from treatment) as well as biological markers (presence of mutations, molecular markers, and lymph node status) when considering surveillance for a specific patient.

Besides these factors, they also described how they consider guidelines and published data. Many providers mention a 5-year cutoff for surveillance unless there is a presence of a nodule, while others stated they screen every year indefinitely. The United States Preventive Services Task Force and National Comprehensive Cancer Network guidelines were mentioned as resources they use as well. When asked if they currently use a risk stratification tool for risk of recurrence and/or SPLC, the providers denied using one.

Beyond the guidelines and published data, providers listed other non–cancer-specific clinical factors, such as the patient's overall health, comorbidities, other medical illnesses, and overall life expectancy. One provider expressed how:

> *...with older people over the age of 75, they have competing risks. Not just their lung cancer recurrence but also dying from other causes so we certainly take that in account and a patient's ability to continue coming in. It's really a shared decision making with the patient.* [P04]

One provider mentioned how they look at life expectancy tables to help decide if focusing on lung cancer surveillance is in the patient's best interest.

Nonclinical Factors (Socioeconomic and Patient Sentiments)

Factors such as patient social circumstances and patient sentiments toward surveillance were highlighted. One provider stated,

XSL•FO
RenderX

> *People have a certain bandwidth for their healthcare, and I think they should be able to put their energy on what specific condition they should focus on [ie, heart condition or diabetes].* [P01]

They explained that patients' factors, such as differing emotional capacities, financial situations, and insurance coverage, all play a role in their clinical judgment on whether surveillance is beneficial. For instance, it was mentioned,

> *I also tell them things like I don't know if their insurance will cover surveillance after five years and things like that and then generally if they want to, then I'll order the scan.* [P08]

The provider explains that financial capability is also a part of the shared decision-making process in considering surveillance. Another provider described how they consider "healthcare spending [and] patients' bank accounts" (P10) as well.

Emotional well-being, particularly patient anxiety, was stated a few times. One provider explained how there's "still a lot of anxiety result[ing] from living as a cancer survivor" (P07). They noted there are significant emotional challenges that patients face as survivors of lung cancer and expressed great caution with how they bring up the topic of risk of recurrence and/or SPLC with their patients. One provider shared how "...many of them have anxieties, of course, some of them won't want the scans sooner and then some people don't want to know" (P11).

Another provider brought up a similar sentiment, stating:

> *There is a lot of anxiety about just having nodules knowing that one of them turned into a lung cancer group [in the past].* [P08]

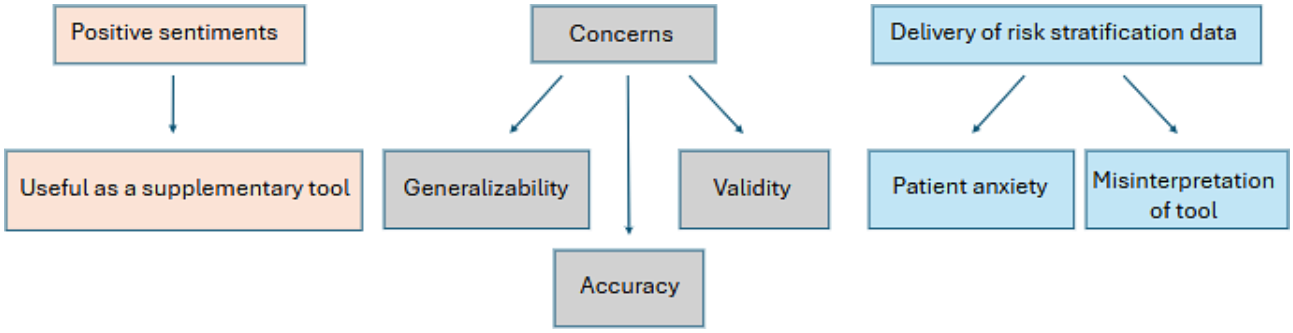
The providers explained how this is a shared decision-making process with their patients, acknowledging patients' socioeconomic and emotional bandwidths. One provider acknowledged that at times, they'll need to follow how the patient feels and stated:

> *If the patient decides that they don't want to do surveillance anymore then it is a discussion and then we will stop.* [P11]

Sentiments Toward a Hypothetical Risk Stratification Tool

When asked about a hypothetical risk stratification tool that can use a patient's medical record to estimate a 5-year risk of recurrence and/or SPLC, providers were generally positive, but did express concerns surrounding a risk stratification tool (Figure 2).

Figure 2. Provider sentiments on a risk stratification tool. This figure describes a summary of the provider sentiments on a hypothetical risk stratification tool. The first row of headings is different overarching groups of sentiments. Following each heading ("positive sentiments," "concerns," and "delivery of risk stratification data") are arrows pointing to further explanations and examples of sentiments that were noted by the participants. For the "delivery of risk stratification data," the arrows are pointing to outcomes providers expressed concerns of in relation to how the risk stratification data is delivered to patients.



Positive Sentiments

Many providers expressed interest in the hypothetical risk stratification tool, especially its role in surveillance practices past the 5-year mark, which they consider the cutoff for surveillance in survivors of lung cancer without a nodule as it would be "useful for determining whether [they] should keep surveying after 5 years" (P08). Having a risk stratification tool could be beneficial for providers,

> *...because it gives [a] concrete idea of the risk level and if it tells of a very low risk level, it can give some confidence in doing less surveillance imaging.* [P10]

While these providers expressed positive attitudes toward this hypothetical risk stratification tool, the majority stated that although it could be useful as a supplementary tool to clinical decision making, they would not rely solely on it due to various concerns outlined below.

Provider Concerns

Many providers expressed concerns about the generalizability, accuracy, and validity of the tool. They questioned the data the model would be using, with 1 provider suggesting:

> *...[they would] have to trust the data that is going into it to make sure [they] could be convinced that the information would be helpful.* [P13]

They expressed hesitancies if a tool could consider all the factors they consider and how representative of the patient population the data would be.

Regarding generalizability, 1 provider mentioned the potential shortcoming in the data surrounding socioeconomic status, stating:

> *...[the] model will certainly have more pitfalls in poor and less fortified areas. You're going to have a lot of data for the rich folks and not so much data for poor folks. That might be a serious social justice element*

XSL•FO
RenderX

> *to that – who gets screened for cancer, people who can afford a house in Palo Alto? If you live in Nevada, maybe not so.* [P02]

Not only that, underrepresentation of racial and gender disparities was also highlighted as a concern, with 1 provider expressing that,

> *...[they] worry a bit about racial or gender disparities in terms of information that models are given and whether that will actually apply to all patients.* [P08]

Due to the concerns that the data may not be representative and comprehensive enough, the providers suggested that the model may not be generalizable to their patients and thus may not be a tool they can fully rely on in their clinical decision-making.

Concerns with the accuracy and validity of the input data for the tool were also shared. A participant explained,

> *...as you do this [with] large amounts of data, things that might not be apparent otherwise can come forward. That's where we have the challenge of, is it really there or is it not really there, but you know you're not going to have most people though there will be some who ask about all those little details.* [P07]

Other concerns providers highlighted were the limitations of relying on data from electronic medical records. A provider stated that "there is lots of variability in the patient data like the dates of diagnoses" (P09) and how "the date of diagnosis is not discrete and consistent" (P09). One provider illustrated that,

> *...many patients have incomplete follow-up and survival data which could bias your observed outcomes and could result in an inaccurate model.* [P10]

Therefore, instead of being a primary tool to depend on, the providers emphasized that this tool should have a "disclaimer that of course [is] not predictive" (P12) of whether a patient has a recurrence and/or SPLC and should be used as a supplementary check.

Views on False Positives Versus False Negatives

Providers also discussed their perspectives on false positives and false negatives in a tool, with most providers preferring a model threshold that errs on the side of more false positives. One provider explained,

> *The repercussions of missed opportunity to diagnose recurrence earlier is worse than the risk of overtreating.* [P03]

Other providers, while still agreeing that false positives were better than false negatives, shared that there should be "some Markov modeling to see [an] ideal threshold" (P06) to help minimize the difference. Many providers illustrated that this approach of more false positives would be useful in flagging patients and then verifying if the finding is accurate, but false positives would falsely worry patients who did not have lung cancer, and that "there could be downsides to over biopsying benign things" (P06).

Delivery of Risk Stratification Data to Patients

When asked whether patients should directly receive the risk stratification tool data, most of the providers emphasized that patients should not directly receive it due to the potential for misinterpreting the numbers and developing anxiety about their health status. Many of the providers explained that they generally do not give quantitative values involving risk prediction as they want to avoid worrying their patients. They also worry that if they gave specific numbers, it would not truly resonate with the patient, as they would not understand the numbers in depth. One provider puts it as sharing enough in "making sure that they understand the importance of coming back for their scans regularly" (P08). There were a few providers who stated that they would share exact numbers to patients if the patients pushed for those numbers or used the numbers "...to speak to that patient about the risk of radiation, risk of recurrence of cancer, and really put the numbers in front of them" (P02).

Discussion

Principal Findings

This qualitative pilot study interviewed specialists from 2 medical systems about their current practices surrounding surveillance for risk of recurrence and/or SPLC, their attitudes toward a hypothetical risk stratification tool, and how they would deliver risk stratification data to their patients. Surveillance decisions were found to incorporate diverse clinical and nonclinical factors beyond current guidelines, which may explain variability in imaging receipt observed in population-based studies [15]. While participants showed interest in risk stratification tools to systematize decision-making, they raised concerns about the tool's generalizability, accuracy, validity, and content delivery. These concerns highlight key areas for future tool and intervention development. These pilot findings will also inform the development of a future national survey study of physician attitudes and perspectives of an adapted risk stratification tool for SPLC and lung cancer surveillance.

Providers emphasized how their surveillance plans go beyond following current guidelines, incorporating both clinical and nonclinical risk factors. They consider patients' overall well-being, comorbidities, and sentiments toward surveillance, tailoring efforts to reduce adverse outcomes, such as complications or reduced quality of life [7,8]. Most participants highlighted the importance of shared decision-making, which has been shown to improve knowledge surrounding lung cancer screening and be positively associated with intentions to go through screening again [16]. This personalized approach allows providers to create comprehensive patient-centered surveillance plans.

With innovations, such as computer models and artificial intelligence (AI) in cancer care, understanding provider acceptance and opinions toward a computer-generated risk stratification tool is important [17]. This feedback provides risk stratification tool developers with the necessary elements and concerns clinicians are likely to have. Given the complexity of factors in lung cancer recurrence and SPLC, such models could

XSL•FO
RenderX

guide and standardize surveillance by identifying high-risk patients while reducing unnecessary low-dose computed tomography scans for low-risk individuals.

Providers expressed interest in a hypothetical risk stratification tool for improving clinical care, but stated it was not robust enough to be used independently and emphasized they trust their clinical intuition above all. They viewed it as a supplemental tool that could provide additional validation to their clinical decision-making, citing limitations such as the data quality and model accuracy. Providers emphasized the importance of inclusive input data, particularly for underrepresented groups based on income, race, and gender. This has been shown to be a significant barrier to the widespread adoption of AI models in cancer care [18]. Accuracy concerns were also raised, such as overestimating risk in low-risk individuals or failing to identify high-risk patients. Given these limitations, most providers viewed such tools as secondary to clinical intuition. They stressed that the tool's output should not drive decisions independently, and that providers should understand the input factors and rely on their expertise to override the tool if inaccuracies arise.

Delivery of risk information was another key concern. Most providers indicated that they would avoid sharing exact risk stratification numbers with their patients, citing concerns about potential misinterpretation and anxiety. Studies have shown that patients with cancer already have scan-associated anxiety, and adding to this underlying anxiety could negatively influence their adherence to getting necessary surveillance scans [19,20]. Instead, providers have suggested that they could use the risk stratification tool to initiate conversations about risk without focusing on specific numbers. These sentiments from providers shed light on the importance of considering how patients will receive risk stratification tool information before sharing it, and in response, sharing enough information to ensure that the patient can continue to pursue the best care in a well-informed manner. Although providers viewed this tool as supplementary, they agreed that it could be a helpful double-check for their own clinical judgment and prompt alternative surveillance strategies, such as adjusting scan frequency. Ultimately, providers emphasized that with a tool or not, their goal for patients is to adhere to the most clinically accurate and effective plan to prevent recurrence and/or SPLC.

Limitations

First, the small sample size limits the transferability of findings to lung cancer survivorship specialists nationwide. However, as a pilot study, it gathered perspectives across various settings, including academia, the VA, and satellite clinics, which may have broadened the findings beyond a single clinic or hospital. Nonetheless, these results will inform our planned national study, which will produce more transferable findings. Second, participants were from an affluent, research-focused region, and providers may be more receptive to technology and innovation than those at other locations. This initial pilot study was also limited by recruitment reach. Future work would benefit from the inclusion of a more diverse group of lung cancer providers from various geographic locations, institutions, and hospital settings across the United States, particularly those in rural areas where access to new technology and innovation may be limited or less emphasized. Third, some participants knew the PI before interviews, given her role as a thoracic oncologist, and some participants had previous exposure to SPLC risk stratification tools. Due to narrow eligibility criteria and the interconnectedness of these 2 health care systems, this was inevitable and had the potential to impact study participants' responses. Future work will gather perspectives from a national sample of providers to broaden findings and make them more transferable. Fourth, providers' perceptions were reported based on a hypothetical tool, rather than one they have used or trialed, limiting the applicability of findings. While the hypothetical risk stratification tool concept was explained to providers and some had previous knowledge of risk stratification tools, participants were still inferring what the tool could or could not do. Future work should provide a detailed explanation of a developed risk stratification tool, outlining its exact functionality within surveillance.

Conclusions

This study illustrates the complexity of lung cancer surveillance decision-making, which may contribute to variability in surveillance imaging among survivors. While risk stratification offers a systematic approach to guide decisions, uptake is limited by concerns about data quality, tool accuracy, and delivery and applicability to individual patients. Addressing these issues can inform the development and implementation of more acceptable and effective risk stratification tools.

Acknowledgments

The authors would like to thank Alex Sox-Harris (MS, PhD) for his helpful feedback on this paper.

Funding

This work was supported by the Merit Review Award, 3I01HX003215-02S1, and VA Career Development Award, 1IK2HX003949-01A1, from the United States Department of Veterans Affairs, Health Services Research and Development Service (HSR&D).The contents do not represent the views of the US Department of Veterans Affairs or the United States Government.

XSL•FO
RenderX

Data Availability

The full dataset cannot be made publicly available due to privacy concerns and restrictions imposed by our local institutional review board. For investigators with the appropriate authorizations within the Department of Veterans Affairs, requests for data access can be made to the corresponding author.

Authors' Contributions

Conceptualization: JMW, JTW
Data curation: JMW, JTW, SC
Formal analysis: JMW, JTW, SC, CG
Funding acquisition: JTW, SA
Investigation: JMW, JTW, SC
Methodology: JMW, JTW
Project administration: JMW, JTW, SC
Resources: JTW, SA
Supervision: JTW, CG, SA
Visualization: JMW, SC, JTW
Writing – original draft: JMW, JTW, SC
Writing – review & editing: All authors.

Conflicts of Interest

SA is a consultant for Gore Medical and Novartis. All other authors have no potential conflicts of interest to disclose.

References

1. Schneider BJ, Ismaila N, Aerts J, et al. Lung cancer surveillance after definitive curative-intent therapy: ASCO guideline. J Clin Oncol 2020 Mar 1;38(7):753-766. [doi: 10.1200/JCO.19.02748] [Medline: 31829901]
2. Moran A, Daly ME. Surveillance imaging for non-small cell lung cancer: mounting evidence that less is more. Transl Lung Cancer Res 2019 Dec;8(Suppl 4):S343-S346. [doi: 10.21037/tlcr.2019.05.07] [Medline: 32038911]
3. Postmus PE, Kerr KM, Oudkerk M, et al. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2017 Jul 1;28(suppl_4):iv1-iv21. [doi: 10.1093/annonc/mdx222] [Medline: 28881918]
4. Bostock IC, Hofstetter W, Mehran R, et al. Barriers to surveillance imaging adherence in early-staged lung cancer. J Thorac Dis 2021 Dec;13(12):6848-6854. [doi: 10.21037/jtd-21-1254] [Medline: 35070369]
5. Randle R, Ajawara U, Adams S, et al. P1.28-01 adherence to post-treatment surveillance guidelines in non-small cell lung cancer - a veteran's health administration cohort. J Thorac Oncol 2023 Nov;18(11):S288. [doi: 10.1016/j.jtho.2023.09.494]
6. Jiang C, Zhang Y, Deng P, et al. The overlooked cornerstone in precise medicine: personalized postoperative surveillance plan for NSCLC. JTO Clin Res Rep 2024 Aug;5(8):100701. [doi: 10.1016/j.jtocrr.2024.100701] [Medline: 39188582]
7. Howard DH, Richards TB, Bach PB, Kegler MC, Berg CJ. Comorbidities, smoking status, and life expectancy among individuals eligible for lung cancer screening. Cancer 2015 Dec 15;121(24):4341-4347. [doi: 10.1002/cncr.29677] [Medline: 26372542]
8. Robinson EM, Liu BY, Sigel K, Yin C, Wisnivesky J, Kale MS. Impact of comorbidities on lung cancer screening evaluation. Clin Lung Cancer 2022 Jul;23(5):402-409. [doi: 10.1016/j.cllc.2022.03.012] [Medline: 35641376]
9. Wu JY, Wakelee HA, Han SS. Optimizing lung cancer screening with risk prediction: current challenges and the emerging role of biomarkers. JCO 2023 Sep 20;41(27):4341-4347. [doi: 10.1200/JCO.23.01060]
10. Choi E, Sanyal N, Ding VY, et al. Development and validation of a risk prediction model for second primary lung cancer. J Natl Cancer Inst 2022 Jan 11;114(1):87-96. [doi: 10.1093/jnci/djab138] [Medline: 34255071]
11. Caverly TJ, Wiener RS, Kumbier K, Lowery J, Fagerlin A. Prediction-augmented shared decision-making and lung cancer screening uptake. JAMA Netw Open 2024 Jul 1;7(7):e2419624. [doi: 10.1001/jamanetworkopen.2024.19624] [Medline: 38949809]
12. Parikh RB, Manz CR, Nelson MN, et al. Clinician perspectives on machine learning prognostic algorithms in the routine care of patients with cancer: a qualitative study. Support Care Cancer 2022 May;30(5):4363-4372. [doi: 10.1007/s00520-021-06774-w] [Medline: 35094138]
13. Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. Field methods 2006;18(1). [doi: 10.1177/1525822X05279903]
14. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
15. Thakur MK, Ruterbusch J, Schwartz AG, et al. Utility of surveillance imaging in stage III non-small cell lung cancer (NSCLC) patients: A SEER-Medicare analysis. JCO 2016 May 20;34(15_suppl):8531-8531. [doi: 10.1200/JCO.2016.34.15_suppl.8531]

XSL•FO
RenderX

16. Tan NQP, Nishi SPE, Lowenstein LM, et al. Impact of the shared decision-making process on lung cancer screening decisions. Cancer Med 2022 Feb;11(3):790-797. [doi: 10.1002/cam4.4445] [Medline: 34964284]
17. Jaber N. Can artificial intelligence help see cancer in new, and better, ways? National Cancer Insitute. URL: https://www.cancer.gov/news-events/cancer-currents-blog/2022/artificial-intelligence-cancer-imaging [accessed 2022-03-22]
18. Sebastian AM, Peter D. Artificial intelligence in cancer research: trends, challenges and future directions. Life (Basel) 2022 Nov 28;12(12):1991. [doi: 10.3390/life12121991] [Medline: 36556356]
19. Brandzel S, Rosenberg DE, Johnson D, et al. Women's experiences and preferences regarding breast imaging after completing breast cancer treatment. Patient Prefer Adherence 2017;11:199-204. [doi: 10.2147/PPA.S122244] [Medline: 28203064]
20. Bui KT, Liang R, Kiely BE, Brown C, Dhillon HM, Blinman P. Scanxiety: a scoping review about scan-associated anxiety. BMJ Open 2021 May 26;11(5):e043215. [doi: 10.1136/bmjopen-2020-043215] [Medline: 34039571]

Abbreviations:

AI: artificial intelligence
SPLC: second primary lung cancer
SPLC-RAT: Second Primary Lung Cancer–Risk Assessment Tool
VA: Veterans Affairs

Edited by A Schwartz, M Balcarras, N Cahill; submitted 14.Jul.2025; peer-reviewed by L Schuttner, L McVey; accepted 22.Dec.2025; published 05.Feb.2026.

Please cite as:
Woo JM, Conover S, Gray C, Arya S, Wu JT
Navigating the Complexity of Lung Cancer Surveillance Practices: Qualitative Pilot Study on Provider Perspectives
JMIR Cancer 2026;12:e80659
URL: https://cancer.jmir.org/2026/1/e80659
doi:10.2196/80659

© Jenny M Woo, Sydney Conover, Caroline Gray, Shipra Arya, Julie T Wu. Originally published in JMIR Cancer (https://cancer.jmir.org), 5.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Awareness of Lung Cancer Signs and Symptoms, Risk Factors, and Screening in Jordan: Cross-Sectional Study

Ahmad Ayman Yousef[1*], MD; Mohammad Yousef Saleh[2*], PhD; Hani Shihadeh[1], MD; Sanad Ahmad[1], MD; Mo'tasem Alaiwah[1], MD; Heba Shbailat[1], MD; Karen Zaher Alsyouf[1], MD; Malek Altaamreh[1], MD; Eman Alshawish[3], PhD; Rasmieh Al amer[4], PhD

[1]School of Medicine, University of Jordan, Amman, Jordan

[2]Clinical Nursing Department, School of Nursing, University of Jordan, Amman, Jordan

[3]Faculty of Nursing, Pediatric Nursing and Midwifery Division, An-Najah National University, Nablus, Occupied Palestinian Territory

[4]faculty of nursing, Yarmouk University, Irbid, Jordan

[*]these authors contributed equally

Corresponding Author:
Eman Alshawish, PhD
Faculty of Nursing
Pediatric Nursing and Midwifery Division
An-Najah National University
Nablus, Palestine
Nablus, 1159
Occupied Palestinian Territory
Phone: 970 0797027809
Email: alshawish@najah.edu

Abstract

Background: Lung cancer is one of the most commonly diagnosed cancers worldwide and the leading cause of cancer-related mortality. Global studies have highlighted the importance of awareness of lung cancer signs and symptoms, as it influences health care seeking, diagnosis timing, and treatment outcomes. Despite the high prevalence of both lung cancer and smoking in Jordan, no studies have assessed lung cancer awareness among the Jordanian population. This study is the first in Jordan to reveal the levels of awareness of lung cancer signs and symptoms, risk factors, and screening and to identify factors associated with the awareness level.

Objective: This study aimed to explore the awareness of lung cancer risk factors, signs and symptoms, and screening in the Jordanian population and the factors determining the levels of awareness.

Methods: A cross-sectional survey was conducted across all 12 governorates of Jordan, including 498 adult participants. Data were collected using a structured questionnaire covering sociodemographic characteristics and awareness about lung cancer signs, symptoms, risk factors, and screening. Items were derived from the Lung Cancer Awareness Measure Toolkit, version 2.1, 2011. Awareness scores for lung cancer risk factors, signs and symptoms, and overall lung cancer awareness were calculated. Multivariate linear regression was used to identify significant predictors of awareness scores. Multivariate logistic regression was used to determine significant predictors of lung cancer screening awareness.

Results: The mean lung cancer signs and symptoms awareness score was 5.4 (SD 4.9) out of 14, while the mean risk factor awareness score was 33.0 (SD 5.6) out of 45. The most recognized symptoms were coughing up blood (n=297, 59.6%) and pain upon breathing (n=237, 47.6%), while cigarette smoking (n=403, 80.9%) and exposure to air pollution (n=339, 68.1%) were the most recognized risk factors. Only 41.6% (n=207) of the participants had heard of lung cancer screening, while 79.3% (n=395) believed that screening improves survival. Being a smoker was predictive of lower awareness of lung cancer signs, symptoms, and risk factors, while a higher educational level and having a friend or relative previously diagnosed with lung cancer were associated with higher levels of awareness.

Conclusions: This study revealed a low level of awareness of lung cancer signs and symptoms, risk factors, and screening in the Jordanian population. Further studies are needed to identify effective strategies for improving awareness, especially among high-risk populations.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e86390) doi:10.2196/86390

KEYWORDS

lung cancer awareness; lung cancer risk factors; lung cancer signs and symptoms; lung cancer screening; Jordan

Introduction

According to recent statistics, lung cancer is one of the most commonly diagnosed cancers worldwide and the most common cause of cancer-related mortality [1]. Thus, it remains an extremely significant global health problem. In the Hashemite Kingdom of Jordan, lung cancer is also one of the most commonly diagnosed cancers and the most common cause of cancer-related mortality [1], which could be attributed to the very high tobacco smoking prevalence in the country [2]. Multiple studies have shown that low awareness of lung cancer signs and symptoms may be associated with longer delays in diagnosis [3-5], which in turn can lead to worse treatment outcomes and higher mortality rates [6]. In addition, research suggests that better awareness of lung cancer signs and symptoms and risk factors may lead to better attitudes toward smoking and alcohol habits [7]. This highlights the importance of promoting lung cancer awareness in different populations.

Several studies have evaluated awareness of lung cancer signs and symptoms and risk factors across various populations. In high-income nations, studies conducted in the United Kingdom, Australia, and Estonia revealed a low level of awareness. In all 3 studies, smoking was the most recognized risk factor, while shortness of breath and coughing up blood were the most recognized signs and symptoms. High socioeconomic status was predictive of higher awareness of signs, symptoms, and risk factors, while smokers had lower awareness levels than nonsmokers. Contradictory results were found regarding age and sex. In the United Kingdom and Australia, female participants had better awareness, while in Estonia, male participants demonstrated higher awareness. In the United Kingdom, older participants had better awareness, while in Australia and Estonia, younger people had better awareness [8-10]. Awareness of lung cancer signs, symptoms, and risk factors was also investigated in many low- and middle-income countries, including Nepal, Malaysia, India, and Nigeria. In concordance with results from high-income nations, tobacco smoking was the most commonly recognized risk factor, while other risk factors were poorly acknowledged, and coughing up blood was the most commonly recognized symptom [7,11-14].

Recently, research on awareness of lung cancer signs, symptoms, and risk factors has been conducted in several countries in the Middle East. In nations such as Syria, Saudi Arabia, and Palestine that are geographically and culturally similar to Jordan, results were also consistent with studies in other low- and middle-income nations. Overall levels of awareness were low, and higher levels of education and socioeconomic status were associated with higher levels of awareness [15-17]. The lack of such data in Jordan highlights the need for similar studies that explore the degree of awareness among the Jordanian population.

Considering the high prevalence and high mortality rate of lung cancer, along with the propensity for patients to seek medical attention at an advanced stage when treatment becomes more challenging [18], many resource-rich countries have implemented national screening programs for early detection of asymptomatic lung cancer using low-dose chest computed tomography scans. The survival benefit of lung cancer screening in populations at risk was confirmed by multiple randomized controlled trials, notably among individuals aged 50 to 80 years with a ≥20-pack-year smoking history [19-21], highlighting the importance of lung cancer screening as a public health measure.

Many research efforts were directed toward investigating lung cancer screening awareness among both smokers and nonsmokers. In low- and middle-income nations, research in China found that only approximately 25% of high-risk individuals participated in lung cancer screening. Additionally, awareness of lung cancer screening benefits was low, with 72.1% of participants perceiving it as unnecessary [22]. Notably, Lin et al [23] reported that a higher level of awareness of lung cancer signs and symptoms was predictive of more positive lung cancer screening behaviors. Studies in Malaysia and Saudi Arabia revealed high rates of willingness to undergo screening (78.8%-94.7%) among high-risk individuals [17,24]. Despite the high rate of willingness to undergo lung cancer screening, significant barriers, including a lack of medical insurance or awareness of lung cancer risk factors, might limit participation in screening programs. In a high-income nation such as the United States, studies have also reported low levels of lung cancer screening awareness [25,26]. Various strategies designed to enhance awareness of lung cancer screening and perhaps increase participation rates in lung cancer screening programs have been examined. Multiple studies found that various educational programs successfully raised awareness of lung cancer screening and improved participation rates for lung cancer screening [27,28].

As per our knowledge, studies exploring awareness of lung cancer signs, symptoms, risk factors, and screening have not been conducted in Jordan, despite the high incidence of this cancer and the very high smoking prevalence. In addition, very few studies have investigated the factors associated with lung cancer screening awareness. The main aims of this study are to explore the levels of awareness of lung cancer signs and symptoms, risk factors, and screening in the Jordanian population and to explore the factors that predict these levels of awareness.

Methods

Study Design

This cross-sectional study was conducted across all 12 governorates of the Hashemite Kingdom of Jordan, covering both rural and urban regions, from March 2024 to September 2025. In total, 498 adult Jordanian citizens participated in the

XSL•FO
RenderX

study. Due to limitations in time and resources, participants were recruited using convenience sampling. Data were collected through an online self-administered questionnaire distributed via Google Forms, which was shared on various social media platforms, including Facebook (Meta Platforms Inc) groups and WhatsApp (Meta Platforms Inc) channels. Only adult Jordanian citizens were eligible to complete the study. Individuals who were studying or working in health care–related fields were excluded from the study to reduce potential bias related to prior knowledge.

Ethical Considerations

Ethics approval for this study was obtained from the institutional review board ethics committee at the University of Jordan (JU-2024/32). All participants provided electronic informed consent prior to completing the questionnaire. They were informed that all collected data would be collected anonymously, kept confidential, and stored securely in a nonidentifiable format. They were also informed that participation was voluntary and that they could withdraw from the study at any time. This study was conducted in accordance with institutional and national ethical standards and the Declaration of Helsinki (1975, updated 2013). Participants did not receive any compensations or financial incentives.

Study Tool

Participants completed a questionnaire administered through Google Forms. The form consisted of 3 sections.

Section 1: Socioeconomic and Demographic Characteristics

This section gathered data on participants' age, sex, place of residence, level of education, marital status, job, and smoking status.

Section 2: Awareness of Lung Cancer Signs, Symptoms, and Risk Factors

Items were adapted from the Cancer Research UK Lung Cancer Awareness Measure Toolkit, version 2.1. 2011, which had shown good validity and reliability [10]. The Arabic translation of the questionnaire followed the Chapman and Carter translation procedure [29], which included expert forward translation, review by bilingual experts, and cultural adaptation to ensure clarity and appropriateness.

The signs and symptoms awareness subscale included 14 items, with responses scored as 1 for "yes" and 0 for "no or don't know." Those who answered "yes" were deemed to recognize the sign or symptom. Included signs and symptoms were unexplained weight loss, persistent chest infection (≥3 weeks), persistent cough (2 to 3 weeks or longer), persistent tiredness or lack of energy, persistent shortness of breath, persistent chest pain, persistent shoulder pain, coughing up blood, ache or pain when breathing, loss of appetite, painful cough, finger or nail changes, a high-pitched sound when breathing, and progression or new development in an existing cough. The signs and symptoms awareness score was calculated by summing the scores across all 14 items. Cronbach α for the lung cancer signs and symptoms awareness subscale in this study was 0.936.

The lung cancer risk factor awareness subscale consisted of 9 items scored on a 5-point Likert scale (1="strongly disagree," 2="disagree," 3="neutral or uncertain," 4="agree," and 5="strongly agree"). Participants who responded with "strongly agree" or "agree" were deemed to recognize the risk factor. Included risk factors were exposure to radon gas, exposure to passive smoking, previous cancer treatment, having a close relative with lung cancer, exposure to chemicals such as asbestos fibers, having a previous history of another cancer, such as head and neck cancer, exposure to air pollution, smoking, and having a history of chronic lung diseases such as chronic obstructive lung disease. A total risk factor awareness score was calculated for each participant by summing the scores across all 9 items, with higher scores indicating a higher level of awareness. Cronbach α for the lung cancer risk factors awareness subscale in this study was 0.901.

A total lung cancer awareness score was calculated by summing the total signs and symptoms awareness score and the total risk factors awareness score.

Section 3: Awareness of Lung Cancer Screening

Participants were asked whether they had ever heard about lung cancer screening and whether they believed that early detection of lung cancer improves cure and survival rates.

Data Analysis

Statistical analysis was performed using SPSS (version 27.0.1; IBM Corp). Descriptive statistics were used to summarize the characteristics of the study population. Inferential statistical tests, including 2-tailed independent samples *t* tests and 1-way ANOVA, were used to compare mean awareness scores across demographic subgroups. Multivariate linear regression was conducted to identify significant predictors of awareness scores related to lung cancer signs, symptoms, and risk factors, while multivariate logistic regression analysis was used to determine the predictors of awareness related to lung cancer screening. A *P* value of <.05 was considered statistically significant for all analyses.

Results

Descriptive Statistics

Table 1 summarizes the descriptive statistics of the study participants. The mean age of participants was 35.2 (SD 13.2) years. Most participants resided in an urban area, and most had completed either high school or a bachelor's degree or diploma. Overall, 39.6% (n=197) of participants were active smokers. The mean lung cancer risk factors awareness score was 33.0 (SD 5.6, range 9-45). The mean lung cancer signs and symptoms awareness score was 5.4 (SD 4.9, range 0-14). The mean total lung cancer awareness score was 38.4 (SD 8.4, range 9-59). In total, 41.6% (n=207) of the study participants had ever heard of lung cancer screening, while 79.3% (n=395) believed that lung cancer screening for early detection improves survival.

XSL•FO
RenderX

Table 1. Descriptive statistics of the study participants (N=498).

Variable	Values
Age (years), mean (SD)	35.2 (13.2)
Sex, n (%)	
Female	314 (63.1)
Male	184 (36.9)
Place of residence, n (%)	
Urban	449 (90.2)
Rural	49 (9.8)
Educational level, n (%)	
Elementary school	38 (7.6)
High school	74 (14.9)
Bachelor's degree or diploma	337 (67.7)
Postgraduate	49 (9.8)
Marital status, n (%)	
Single	200 (40.2)
Married	298 (59.8)
Smoking status, n (%)	
Nonsmoker	301 (60.4)
Smoker	197 (39.6)
Lung cancer risk factors awareness score[a]	
Mean (SD)	33.0 (5.6)
Median (IQR)	33.0 (29.0-36.0)
Lung cancer signs and symptoms awareness score[b]	
Mean (SD)	5.4 (4.9)
Median (IQR)	4.0 (0.0-10.0)
Total lung cancer awareness score[c]	
Mean (SD)	38.4 (8.4)
Median (IQR)	38.0 (32.0-44.0)
"Have you ever heard of lung cancer screening for early detection?" n (%)	
Yes	207 (41.6)
No	291 (58.4)
"Do you think that lung cancer screening for early detection can improve survival?" n (%)	
Yes	395 (79.3)
No or do not know	103 (20.7)

[a]Range for possible lung cancer risk factors awareness scores was 9-45.

[b]Range for possible lung cancer signs and symptoms awareness scores was 0-14.

[c]Range for possible total lung cancer awareness scores was 9-59.

Recognition of Lung Cancer Signs and Symptoms and Risk Factors

Table 2 summarizes the frequencies of recognition of the different lung cancer signs, symptoms, and risk factors. The most recognized signs and symptoms were coughing up blood (n=297, 59.6%) and pain when breathing (n=237, 47.6%), while the least recognized signs and symptoms were persistent shoulder pain (n=108, 21.7%) and changes in the shape of fingers or nails (n=122, 24.5%). The most recognized risk factors were being a smoker (n=403, 80.9%) and exposure to air pollution (n=339, 68.1%), while the least recognized risk factors were receiving treatment for a previous cancer (n=175, 35.1%) and having a past history of cancer (n=204, 41%).

XSL•FO
RenderX

Table 2. Recognition of signs, symptoms, and risk factors of lung cancer.

Variable	Values, n (%)
Symptoms	
Unexplained weight loss	152 (30.5)
Persistent chest infection	195 (39.2)
Cough that does not go away	168 (33.7)
Shortness of breath	224 (45)
Persistent tiredness	168 (33.7)
Persistent chest pain	216 (43.4)
Persistent shoulder pain	108 (21.7)
Coughing up blood	297 (59.6)
Ache or pain when breathing	237 (47.6)
Loss of appetite	176 (35.3)
Painful cough	228 (45.8)
Changes in the shape of the fingernails	122 (24.5)
High-pitched sound when breathing	205 (41.2)
Change in existing cough	203 (40.8)
Risk factors	
Exposure to radon gas	233 (46.8)
Passive cigarette smoke exposure	326 (65.5)
Past cancer treatment	175 (35.1)
Close relative with lung cancer	219 (44)
Exposure to chemicals (eg, asbestos)	249 (50)
Past history of cancer (eg, head and neck cancer)	204 (41)
Exposure to air pollution	339 (68.1)
Being a smoker	403 (80.9)
History of lung disease (eg, chronic obstructive pulmonary disease)	337 (67.7)

Association Among Risk Factors Awareness Score, Signs and Symptoms Awareness Score, Total Lung Cancer Awareness Score, and Demographic Characteristics

Table 3 summarizes lung cancer risk factors awareness score, signs and symptoms awareness score, and total lung cancer awareness score across the different demographic categories. Participants aged <35 years had a significantly higher level of awareness of lung cancer signs and symptoms (P=.04). Smokers had a significantly lower level of awareness of both lung cancer risk factors (P<.001) and signs and symptoms (P=.003). No significant differences in awareness of lung cancer risk factors and symptoms were observed across sex, place of residence, level of education, and marital status categories.

XSL•FO
RenderX

Table 3. Comparison of risk factors awareness scores, signs and symptoms awareness scores, and total lung cancer awareness scores across demographic categories.

Variable	Risk factors awareness score[a]		Signs and symptoms awareness score[b]		Total lung cancer awareness score[c]	
	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value
Age (years)		.34		.04		.06
≥35	32.7 (4.6)		5.0 (5.0)		37.7 (7.9)	
<35	33.2 (6.5)		5.9 (4.8)		39.1 (8.9)	
Gender		.18		.54		.19
Female	33.3 (4.7)		5.5 (4.7)		38.8 (7.7)	
Male	32.5 (6.9)		5.2 (5.4)		37.7 (9.4)	
Place of residence		.74		.15		.28
Rural	33.2 (6.5)		6.4 (5.2)		39.6 (9.3)	
Urban	32.9 (5.5)		5.3 (4.9)		38.3 (8.3)	
Level of education		.49		.17		.14
Elementary school	31.7 (6.4)		3.9 (5.2)		35.6 (9.6)	
High school	32.8 (4.9)		5.0 (5.1)		37.8 (8.2)	
Bachelor’s degree or diploma	33.1 (5.7)		5.7 (4.9)		38.8 (8.2)	
Postgraduate	33.3 (5.7)		5.6 (4.9)		38.9 (8.6)	
Marital status		.73		.33		.42
Single	33.1 (6.7)		5.7 (4.9)		38.8 (9.1)	
Married	32.9 (4.7)		5.2 (5.0)		38.1 (7.9)	
Smoking status		<.001		.003		<.001
Nonsmoker	33.9 (5.1)		6.0 (4.7)		39.9 (7.8)	
Smoker	31.6 (6.1)		4.6 (5.2)		36.2 (8.9)	

[a]Range for possible risk factors awareness scores was 9-45.

[b]Range for possible signs and symptoms awareness scores was 0-14.

[c]Range for possible total lung cancer awareness scores was 9-59.

Linear Regression Analysis

Table 4 summarizes the results of linear regression analysis used to predict risk factors awareness score, signs and symptoms awareness score, and total lung cancer awareness score. Having a bachelor’s degree or diploma level of education was significantly associated with a higher lung cancer signs and symptoms awareness score (β=1.89, 95% CI 0.23-3.55; *P*=.03). Being a smoker was significantly predictive of a lower total lung cancer awareness score (β=−3.74, 95% CI −5.29-−2.19; *P*<.001) as well as lower scores in both the risk factors (β=−2.23, 95% CI −3.29-−1.17; *P*<.001) and signs and symptoms (β=−1.51, 95% CI −2.42-−0.59; *P*<.001) domains. Having a friend or relative diagnosed with lung cancer was significantly predictive of higher symptom awareness scores (β=2.74, 95% CI 1.79-3.68; *P*<.001) and total lung cancer awareness scores (β=3.69, 95% CI 2.08-5.29; *P*<.001). There was no significant association between age, gender, place of residence, educational level, marital status, or having friends or relatives working in the health care sector, and total lung cancer awareness score.

XSL•FO
RenderX

Table 4. Predictors of lung cancer risk factors awareness score, signs and symptoms awareness score, and total lung cancer awareness score (multivariate linear regression analysis).

Predictor variable	Risk factors awareness score[a]		Signs and symptoms awareness score[b]		Total lung cancer awareness score[c]	
	β coefficient (95% CI)	*P* value	β coefficient (95% CI)	*P* value	β coefficient (95% CI)	*P* value
Age (years)	−0.04 (−0.10-0.01)	.13	−0.01 (−0.05-0.04)	.84	−0.05 (−0.13-0.03)	.25
Sex						
Female	Reference	Reference	Reference	Reference	Reference	Reference
Male	0.11 (−1.09-1.11)	.98	0.36 (−0.60-1.31)	.46	0.37 (−1.24-1.98)	.65
Place of residence						
Rural	Reference	Reference	Reference	Reference	Reference	Reference
Urban	−0.28 (−1.96-1.40)	.74	−1.28 (−2.74-0.18)	.08	−1.56 (−4.03-0.91)	.22
Educational level						
Elementary school	Reference	Reference	Reference	Reference	Reference	Reference
High school	0.64 (−1.54-2.83)	.56	0.93 (−0.96-2.82)	.33	1.58 (−1.63-4.78)	.33
BSc or diploma	0.71 (−1.21-2.63)	.47	1.89 (0.23-3.55)	.03	2.60 (−0.22-5.42)	.07
Postgraduate	0.47 (−1.94-2.88)	.70	1.47 (−0.61-3.56)	.17	1.94 (−1.59-5.48)	.28
Marital status						
Single	Reference	Reference	Reference	Reference	Reference	Reference
Married	0.66 (–0.80-2.12)	.37	−0.18 (–1.44-1.09)	.79	0.49 (–1.66-2.63)	.66
Smoking status						
Nonsmoker	Reference	Reference	Reference	Reference	Reference	Reference
Smoker	−2.23 (−3.29-−1.17)	<.001	−1.51 (−2.42-−0.59)	<.001	−3.74 (−5.29-−2.19)	<.001
"Do you have any friends or relatives who work in the health care sector?"						
No	Reference	Reference	Reference	Reference	Reference	Reference
Yes	0.65 (–0.71-2.01)	.35	−0.35 (–1.53-0.83)	.56	0.31 (–1.69-2.30)	.76
"Have you ever had a friend or a relative who was diagnosed with lung cancer?"						
No	Reference	Reference	Reference	Reference	Reference	Reference
Yes	0.95 (–0.14-2.05)	.09	2.74 (1.79-3.68)	<.001	3.69 (2.08-5.29)	<.001

[a]Range for possible risk factors awareness scores was 9-45. R^2=0.054; adjusted R^2=0.034; $F_{10,485}$=2.75; *P*=.003. All variance inflation factor values were <5, indicating an acceptable degree of collinearity.

[b]Range for possible signs and symptoms awareness scores was 0-14. R^2=0.093; adjusted R^2=0.074; $F_{10,485}$=4.96; *P*<.001. All variance inflation factor values were <5, indicating an acceptable degree of collinearity.

[c]Range for possible total lung cancer awareness scores was 9-59. R^2=0.095; adjusted R^2=0.076; $F_{10,485}$=5.05; *P*<.001. All variance inflation factor values were <5, indicating an acceptable degree of collinearity.

Logistic Regression Analysis

Table 5 summarizes the results of logistic regression analysis used to determine the factors predictive of the odds of recognition of lung cancer screening and believing that screening for early detection improves survival. Having a friend or relative diagnosed with lung cancer (odds ratio 2.90, 95% CI 1.87- 4.48; *P*<.001) and a higher symptom awareness score (odds ratio 1.10, 95% CI 1.06-1.15; *P*<.001) were significantly associated with higher odds of recognition of lung cancer screening. Having a higher educational level (Bachelor's degree or diploma: odds ratio 2.73, 95% CI 1.23-6.07; *P*=.01 vs postgraduate odds ratio 3.96, 95% CI 1.28-12.25; *P*=.02), a higher symptom awareness score (odds ratio 1.14, 95% CI 1.07-1.21; *P*<.001), and a higher risk factor awareness score (odds ratio 1.08, 95% CI 1.03-1.13; *P*<.001) were significantly associated with higher odds of believing that lung cancer screening is effective for early diagnosis and improving survival.

XSL•FO
RenderX

Table 5. Factors associated with the odds of recognition of lung cancer screening and the belief that early detection improves survival in patients with lung cancer (logistic regression analysis).

Predictor variables	"Have you ever heard of lung cancer screening for early detection?"[a]		"Do you think that lung cancer screening for early diagnosis improves survival in patients with lung cancer?"[b]	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Age (years)	1.01 (0.99-1.03)	.27	1.00 (0.98-1.03)	.95
Gender				
Female	Reference	Reference	Reference	Reference
Male	1.38 (0.90-2.12)	.14	0.88 (0.52-1.48)	.62
Place of residence				
Rural	Reference	Reference	Reference	Reference
Urban	1.23 (0.63-2.41)	.55	0.58 (0.23-1.43)	.23
Educational level				
Elementary school	Reference	Reference	Reference	Reference
High school	0.73 (0.30-1.73)	.47	2.51 (1.00-6.32)	.05
Bachelor's degree or diploma	0.96 (0.45-2.06)	.91	2.73 (1.23-6.07)	.01
Postgraduate	1.23 (0.48-3.16)	.67	3.96 (1.28-12.25)	.02
Marital status				
Single	Reference	Reference	Reference	Reference
Married	0.65 (0.36-1.15)	.14	0.99 (0.50-1.99)	.99
Smoking status				
Nonsmoker	Reference	Reference	Reference	Reference
Smoker	1.02 (0.67-1.57)	.93	1.20 (0.71-2.01)	.50
"Have you ever had any friends or relatives who were diagnosed with lung cancer?"				
No	Reference	Reference	Reference	Reference
Yes	2.90 (1.87-4.48)	<.001	1.28 (0.72-2.28)	.40
"Do you have any friends or relatives who work in the health care sector?"				
No	Reference	Reference	Reference	Reference
Yes	1.25 (0.72-2.17)	.43	1.07 (0.57-2.00)	.84
Risk factors awareness score[c]	1.01 (0.97-1.04)	.74	1.08 (1.03-1.13)	<.001
Signs and symptoms awareness score[d]	1.10 (1.06-1.15)	<.001	1.14 (1.07-1.21)	<.001

[a]Yes=1 and no=0. Hosmer-Lemeshow test: *P*=.82.

[b]Yes=1 and no or do not know=0. Hosmer-Lemeshow test: *P*=.05.

[c]Range for possible risk factors awareness scores was 9-45.

[d]Range for possible signs and symptom awareness scores was 0-14.

No significant association was found between age, gender, residence, marital status, smoking status, or having friends or family in the health care profession and the recognition of lung cancer screening or believing that screening is effective for enhancing survival through early diagnosis.

Discussion

Principal Findings

The primary objective of this study was to assess population-level awareness of lung cancer signs, symptoms, risk factors, and screening in Jordan. Our study revealed a lung cancer signs and symptoms awareness score of 5.4, indicating that, on average, participants could recognize approximately 5 to 6 signs or symptoms of lung cancer, reflecting a low level of awareness despite the relatively high incidence of lung cancer in Jordan. This score is comparable to the score observed in Nigeria (5) [13] but is significantly lower than the scores reported in the United Kingdom (8.78) [10] and Syria (8.63) [15]. This could be attributed to the lack of national awareness initiatives focused on lung cancer in Jordan and the nonspecific nature of lung cancer signs and symptoms, which could be easily

XSL•FO
RenderX

confused with symptoms or signs of other lung diseases that are prevalent in the country (eg, chronic obstructive pulmonary disease). The mean risk factor awareness score was 33.0 (SD 5.6), which is comparable to that reported in the United Kingdom (35.6) [10].

Smokers had a significantly lower level of awareness of both lung cancer risk factors (P<.001) and signs and symptoms (P=.003), aligning with most existing literature [8-10,15]. This association might reflect multiple factors. Smokers might be more likely to avoid or disengage from health awareness–promoting programs due to perceived threat or denial. In addition, tobacco smoking may be more common in populations with lower socioeconomic status and educational attainment, which might also limit their awareness. This finding elucidates the importance of focusing awareness programs on smokers due to their increased susceptibility to lung cancer.

Younger participants had greater awareness of lung cancer signs and symptoms, consistent with studies from Australia [8], Estonia [9], and Saudi Arabia [17], but divergent from findings from the United Kingdom [10], where older individuals exhibited greater awareness. This could be attributed to younger populations having greater access to social media and online information platforms, which promote their health awareness. However, in other countries, cancer awareness programs might focus on older populations due to their increased susceptibility.

Having a friend or relative diagnosed with lung cancer was significantly associated with a higher lung cancer signs and symptoms awareness score (β=2.74; P<.001), highlighting the influence of personal connections on an individual's perception of their own signs and symptoms.

Consistent with studies from high-, low-, and middle-income countries, smoking was the most recognized risk factor for lung cancer in Jordan, reflecting the global success of antitobacco campaigns in raising awareness of smoking as the leading preventable cause of lung cancer. Other risk factors (eg, exposure to asbestos fibers, radon gas, and receiving treatment for a previous cancer) were much less recognizable, likely due to their relative rarity compared to tobacco smoking. This highlights the importance of focusing health awareness programs on the most overlooked risk factors, particularly in populations with distinct exposures (eg, individuals whose occupations involve asbestos exposure).

Signs and symptoms recognition was variable across different countries. In Jordan and the United Kingdom, the most recognized symptom was coughing up blood [10], while shortness of breath was the most recognized symptom in Syria [15], Saudi Arabia [17], and Australia [8]. Change in a pre-existing cough and prolonged cough were the most commonly recognized in Palestine and Estonia, respectively [9,16]. This variability could be attributed to the varying prevalence of other respiratory diseases across the globe. In populations with a high prevalence of tuberculosis, for example, coughing up blood may be less recognized as a symptom of lung cancer, while in populations with a high prevalence of chronic obstructive pulmonary disease, shortness of breath and a change in a pre-existing cough may be less perceived as serious.

The least recognized lung cancer signs and symptoms in our study were persistent shoulder pain and changes in the shape of the nails. Unlike the more prominent respiratory symptoms (eg, coughing up blood), these symptoms may be seen as unrelated to the lungs in common perception; rather, they might be linked to the musculoskeletal system.

A small proportion of participants (n=207, 41.6%) in our study recognized the concept of lung cancer screening, while the majority (n=395, 79.3%) believed that screening improves survival. These results are comparable to findings from the United States [25], where 44% of participants had heard of and were aware of lung cancer screening, and 93% believed that catching it early results in better outcomes. The high proportion of participants believing that screening for early detection improves survival despite the low level of lung cancer screening recognition might be due to the common misconception that early diagnosis improves outcomes in all cancers, while the low level of recognition of lung cancer screening might be due to its relative novelty compared to screening for other cancers (eg, breast cancer). Having a friend or relative diagnosed with lung cancer was associated with higher odds of recognition of lung cancer screening, again highlighting the importance of close people's experiences in perceptions of an individual's own health.

Higher lung cancer awareness scores were significantly associated with higher odds of recognizing lung cancer screening and believing that it improves survival, highlighting the potential role of improving lung cancer awareness in promoting positive attitudes toward lung cancer screening.

To the best of our knowledge, this is the first study to assess population awareness of lung cancer symptoms, risk factors, and screening in the Hashemite Kingdom of Jordan, and this is the first study to identify and quantify that having a relative or friend diagnosed with lung cancer increases awareness of lung cancer signs, symptoms, and screening. Further research should be done to identify effective interventions to improve knowledge and awareness of lung cancer, minimize diagnostic delay, and improve early detection and participation rates in lung cancer screening, especially among high-risk groups. In addition, further studies are warranted to explore characteristics of patients with lung cancer in Jordan, including involvement in screening programs, stage at diagnosis, and unique exposures.

Limitations

Several limitations should be acknowledged when considering the findings of our study. The use of an online form to collect data excludes individuals with no access to social media or the internet, which might induce selection bias, resulting in an overestimation of awareness levels. This approach may also have introduced bias toward younger participants, as they are more likely to use social media, thereby limiting the representation of older adults who are at higher risk of developing lung cancer. The small sample size limits the generalizability and the representativeness of our sample. Participants' governorates of residence were not recorded, limiting the ability to determine geographic representativeness and identify potential regional variation in lung cancer awareness across Jordan. The reliance on a self-administered

XSL•FO
RenderX

questionnaire and the lack of concise explanations limited deeper exploration of participants' understanding. Moreover, the inability to completely exclude health care professionals might bias the results toward higher awareness levels. Our study had a pure quantitative design and lacked a qualitative approach to evaluate lung cancer awareness, depended on recognition rather than recall, and used closed-ended rather than open-ended questions. Our study included Jordanian adults of all ages, with no focus on older people despite their higher susceptibility.

Conclusions

Our study revealed a low level of awareness of lung cancer signs and symptoms, risk factors, and screening. Higher awareness was associated with younger age, having a friend or relative diagnosed with lung cancer, while smoking was linked with a lower level of awareness. These findings underscore the importance of implementing targeted, age-inclusive public health strategies to improve lung cancer awareness and promote early detection, particularly among high-risk groups.

Acknowledgments

All authors disclose that no artificial intelligence tools were used in the ideation, writing, data analysis, or revision of this manuscript. All ideas and text were generated solely by human authors.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Funding

The authors declare that no financial support was received for this study.

Conflicts of Interest

None declared.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Alkouri O, Khader Y, Al-Bashaireh AM. Prevalence of cigarettes and waterpipe smoking among Jordanians, refugees, and migrants in Jordan and its associated factors: a secondary data analysis. Int J Environ Res Public Health 2022 Dec 21;20(1):82 [FREE Full text] [doi: 10.3390/ijerph20010082] [Medline: 36612400]
3. Corner J, Hopkinson J, Roffe L. Experience of health changes and reasons for delay in seeking care: a UK study of the months prior to the diagnosis of lung cancer. Soc Sci Med 2006 Mar;62(6):1381-1391. [doi: 10.1016/j.socscimed.2005.08.012] [Medline: 16168544]
4. Ellis PM, Vandermeer R. Delays in the diagnosis of lung cancer. J Thorac Dis 2011 Sep;3(3):183-188 [FREE Full text] [doi: 10.3978/j.issn.2072-1439.2011.01.01] [Medline: 22263086]
5. Smith SM, Campbell NC, MacLeod U, Lee AJ, Raja A, Wyke S, et al. Factors contributing to the time taken to consult with symptoms of lung cancer: a cross-sectional study. Thorax 2009 Jun;64(6):523-531. [doi: 10.1136/thx.2008.096560] [Medline: 19052045]
6. Romine PE, Sun Q, Fedorenko C, Li L, Tang M, Eaton KD, et al. Impact of diagnostic delays on lung cancer survival outcomes: a population study of the US SEER-Medicare database. JCO Oncol Pract 2022 Jun;18(6):e877-e885 [FREE Full text] [doi: 10.1200/OP.21.00485] [Medline: 35119911]
7. Shankar A, Roy S, Malik A, Rath GK, Julka PK, Kamal VK, et al. Level of awareness of various aspects of lung cancer among college teachers in India: impact of cancer awareness programmes in prevention and early detection. J Cancer Educ 2016 Dec;31(4):709-714. [doi: 10.1007/s13187-015-0960-7] [Medline: 26687206]
8. Crane M, Scott N, O'Hara BJ, Aranda S, Lafontaine M, Stacey I, et al. Knowledge of the signs and symptoms and risk factors of lung cancer in Australia: mixed methods study. BMC Public Health 2016 Jun 13;16:508 [FREE Full text] [doi: 10.1186/s12889-016-3051-8] [Medline: 27296668]
9. Hanson HM, Raag M, Adrat M, Laisaar T. Awareness of lung cancer symptoms and risk factors in general population. Open J Respir Dis 2017 Feb;07(01):1-11. [doi: 10.4236/ojrd.2017.71001]
10. Simon AE, Juszczyk D, Smyth N, Power E, Hiom S, Peake MD, et al. Knowledge of lung cancer symptoms and risk factors in the U.K.: development of a measure and results from a population-based survey. Thorax 2012 May;67(5):426-432. [doi: 10.1136/thoraxjnl-2011-200898] [Medline: 22426791]
11. Al-Naggar RA, Kadir SY. Lung cancer knowledge among secondary school male teachers in Kudat, Sabah, Malaysia. Asian Pac J Cancer Prev 2013;14(1):103-109 [FREE Full text] [doi: 10.7314/apjcp.2013.14.1.103] [Medline: 23534705]
12. Chawla R, Sathian B, Mehra A, Kiyawat V, Garg A, Sharma K. Awareness and assessment of risk factors for lung cancer in residents of Pokhara Valley, Nepal. Asian Pac J Cancer Prev 2010;11(6):1789-1793 [FREE Full text] [Medline: 21338234]

XSL•FO
RenderX

13. Desalu OO, Fawibe AE, Sanya EO, Ojuawo OB, Aladesanmi AO, Salami AK. Lung cancer awareness and anticipated delay before seeking medical help in the middle-belt population of Nigeria. Int J Tuberc Lung Dis 2016 Apr;20(4):560-566. [doi: 10.5588/ijtld.15.0497] [Medline: 26970168]
14. Loh JF, Tan SL. Lung cancer knowledge and screening in the context of the Malaysian population. J Pharm Pract Res 2018 Feb 15;48(1):56-64. [doi: 10.1002/jppr.1341]
15. Alibrahim H, Bohsas H, Swed S, Abouainain Y, Almoshantaf M, Nasif M, et al. Awareness of lung cancer risk factors and symptoms in Syria: an online cross-sectional study. F1000Research 2024;13:172 [FREE Full text] [doi: 10.12688/f1000research.144432.1]
16. Elshami M, Abukmail H, Aqel W, Alser M, Al-Slaibi I, Shurrab H, et al. Awareness of Palestinians about lung cancer symptoms: a national cross-sectional study. BMC Pulm Med 2022 Apr 08;22(1):135 [FREE Full text] [doi: 10.1186/s12890-022-01923-1] [Medline: 35395762]
17. Madkhali MA, Alhazmi E, Hakami F, Darraj H, Hamdi S, Hakami KM, et al. A cross-sectional study on the knowledge and awareness of lung cancer and screening in Jazan region, Saudi Arabia. J Multidiscip Healthc 2023 Dec 5;16:3857-3870 [FREE Full text] [doi: 10.2147/JMDH.S435129] [Medline: 38076592]
18. Prado MG, Kessler LG, Au MA, Burkhardt HA, Zigman Suchsland M, Kowalski L, et al. Symptoms and signs of lung cancer prior to diagnosis: case-control study using electronic health records from ambulatory care within a large US-based tertiary care centre. BMJ Open 2023 Apr 20;13(4):e068832 [FREE Full text] [doi: 10.1136/bmjopen-2022-068832] [Medline: 37080616]
19. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med 2020 Feb 06;382(6):503-513. [doi: 10.1056/NEJMoa1911793] [Medline: 31995683]
20. Henschke CI, Yip R, Shaham D, Markowitz S, Cervera Deval J, Zulueta JJ, et al. A 20-year follow-up of the International Early Lung Cancer Action Program (I-ELCAP). Radiology 2023 Nov;309(2):e231988. [doi: 10.1148/radiol.231988] [Medline: 37934099]
21. Ru Zhao Y, Xie X, de Koning HJ, Mali WP, Vliegenthart R, Oudkerk M. NELSON lung cancer screening study. Cancer Imaging 2011 Oct 03;11 Spec No A(1A):S79-S84 [FREE Full text] [doi: 10.1102/1470-7330.2011.9020] [Medline: 22185865]
22. Zhang T, Mao Z, Li W, Ma M, Li G, Qiao X, et al. Knowledge, attitude, and practice of lung cancer screening and associated factors among high-risk population in Lanzhou, China: a cross-sectional study. Medicine (Baltimore) 2024 Mar 01;103(9):e37431 [FREE Full text] [doi: 10.1097/MD.0000000000037431] [Medline: 38428855]
23. Lin YA, Lin X, Li Y, Wang F, Arbing R, Chen W, et al. Screening behaviors of high-risk individuals for lung cancer: a cross-sectional study. Asia Pac J Oncol Nurs 2024 Feb 10;11(4):100402 [FREE Full text] [doi: 10.1016/j.apjon.2024.100402] [Medline: 38495639]
24. Nyanti LE, Chua CZ, Loo HC, Khor CZ, Toh ES, Gill RS, et al. Determinants of willingness to undergo lung cancer screening among high-risk current and ex-smokers in Sabah, Malaysia: a cross-sectional pilot study. Tuberc Respir Dis (Seoul) 2023 Oct;86(4):284-293 [FREE Full text] [doi: 10.4046/trd.2023.0051] [Medline: 37643812]
25. Monu J, Triplette M, Wood DE, Wolff EM, Lavallee DC, Flum DR, et al. Evaluating knowledge, attitudes, and beliefs about lung cancer screening using crowdsourcing. Chest 2020 Jul;158(1):386-392 [FREE Full text] [doi: 10.1016/j.chest.2019.12.048] [Medline: 32035910]
26. Welch AC, London SM, Wilshire CL, Gilbert CR, Buchwald D, Ferguson G, et al. Access to lung cancer screening among American Indian and Alaska native adults: a qualitative study. Chest 2024 Mar;165(3):716-724. [doi: 10.1016/j.chest.2023.10.025] [Medline: 37898186]
27. Dignan M, Cina K, Sargent M, O'Connor M, Tobacco R, Burhansstipanov L, et al. Increasing lung cancer screening for high-risk smokers in a frontier population. J Cancer Educ 2024 Feb;39(1):27-32. [doi: 10.1007/s13187-023-02369-7] [Medline: 37688691]
28. Williams LB, Shelton BJ, Gomez ML, Al-Mrayat YD, Studts JL. Using implementation science to disseminate a lung cancer screening education intervention through community health workers. J Community Health 2021 Feb;46(1):165-173 [FREE Full text] [doi: 10.1007/s10900-020-00864-2] [Medline: 32594413]
29. Chapman DW, Carter JF. Translation procedures for the cross cultural use of measurement instruments. Educ Eval Policy Anal 1979 May;1(3):71-76. [doi: 10.3102/01623737001003071]

XSL•FO
RenderX

Edited by A Schwartz, M Balcarras; submitted 23.Oct.2025; peer-reviewed by A Sahajan, X Ding, O Ajayi, R Akinniranye, S Zaraa; comments to author 19.Jan.2026; revised version received 04.Feb.2026; accepted 05.Feb.2026; published 19.May.2026.

Please cite as:
Yousef AA, Saleh MY, Shihadeh H, Ahmad S, Alaiwah M, Shbailat H, Alsyouf KZ, Altaamreh M, Alshawish E, Al amer R
Awareness of Lung Cancer Signs and Symptoms, Risk Factors, and Screening in Jordan: Cross-Sectional Study
JMIR Cancer 2026;12:e86390
URL: https://cancer.jmir.org/2026/1/e86390
doi:10.2196/86390
PMID:

©Ahmad Ayman Yousef, Mohammad Yousef Saleh, Hani Shihadeh, Sanad Ahmad, Mo'tasem Alaiwah, Heba Shbailat, Karen Zaher Alsyouf, Malek Altaamreh, Eman Alshawish, Rasmieh Al amer. Originally published in JMIR Cancer (https://cancer.jmir.org), 19.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cancer Clinician Experiences With Telehealth Across Oncology Specialties: Cross-Sectional Survey Study

Benjamin Baker[1], MD; Sakshith Reddy Chintala[1], BS; Sarah Azari[1], MD; James D Murphy[2], MD; Hala Madanat[3], PhD; Maria Elena Martinez[4], PhD; Richard M Cripps[3], DPhil; Aditya Bagrodia[1], MD; Humberto Parada Jr[5], PhD; Juan Javier-DesLoges[1], MD

[1]Department of Urology, University of California, San Diego, 200 W. Arbor Drive #8897, San Diego, CA, United States

[2]Department of Radiation Medicine and Applied Sciences, Unviersity of California San Deigo, La Jolla, CA, United States

[3]Department of Biology, San Diego State University, San Diego, CA, United States

[4]Herbert Wertheim School of Public Health and Moores Cancer Center, University of California San Diego, La Jolla, CA, United States

[5]Department of Epidemiology and Biostatistics, School of Public Health, San Diego State University, San Diego, CA, United States

Corresponding Author:
Benjamin Baker, MD
Department of Urology, University of California, San Diego, 200 W. Arbor Drive #8897, San Diego, CA, United States

Abstract

Background: Telehealth use expanded rapidly during the COVID-19 pandemic and has since become integrated into routine oncology care. However, clinician experiences and confidence with telehealth may vary across oncology subspecialties, and these differences remain incompletely characterized.

Objective: This study aimed to evaluate telehealth perceptions, decision-making factors, and confidence across oncology clinicians from multiple specialties at a single academic cancer center.

Methods: We conducted a cross-sectional survey of oncology clinicians at the University of California San Diego Moores Cancer Center. A 27-item Qualtrics questionnaire was distributed by email to attending physicians, fellows, and advanced practice providers across surgical, medical, radiation, and palliative oncology. Survey domains included telehealth coordination, technical quality, communication effectiveness, confidence in remote physical examination, and factors influencing visit modality selection. Quantitative responses were summarized descriptively, and open-ended responses underwent thematic analysis.

Results: Among 114 invited clinicians, 57 (50%) complete responses were analyzed. Among 57 respondents, 86% (n=49) were physicians. Specialties included surgical oncology (n=21, 36.8%), medical oncology (n=18, 31.6%), radiation oncology (n=16, 28.1%), and palliative care (n=2, 3.5%). Telehealth coordination was rated easy or very easy by 78.9% (n=45) of clinicians. Audio and video quality were rated good or very good by 71.9% (n=41) and 57.9% (n=33), respectively. Technical issues prompted at least occasional conversion to audio-only visits for 86% (n=49) of respondents, and 17.5% (n=10) of clinicians reported frequent conversion. Communication satisfaction was high (n=48, 84.2%), yet only 50.9% (n=29) reported moderate or high confidence in performing a virtual physical examination. The need for physical examination (n=53, 93%) and patient travel distance (n=45, 78.9%) were the most influential factors in offering telehealth visits. Confidence varied by specialty and clinical phase: moderate or high confidence was reported by surgical oncology clinicians in 52.4% (11/21) of preoperative and 71.4% (15/21) of postoperative responses, by medical oncology clinicians in 77.8% (14/18) of prechemotherapy and 76.5% (13/17) of during chemotherapy responses, and by radiation oncology clinicians in 87.5% (14/16) of preradiation and 93.8% (15/16) of during radiation responses.

Conclusions: Oncology clinicians reported generally positive experiences with telehealth, although confidence and use patterns varied across specialties and phases of care. Telehealth is viewed as a complementary modality when physical examination is not essential. Improvements in technical reliability and interpreter integration may enhance equitable access and support continued telehealth use in oncology care.

(JMIR Cancer 2026;12:e94610) doi:10.2196/94610

KEYWORDS

telehealth; oncology; telemedicine; digital health; cancer care delivery; clinician experience

XSL•FO
RenderX

Introduction

Telehealth use expanded rapidly during the COVID-19 pandemic and has remained integrated into many areas of routine clinical care [1-3]. In oncology, telehealth has particular appeal because patients often require longitudinal care, frequent follow-up, multidisciplinary coordination, and review of laboratory, imaging, pathology, and treatment-related information. For selected visits, virtual care may reduce travel burden, decrease time away from work or caregiving responsibilities, and improve access for patients who live far from cancer centers [1,2].

Despite these advantages, telehealth in oncology is not uniformly applicable across all clinical settings. Cancer care often depends on physical examination, assessment of treatment toxicity, symptom evaluation, shared decision-making, and coordination across medical, surgical, radiation, and supportive care teams. The relative importance of these factors may differ substantially by oncology subspecialty and by phase of care. For example, surgical oncologists may have different levels of confidence in telehealth before versus after an operation, whereas medical and radiation oncologists may use telehealth differently during treatment planning, treatment monitoring, and survivorship care. Prior studies have described telehealth implementation in cancer centers, transition to telehealth in surgical practice, and clinician attitudes toward telehealth in multidisciplinary cancer care, but specialty-specific clinician confidence across different phases of oncology care remains incompletely characterized [4-6].

Understanding clinician experience is important because clinicians play a central role in determining whether telehealth is offered, whether it is considered clinically appropriate, and how technical or communication barriers are managed. Clinician confidence may be shaped by visit type, need for physical examination, institutional technology, interpreter availability, patient age, and other patient-specific factors. Broader telehealth use data also suggest that adoption and experience may be shaped by patient characteristics, clinical setting, and health system infrastructure [7]. These issues are especially relevant because telehealth may improve access for some patients while potentially widening disparities for others if language, technology, or workflow barriers are not addressed.

However, limited data describe how oncology clinicians across different specialties perceive telehealth, what factors influence their decision to offer virtual visits, and how confident they are using telehealth during different phases of cancer care. To address this gap, we conducted a cross-sectional survey of medical, surgical, radiation, and palliative oncology clinicians at a single academic cancer center. We evaluated clinician experiences with telehealth, factors influencing visit modality selection, perceived barriers and facilitators, and confidence in telehealth across specialty-specific clinical scenarios.

Methods

Study Design and Setting

We conducted a cross-sectional survey study of oncology clinicians at the University of California San Diego Moores Cancer Center, a National Cancer Institute–designated academic cancer center. The survey was designed to assess clinician experiences with telehealth across multiple oncology subspecialties, including perceived visit quality, technical challenges, communication, confidence in remote assessment, and factors influencing visit modality selection.

Participants and Recruitment

Medical, surgical, palliative, and radiation oncology physicians, fellows, and advanced practice clinicians were invited to participate by email. The survey was distributed using a combination of individual provider email lists and specialty or institutional email distribution lists. A total of 114 providers were invited, including 32 (28.1%) surgical oncology clinicians, 49 (43%) medical oncology clinicians, 26 (22.8%) radiation oncology clinicians, and 7 (6.1%) palliative care clinicians. The survey was open from October 1, 2024, through January 31, 2025. Participation was voluntary, and no incentive was provided. Responses were treated confidentially and reported in aggregate. The primary analytic cohort included respondents with complete survey data for the main telehealth experience items. Blank or incomplete survey records were excluded from the primary analysis.

Telehealth Context

At the time of the survey, telehealth visits at our institution were primarily conducted through Epic or MyChart video visits. Telephone visits were used when audiovisual connection could not be established or when technical issues required conversion from video to audio-only care. Survey items therefore focused primarily on video-based telehealth, while also capturing the frequency of conversion to audio-only visits due to technical difficulties.

Survey Instrument Development

The survey instrument was developed by the study investigators and was administered using Qualtrics. The survey was generally inspired by previously published surveys evaluating telehealth experiences, satisfaction, communication, and technology use [8-11]. Items were newly adapted for oncology-specific workflows, including visit modality decision-making, perceived adequacy of remote physical examination, and clinician confidence across phases of cancer care, such as preoperative and postoperative care, chemotherapy, and radiation treatment. The final instrument included 27 closed- and open-ended items and is provided as Multimedia Appendix 1. The survey was reviewed by study investigators with expertise in oncology, telehealth, and cancer care delivery for face validity, clinical relevance, and clarity before distribution. Formal pilot testing, cognitive interviewing, psychometric validation, and reliability testing were not performed.

Quantitative Analysis

Descriptive statistics were used to summarize respondent characteristics and survey responses. Categorical variables were reported as counts and percentages. Continuous variables were summarized using mean and SD or median and IQR, as appropriate. For key proportions, 95% CIs were calculated using the Wilson method. Likert-scale responses were collapsed for selected analyses to improve interpretability; responses

XSL•FO
RenderX

indicating at least moderate confidence were grouped as moderate or high confidence.

Exploratory chi-square tests were used to compare collapsed confidence proportions across specialty groups, McNemar tests were used for paired within-specialty comparisons across phase-of-care items, and a Friedman test was used to evaluate clinician likelihood of offering telehealth across increasing patient age categories. Given the modest sample size and exploratory design, statistical comparisons were interpreted cautiously.

Qualitative Analysis

Open-ended responses were analyzed using manual inductive coding informed by grounded theory principles. Two investigators independently reviewed and coded the open-ended responses. Codes were compared, and disagreements were resolved through consensus.

Recurring concepts related to telehealth decision-making, barriers, facilitators, and clinician experience were grouped into themes when similar concepts were identified across responses from at least 2 clinicians. Themes were then reviewed and refined to ensure that they reflected the underlying responses and were clinically interpretable. Representative quotations were selected to illustrate major themes. Qualitative analysis was conducted using exported survey responses rather than dedicated qualitative analysis software.

Ethical Considerations

This study was approved by the University of California San Diego Institutional Review Board (IRB 810808). Consent information was included within the survey, and respondents were required to proceed through the consent portion before accessing the survey. Participation was voluntary, no incentive was provided, and responses were treated confidentially and reported in aggregate.

Results

Respondent Characteristics

A total of 114 clinicians were invited to participate, including 32 (28.1%) surgical oncology clinicians, 49 (43%) medical oncology clinicians, 26 (22.8%) radiation oncology clinicians, and 7 (6.1%) palliative care clinicians. Overall, 65 (50%) survey records were captured, of which 57 (87.7%) were complete and included in the primary analysis, yielding a complete response rate of 50% (57/114). Eight blank or incomplete records were excluded.

Among the 57 respondents, 49 (86%) were physicians and 8 (14%) were advanced practice clinicians. Respondents included surgical oncology clinicians (21/57, 36.8%), medical oncology clinicians (18/57, 31.6%), radiation oncology clinicians (16/57, 28.1%), and palliative care clinicians (2/57, 3.5%). Specialty-specific complete response rates were 65.6% (21/32) for surgical oncology, 36.7% (18/49) for medical oncology, 61.5% (16/26) for radiation oncology, and 28.6% (2/7) for palliative care. Mean respondent age was 46.3 (SD 8.7) years. Most respondents were male (34/57, 59.6%) and non-Hispanic (54/57, 94.7%). Full demographic and professional characteristics are shown in Table 1.

XSL•FO
RenderX

Table . Demographic and professional characteristics of oncology clinicians participating in the telehealth perceptions survey (N=57).

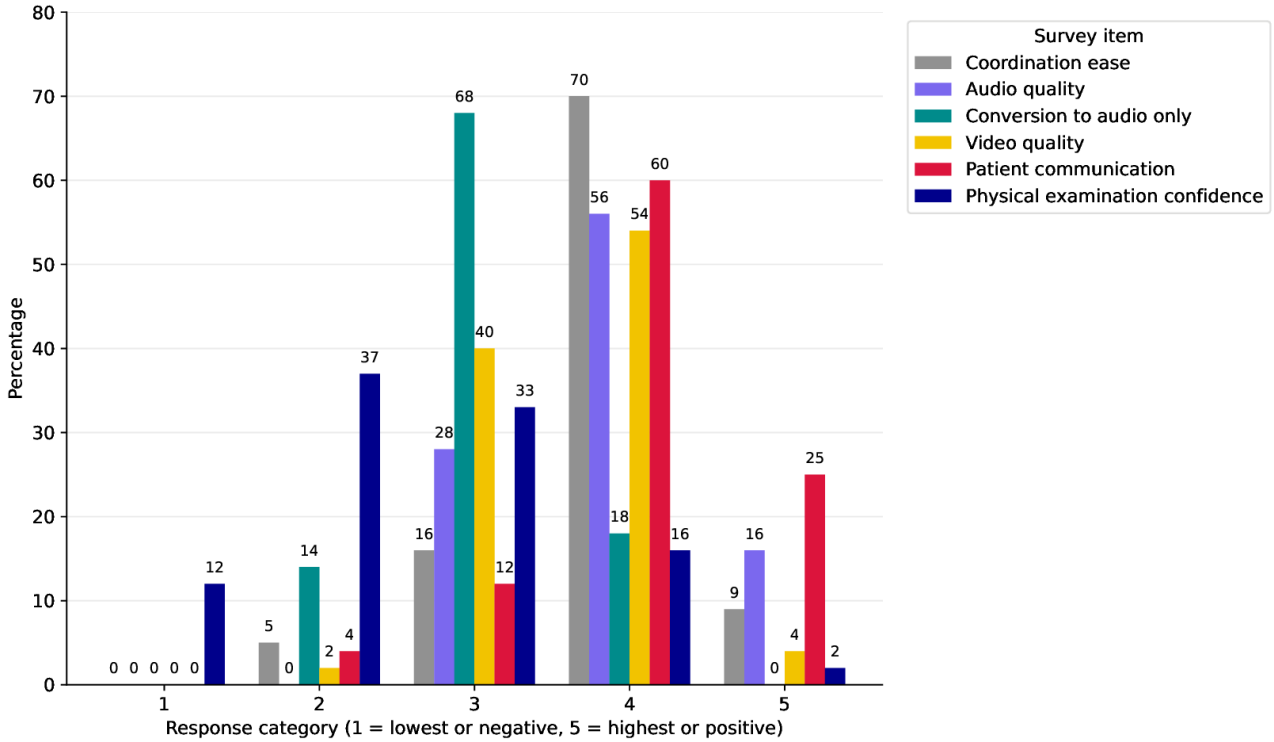
Characteristics	Participants
Age (y), mean (SD)	46.3 (8.7)
Sex, n (%)	
Male	34 (59.6)
Female	23 (40.4)
Race, n (%)	
Asian	12 (21.1)
Black or African American	4 (7)
Native Hawaiian or Pacific Islander	1 (1.8)
White	35 (61.4)
Other	5 (8.8)
Ethnicity, n (%)	
Hispanic	3 (5.3)
Non-Hispanic	54 (94.7)
Professional role, n (%)	
Physician	49 (86)
Advanced practice provider	8 (14)
Specialty, n (%)	
Medical oncology	18 (31.6)
Palliative care	2 (3.5)
Radiation oncology	16 (28.1)
Surgical oncology	21 (36.8)

Experience With Telehealth

Clinician experience with telehealth is summarized in Figure 1 and Table 2. In Figure 1, responses were reported on 5-point Likert scales. For coordination, audio quality, video quality, communication, and physical examination confidence, higher scores indicated more favorable responses. For audio-only conversion, higher scores indicated more frequent conversion. Most respondents reported that coordinating a telehealth visit was easy or very easy (45/57, 78.9%; 95% CI 66.7% - 87.5%). Audio quality was rated good or very good by 41 of 57 (71.9%; 95% CI 59.2% - 81.9%) respondents, while video quality was rated good or very good by 33 of 57 (57.9%; 95% CI 45.0% - 69.8%) respondents. Technical issues were common: 49 of 57 (86%; 95% CI 74.7% - 92.7%) clinicians reported at least occasional conversion to audio-only visits due to audiovisual problems, including 10 (17.5%; 95% CI 9.8% - 29.4%) clinicians who reported frequent conversion.

XSL•FO
RenderX

Figure 1. Clinician experience with telehealth across six domains.



XSL•FO
RenderX

Table . Detailed survey responses for telehealth experience and visit modality decision factors[a].

Domain or survey item	Response summary	Responses, n (%)	95% CI
Telehealth experience (n=57)			
Telehealth coordination	Easy or very easy	45 (78.9)	66.7 - 87.5
Audio quality	Good or very good	41 (71.9)	59.2 - 81.9
Video quality	Good or very good	33 (57.9)	45.0 - 69.8
Audio-only conversion due to technical issues	Occasionally, frequently, or very frequently	49 (86)	74.7 - 92.7
Patient communication by telehealth	Somewhat or extremely satisfied	48 (84.2)	72.6 - 91.5
Confidence in telehealth physical examination	Moderate or high confidence	29 (50.9)	38.3 - 63.4
Visit modality decision factors (n=57)			
Reviewed patient address before visit	Yes	22 (38.6)	27.1 - 51.6
Patient distance influenced decision to offer telehealth	Yes	45 (78.9)	66.7 - 87.5
Need for physical examination influenced decision to offer telehealth	Yes	53 (93)	83.3 - 97.2
Race or ethnicity influenced decision to offer telehealth	Yes	0 (0)	0.0 - 6.3
Primary language influenced decision to offer telehealth	Yes	25 (43.9)	31.8 - 56.7
Interpreter availability and coordination for telehealth	Somewhat or extremely satisfied	12 (21.1)	12.5 - 33.3
Interpreter availability and coordination for telehealth	Somewhat or extremely dissatisfied	31 (54.4)	41.6 - 66.6
Likelihood of offering telehealth by patient age (y), (n=57)			
Patients younger than 65 years	Somewhat or extremely likely	34 (59.6)	46.7 - 71.4
Patients aged 65-75 years	Somewhat or extremely likely	31 (54.4)	41.6 - 66.6
Patients aged 75-85 years	Somewhat or extremely likely	28 (49.1)	36.6 - 61.7
Patients older than 85 years	Somewhat or extremely likely	24 (42.1)	30.2 - 55.0
Specialty-specific confidence in telehealth[b]			
Surgical oncology: preoperative (n=21)	Moderate or high confidence	11 (52.4)	32.4 - 71.7
Surgical oncology: postoperative (n=21)	Moderate or high confidence	15 (71.4)	50.0 - 86.2
Medical oncology: prechemotherapy (n=18)	Moderate or high confidence	14 (77.8)	54.8 - 91.0
Medical oncology: during chemotherapy (n=17)	Moderate or high confidence	13 (76.5)	52.7 - 90.4
Radiation oncology: preradiation (n=16)	Moderate or high confidence	14 (87.5)	64.0 - 96.5
Radiation oncology: during radiation (n=16)	Moderate or high confidence	15 (93.8)	71.7 - 98.9

[a]Percentages may not total 100 due to rounding. 95% CIs were calculated using the Wilson method. Moderate or high confidence includes responses indicating at least moderate confidence.

[b]For specialty-specific confidence items, “not applicable” responses were excluded from denominators.

XSL•FO
RenderX

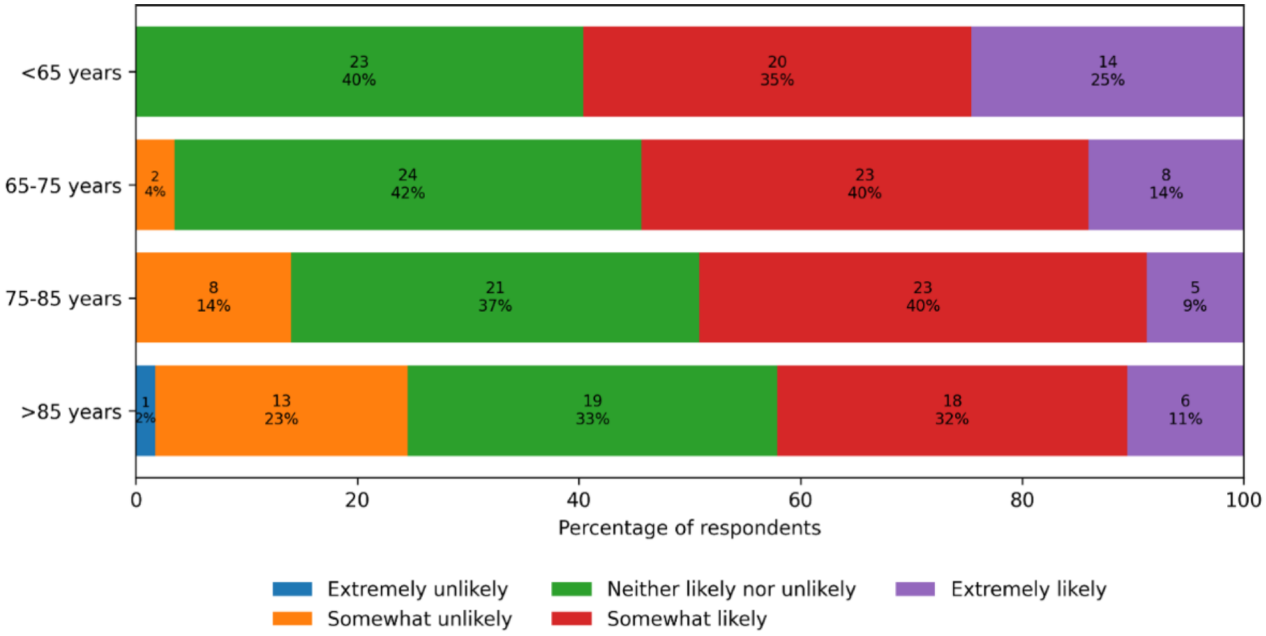
Despite these technical limitations, communication satisfaction was high. Of 57 respondents, 48 (84.2%; 95% CI 72.6% - 91.5%) were somewhat or extremely satisfied with patients' ability to communicate questions and concerns by telehealth. In contrast, confidence in the remote physical examination was more limited: 29 of 57 (50.9%; 95% CI 38.3% - 63.4%) respondents reported moderate or high confidence in performing a physical examination by telehealth, while 28 of 57 (49.1%; 95% CI 36.6% - 61.7%) respondents reported being not confident or not very confident.

Decision Factors for Offering Telehealth

Factors influencing clinicians' decisions to offer telehealth are summarized in Table 2. The need for a physical examination was the most frequently reported factor influencing visit modality selection, cited by 53 of 57 (93%; 95% CI 83.3% - 97.2%) clinicians. Patient distance from clinic also strongly influenced modality selection, cited by 45 of 57 (78.9%; 95% CI 66.7% - 87.5%) clinicians. Although distance influenced decision-making for most respondents, only 22 of 57 (38.6%; 95% CI 27.1% - 51.6%) clinicians reported reviewing a patient's address before a scheduled visit.

Patient age also influenced telehealth decision-making (Figure 2). The proportion of clinicians reporting that they were somewhat or extremely likely to offer telehealth declined across increasing age categories: 34 of 57 (59.6%; 95% CI 46.7% - 71.4%) for patients younger than 65 years, 31 of 57 (54.4%; 95% CI 41.6% - 66.6%) for patients aged 65 to 75 years, 28 of 57 (49.1%; 95% CI 36.6%-61.7%) for patients aged 75 to 85 years, and 24 of 57 (42.1%; 95% CI 30.2% - 55.0%) for patients older than 85 years. This decline across age categories was statistically significant in exploratory ordinal analysis (Friedman test: $P<.001$). No clinician reported that race or ethnicity influenced their decision to offer telehealth. In contrast, 25 of 57 (43.9%; 95% CI 31.8% - 56.7%) clinicians reported that a patient's primary language influenced their decision to offer telehealth.

Figure 2. Likelihood of offering telehealth by patient age group. Bars show the distribution of clinician responses across likelihood categories for each patient age group. Numbers within bars indicate respondent counts and percentages. Percentages may not total 100 due to rounding.



Interpreter-related responses were less favorable than other telehealth domains. Only 12 of 57 (21.1%; 95% CI 12.5% - 33.3%) respondents reported being somewhat or extremely satisfied with obtaining an interpreter for telehealth visits, while 31 of 57 (54.4%; 95% CI 41.6% - 66.6%) respondents reported being somewhat or extremely dissatisfied. Open-ended responses emphasized that interpreter integration into telehealth visits was technically and logistically challenging and that in-person interpretation was often preferred for complex conversations.

Qualitative Themes

Thematic analysis of open-ended responses identified several recurring themes related to telehealth decision-making. When discussing patient distance from the clinic, clinicians emphasized travel burden, cost and time toxicity, access to care, patient preference, and clinical appropriateness. Telehealth was viewed as particularly useful when the visit involved review of laboratory results, imaging, treatment planning, routine follow-up, or postoperative care without a need for physical examination. Representative comments included that telehealth can reduce unnecessary travel and improve follow-up access for patients who live far from the cancer center.

When discussing primary language, clinicians described interpreter use during telehealth visits as cumbersome, time-consuming, and technically unreliable. Several respondents stated that communication was easier when interpreters were present in person, particularly for visits requiring complex counseling or visual explanations. These themes are summarized with representative quotations in Table 3.

XSL•FO
RenderX

Table . Thematic analysis of qualitative responses regarding patient distance and primary language as factors influencing the decision to offer telehealth.

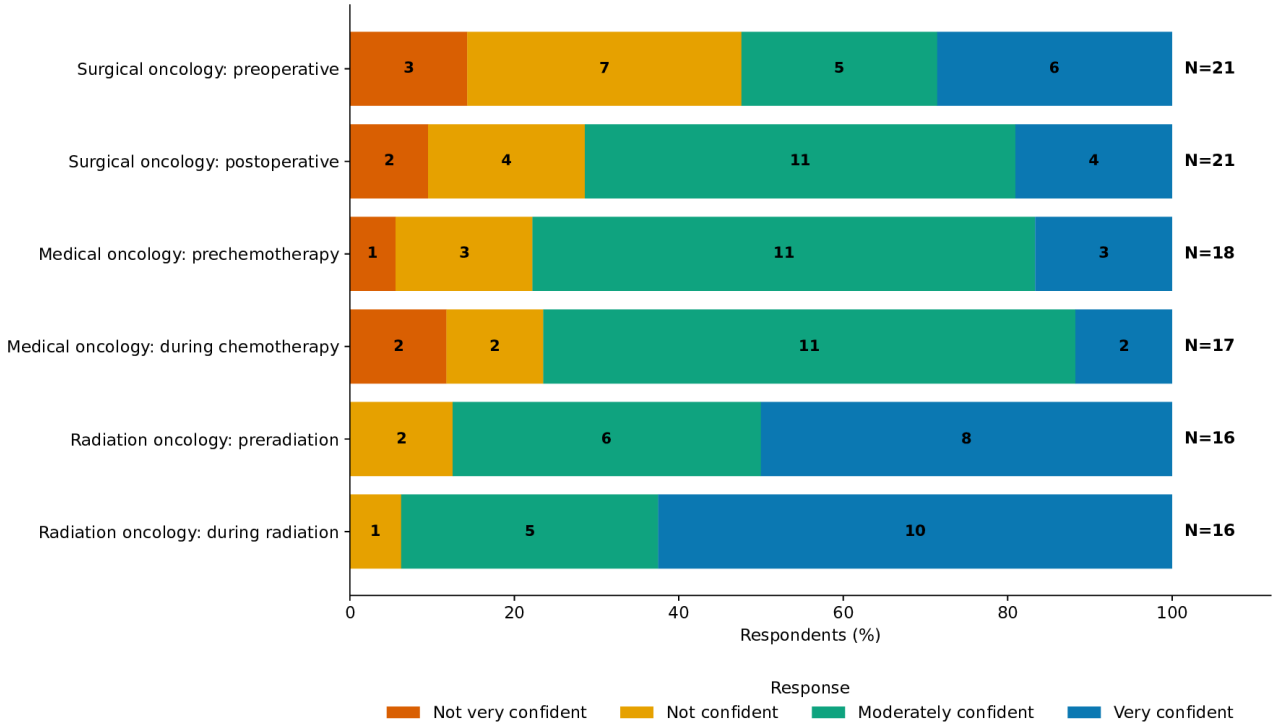
Theme summary	Representative quotes
Does a patient's distance from your clinic influence your decision to offer a telehealth visit?	
Cost, time, and financial toxicity	• "Telehealth avoids the ridiculous parking fee the university charges our patients." • "Unnecessary time spent in a car or bus can be avoided."
Clinical judgment and visit type	• "I feel more confident in-person for breast patients but will offer telehealth for initial consults if travel is difficult." • "A physical exam is not necessary and they live far away, I offer a telehealth visit."
Access, compliance, and equity	• "I want their distance to not be a barrier for follow-up care." • "It allows better access to care and increases patient compliance."
Legal and geographic considerations	• "I was informed that the patient needs to legally be in California at the time of telehealth visit."
Patient preference and flexibility	• "Patients often try to switch their visit to telemed due to distance or traffic." • "If patient mentions travel time, I consider telehealth."
Does a patient's primary language influence your decision to offer a telehealth visit?	
Interpreter use is challenging via telehealth	• "Translation services are spotty for telehealth visits at best, and rates of technical difficulties are much higher." • "I need a translator for telehealth so find it very difficult to do that."
Preference for in-person interpretation	• "Interpreters are best in person with actual live humans." • "Understanding each other is much easier if at least some of the people involved are in person."
Communication barriers and patient comprehension	• "I often like to draw figures/pictures for my patients which help them better understand certain concepts." • "More difficult to coordinate translator on video. Communication and connection issues make it hard."
Impact on visit quality and duration	• "It takes longer. Difficult to use an interpreter via telehealth." • "It is cumbersome to request an interpreter over telephone and/or video."
Workarounds and limitations	• "If the patient has a family member who is willing/able to translate, then I am more keen on using a telehealth visit." • "I have not found a good way to do video telehealth with interpreters."

Experience by Specialty and Phase of Care

Confidence in telehealth varied across specialty-specific clinical scenarios (Figure 3). Among surgical oncology clinicians, 11 of 21 (52.4%; 95% CI 32.4% - 71.7%) reported moderate or high confidence in assessing patients by telehealth in the preoperative setting, compared with 15 of 21 (71.4%; 95% CI 50.0% - 86.2%) in the postoperative setting. This within-specialty increase was descriptive and was not statistically significant in paired exploratory analysis (McNemar test: P=.39).

XSL•FO
RenderX

Figure 3. Clinician confidence in telehealth by oncology specialty and phase of care. Bars represent the distribution of provider responses for each specialty-specific clinical scenario. Numbers within bars indicate respondent counts, and denominators on the right indicate the number of applicable respondents for each scenario. “Not applicable” responses were excluded from denominators.



Among medical oncology clinicians, 77.8% (14/18; 95% CI 54.8% - 91.0%) reported moderate or high confidence in assessing patients by telehealth before chemotherapy. For assessment while receiving chemotherapy, 76.5% (13/17; 95% CI 52.7% - 90.4%) of respondents with nonmissing and applicable responses reported moderate or high confidence. Among radiation oncology clinicians, 87.5% (14/16; 95% CI 64.0% - 96.5%) reported moderate or high confidence before radiation therapy and 93.8% (15/16; 95% CI 71.7% - 98.9%) reported moderate or high confidence during radiation therapy.

In exploratory comparisons across specialty-specific pretreatment scenarios, moderate or high confidence differed across surgical, medical, and radiation oncology clinicians (11/21, 52.4% vs 14/18, 77.8% vs 14/16, 87.5%; chi-square test: P=.049). Differences across postoperative, during chemotherapy, and during radiation therapy scenarios were not statistically significant (15/21, 71.4% vs 13/17, 76.5% vs 15/16, 93.8%; chi-square test: P=.23).

Overall Experience

Overall, clinicians viewed telehealth as a useful adjunct to in-person oncology care when used for clinically appropriate visits. Commonly cited advantages included improved access, reduced travel burden, greater flexibility, and support for longitudinal follow-up. Commonly cited areas for improvement included audiovisual reliability, interpreter integration, patient preparation and expectations, and improvements to the clinician-facing telehealth platform. Clinicians consistently emphasized that telehealth should complement rather than replace in-person care, particularly when physical examination or complex communication is required.

Discussion

Principal Findings

In this cross-sectional survey of oncology clinicians at an academic cancer center, telehealth was generally viewed as a useful adjunct to traditional in-person care. Clinicians reported high satisfaction with visit coordination and patient communication but identified audiovisual reliability, interpreter integration, and limitations of remote physical examination as important barriers. Confidence in telehealth varied by specialty and phase of care, reflecting differences in clinical workflows and the role of physical examination across surgical, medical, and radiation oncology.

Our findings suggest that clinician confidence with telehealth is context dependent rather than uniform across oncology care. Surgical oncology clinicians reported greater confidence in postoperative than preoperative telehealth assessment, which likely reflects the greater need for physical examination, procedural planning, and initial treatment counseling before surgery. In contrast, medical and radiation oncology clinicians reported relatively high confidence across treatment-related scenarios, where telehealth may be particularly useful for counseling, reviewing laboratory or imaging results, symptom assessment, and longitudinal follow-up. These findings support the concept that telehealth is best viewed as a modality matched to visit type rather than a universal replacement for in-person oncology care.

XSL•FO
RenderX

Telehealth Decision-Making and Workflow Implications

The need for physical examination and patient distance from the clinic were the most influential factors in clinicians' decisions to offer telehealth. This reflects a practical balance between clinical appropriateness and patient access. Clinicians described telehealth as a way to reduce travel burden, time toxicity, parking and transportation costs, and barriers to follow-up for patients living far from the cancer center. However, only a minority of respondents reported routinely reviewing a patient's address before a scheduled visit, suggesting that opportunities exist to make distance-informed telehealth triage more systematic.

Institutions could use these findings to develop clearer workflows for identifying visits that are well suited for telehealth. Examples include standardized visit-type criteria, previsit screening for the need for physical examination, workflows that identify patients with long travel distances, and patient-facing instructions to improve technical readiness before the visit. Embedding these processes into scheduling and electronic health record workflows may help reduce variability in how telehealth is offered while preserving clinician judgment.

Equity Considerations

Although telehealth has the potential to improve access, our findings also highlight equity concerns. Clinicians were less likely to favor telehealth for older patients and reported substantial difficulty using interpreter services during telehealth visits. These findings are consistent with prior work showing that telehealth adoption may vary by patient characteristics, language, technology access, and health system infrastructure [7,12]. Without intentional workflows, telehealth could inadvertently widen disparities among patients who already face barriers to cancer care.

Several practical interventions may mitigate these risks. Institutions could improve interpreter integration into video visit platforms, develop workflows for previsit interpreter scheduling, provide technical support before visits, and offer patient education materials in multiple languages. For older patients or patients with limited digital access, previsit test calls, caregiver inclusion, simplified login processes, and hybrid models that combine in-person and telehealth visits may help preserve access without compromising care quality. These strategies may be especially important in oncology, where complex counseling and shared decision-making are common.

Comparison With Prior Literature

Our results align with prior studies demonstrating broad acceptance of telehealth in cancer care while also identifying persistent technical, communication, and patient safety concerns [8,12,13]. Similar to prior work, we found that telehealth use depends on clinical context and specialty-specific workflows [6,13]. Our study adds to this literature by comparing clinician experiences across oncology subspecialties and by evaluating specialty-specific confidence across phases of cancer care. This provides a more nuanced view of where providers perceive telehealth to be most clinically useful and where in-person visits remain important.

Limitations

This study has several limitations. First, it was conducted at a single academic cancer center, and findings may not generalize to community practices, nonacademic settings, or institutions with different telehealth infrastructure. Second, although we were able to calculate an overall complete response rate of 50% (57/114), response rates varied by specialty, and nonresponse bias remains possible. Third, although the survey was reviewed by investigators for face validity and clinical relevance, it did not undergo formal psychometric validation, cognitive testing, or reliability testing. Fourth, subgroup sizes were modest, particularly for palliative care and for specialty-specific phase-of-care analyses, and exploratory statistical comparisons should therefore be interpreted cautiously.

Fifth, the study focused on clinician experience and did not include patient perspectives, objective visit outcomes, completion rates, or measures of clinical quality. Sixth, tumor-site specialty was not collected. This is important because the role of physical examination and perceived appropriateness of telehealth may differ by disease site, such as breast, genitourinary, gastrointestinal, thoracic, or central nervous system malignancies. Finally, telehealth workflows and reimbursement policies continue to evolve, and clinician experiences may change as technology, interpreter integration, and institutional practices improve.

Future Directions

Future studies should validate these findings in larger, multi-institutional cohorts and should include both clinician and patient perspectives. Additional work should evaluate disease site–specific telehealth appropriateness, identify visit types that can be safely and effectively conducted virtually, and test interventions to improve interpreter integration, technical reliability, and equitable access. As telehealth reimbursement policies and digital infrastructure continue to evolve, understanding clinician-level experiences across oncology specialties may help guide workflows that maintain care quality while improving patient access.

Conclusions

Telehealth is a useful adjunct to in-person oncology care across medical, surgical, and radiation oncology, but clinician confidence and perceived utility vary by specialty and phase of care. Our findings support using telehealth as a complementary care modality when physical examination is not essential and when virtual visits can reduce travel burden or improve access. Further integration of telehealth into oncology care will require intentional workflows, improved audiovisual reliability, better interpreter integration, and attention to equitable access for all patients.

XSL•FO
RenderX

Acknowledgments

Generative AI tools (ChatGPT) were used to assist with language editing, organization of revisions, and formatting of manuscript text and reviewer responses. All authors reviewed and approved the final manuscript and take full responsibility for the accuracy, integrity, and content of the work. Generative AI tools were not used to generate or analyze primary survey data or to perform the primary writing of the manuscript.

Funding

This work was supported by the National Institutes of Health (5U54CA285115-03).

Authors' Contributions

JJ-D conceptualized and designed the study. BHB performed data interpretation and drafted the manuscript. SRC contributed to data collection and assisted with data analysis. SA contributed to study oversight and manuscript editing. JDM, HM, MEM, RMC, AB, and HP Jr contributed to methodological guidance, interpretation of results, and critical revision of the manuscript. JJ-D supervised the study and critically revised the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The 27-item Qualtrics survey instrument distributed to oncology providers to assess telehealth experience, visit modality decision factors, interpreter use, likelihood of offering telehealth by patient age, and specialty-specific confidence across phases of cancer care.
[PDF File, 131 KB - cancer_v12i1e94610_app1.pdf]

References

1. Sirintrapun SJ, Lopez AM. Telemedicine in cancer care. Am Soc Clin Oncol Educ Book 2018 May 23;38:540-545. [doi: 10.1200/EDBK_200141] [Medline: 30231354]
2. Patt DA, Wilfong L, Toth S, et al. Telemedicine in community cancer care: how technology helps patients with cancer navigate a pandemic. JCO Oncol Pract 2021 Jan;17(1):e11-e15. [doi: 10.1200/OP.20.00815] [Medline: 33434450]
3. Demeke HB, Merali S, Marks S, et al. Trends in use of telehealth among health centers during the COVID-19 pandemic - United States, June 26-November 6, 2020. MMWR Morb Mortal Wkly Rep 2021 Feb 19;70(7):240-244. [doi: 10.15585/mmwr.mm7007a3] [Medline: 33600385]
4. Lonergan PE, Washington Iii SL, Branagan L, et al. Rapid utilization of telehealth in a comprehensive cancer center as a response to COVID-19: cross-sectional analysis. J Med Internet Res 2020 Jul 6;22(7):e19322. [doi: 10.2196/19322] [Medline: 32568721]
5. Grenda TR, Whang S, Evans NR3. Transitioning a surgery practice to telehealth during COVID-19. Ann Surg 2020 Aug;272(2):e168-e169. [doi: 10.1097/SLA.0000000000004008] [Medline: 32675529]
6. Neeman E, Kumar D, Lyon L, et al. Attitudes and perceptions of multidisciplinary cancer care clinicians toward telehealth and secure messages. JAMA Netw Open 2021 Nov 1;4(11):e2133877. [doi: 10.1001/jamanetworkopen.2021.33877] [Medline: 34817586]
7. Fischer SH, Ray KN, Mehrotra A, Bloom EL, Uscher-Pines L. Prevalence and characteristics of telehealth utilization in the United States. JAMA Netw Open 2020 Oct 1;3(10):e2022302. [doi: 10.1001/jamanetworkopen.2020.22302] [Medline: 33104208]
8. Arem H, Moses J, Cisneros C, et al. Cancer provider and survivor experiences with telehealth during the COVID-19 pandemic. JCO Oncol Pract 2022 Apr;18(4):e452-e461. [doi: 10.1200/OP.21.00401] [Medline: 34714706]
9. Brintzenhofeszoc K, Krok-Schoen JI, Pisegna JL, et al. Survey of cancer care providers' attitude toward care for older adults with cancer during the COVID-19 pandemic. J Geriatr Oncol 2021 Mar;12(2):196-205. [doi: 10.1016/j.jgo.2020.09.028] [Medline: 33144071]
10. Nguyen OT, Mason A, Khanna N, et al. Patient and caregiver experience with telehealth for surgical cancer care: a qualitative study. J Surg Oncol 2023 Jun;127(7):1203-1211. [doi: 10.1002/jso.27228] [Medline: 36883752]
11. Kirby DJ, Fried JW, Buchalter DB, et al. Patient and physician satisfaction with telehealth during the COVID-19 pandemic: sports medicine perspective. Telemed J E Health 2021 Oct;27(10):1151-1159. [doi: 10.1089/tmj.2020.0387] [Medline: 33512302]
12. Stahl M, Cheung J, Post K, Valin JP, Jacobs I. Accelerating virtual health implementation following the COVID-19 pandemic: questionnaire study. JMIR Form Res 2022 May 16;6(5):e32819. [doi: 10.2196/32819] [Medline: 35323115]
13. West HJ, Barzi A, Wong D. Telemedicine in cancer care beyond the COVID-19 pandemic: oncology 2.0? Curr Oncol Rep 2022 Dec;24(12):1843-1850. [doi: 10.1007/s11912-022-01332-x] [Medline: 36208400]

XSL•FO
RenderX

Abbreviations

IRB: institutional review board

Edited by M Balcarras; submitted 14.Mar.2026; peer-reviewed by A Samani, R Ismayilov; revised version received 01.Jul.2026; accepted 07.Jul.2026; published 19.Aug.2026.

Please cite as:
Baker B, Chintala SR, Azari S, Murphy JD, Madanat H, Martinez ME, Cripps RM, Bagrodia A, Parada Jr H, Javier-DesLoges J
Cancer Clinician Experiences With Telehealth Across Oncology Specialties: Cross-Sectional Survey Study
JMIR Cancer 2026;12:e94610
URL: https://cancer.jmir.org/2026/1/e94610
doi:10.2196/94610

© Benjamin Baker, Sakshith Reddy Chintala, Sarah Azari, James D Murphy, Hala Madanat, Maria Elena Martinez, Richard M Cripps, Aditya Bagrodia, Humberto Parada Jr, Juan Javier-DesLoges. Originally published in JMIR Cancer (https://cancer.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Authors' Reply: Urological Content Accuracy as an Unmeasured Dimension in Generative AI Chatbot Communication About Sexual Health After Prostate Cancer

Mats Christiansen[1,2*], RN, PGD, MMSc, PhL; Henrik Eriksson[3*], RN, RNT, PhD; Lisbeth Fagerström[2*], RN, MSc, PhD

[1]Department of Public Health and Caring Sciences, Uppsala University, BMC, Husargatan 3, Uppsala, Sweden

[2]Department of Natural and Health Sciences, Faculty of Science and Engineering, Åbo Akademi University, Vaasa, Ostrobothnia, Finland

[3]Section for Health Promotion and Care Sciences, University West, Trollhättan, Västra Götaland, Sweden

[*]all authors contributed equally

Corresponding Author:
Mats Christiansen, RN, PGD, MMSc, PhL
Department of Public Health and Caring Sciences, Uppsala University, BMC, Husargatan 3, Uppsala, Sweden

Related Articles:

https://cancer.jmir.org/2026/1/e81745

https://cancer.jmir.org/2026/1/e108419

(JMIR Cancer 2026;12:e109004) doi:10.2196/109004

KEYWORDS

auto-netnography; netnography; prostate cancer; sexual minority men; generative AI chatbots

We sincerely thank the authors [1] for their thoughtful engagement with our article [2] and for raising important questions regarding the clinical accuracy of AI-generated health information. We welcome this opportunity for scientific dialogue.

Clinical accuracy is unquestionably an essential dimension of generative AI in health care. The letter and our original article illuminate different dimensions of the same phenomenon. The letter examines the clinical validity and completeness of AI-generated information, whereas our study examined how that information is framed, enacted, and experienced within a situated interaction concerning sexual health after prostate cancer. These perspectives are complementary because they address different dimensions of AI-mediated supportive care.

This distinction was fundamental to the design of our study. As stated in the original article [2], the study was "not designed to evaluate clinical effectiveness, safety, or generalizable performance of chatbots," but instead sought to examine "how responses are framed and experienced in sexual minority contexts." Consistent with contemporary netnographic scholarship, our analytical interest, therefore, lies in the interaction itself rather than in validating the medical correctness of individual chatbot responses [3-5].

The observations presented in the letter [1] illustrate this distinction. The authors note, for example, that none of the chatbots challenged the hypothetical assumption that both radical prostatectomy and radiotherapy would likely include hormonal therapy. We agree that this represents an important clinical observation and precisely the type of finding that a benchmarking study is designed to identify. Within our study [2], however, the hypothetical patient scenario functioned as a standardized interactional prompt rather than a test of guideline concordance. It also highlights an important feature of real-world human-AI interaction: people seeking health information may formulate prompts from incomplete or inaccurate understandings of their condition or treatment. Whether generative AI identifies, challenges, or reproduces such assumptions represents an important question for future research.

Similarly, the observations concerning contextual inaccuracies reflect these complementary perspectives. Our study reported such observations descriptively because they formed part of the interactional experience under investigation rather than constituting outcome measures of clinical reliability.

We, therefore, do not view the letter as challenging the central contribution of our study. Rather, it complements that contribution by examining another important dimension of AI-mediated supportive care. Clinical benchmarking establishes whether AI-generated information is medically reliable. Interpretive inquiry examines how that information becomes meaningful through interaction. Together, these perspectives provide a more comprehensive understanding of generative AI as a digital adjunct in supportive cancer care.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Isler B, Bayraktar AM. Urological content accuracy as an unmeasured dimension in generative AI chatbot communication about sexual health after prostate cancer. JMIR Cancer 2026;12:e108419. [doi: 10.2196/108419]
2. Christiansen M, Eriksson H, Fagerström L. Generative AI chatbots as digital adjuncts for sexual health information after prostate cancer in men who have sex with men: auto-netnographic study. JMIR Cancer 2026 Feb 9;12:e81745. [doi: 10.2196/81745] [Medline: 41662502]
3. Howard L. Auto-netnography in education: unfettered and unshackled. In: Kozinets RV, Gambetti R, editors. Netnography Unlimited: Understanding Technoculture Using Qualitative Social Media Research: Routledge; 2020:217-240. [doi: 10.4324/9781003001430]
4. Kozinets RV. Researching AI chatbots, platforms and the metaverse: understanding today's netnography. In: Belk RW, Otnes C, editors. Handbook of Qualitative Research Methods in Marketing, 2nd edition: Edward Elgar Publishing; 2024:185-196. [doi: 10.4337/9781035302727.00025]
5. Kozinets RV, Gretzel U. Netnography evolved: new contexts, scope, procedures and sensibilities. Ann Tourism Res 2024 Jan;104:103693. [doi: 10.1016/j.annals.2023.103693]

Edited by M Balcarras; submitted 10.Aug.2026; this is a non–peer-reviewed article;accepted 18.Aug.2026; published 15.Sep.2026.

Please cite as:
Christiansen M, Eriksson H, Fagerström L
Authors' Reply: Urological Content Accuracy as an Unmeasured Dimension in Generative AI Chatbot Communication About Sexual Health After Prostate Cancer
JMIR Cancer 2026;12:e109004
URL: https://cancer.jmir.org/2026/1/e109004
doi:10.2196/109004

© Mats Christiansen, Henrik Eriksson, Lisbeth Fagerström. Originally published in JMIR Cancer (https://cancer.jmir.org), 15.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Urological Content Accuracy as an Unmeasured Dimension in Generative AI Chatbot Communication About Sexual Health After Prostate Cancer

Bilgi Isler, MD; Ahmet Murat Bayraktar, MD

Department of Urology, Konya City Hospital, University of Health Sciences, İstiklal, Adana Çevre Yolu Cd. No:135/1, Konya, Türkiye

Corresponding Author:
Bilgi Isler, MD
Department of Urology, Konya City Hospital, University of Health Sciences, İstiklal, Adana Çevre Yolu Cd. No:135/1, Konya, Türkiye

Related Articles:

https://cancer.jmir.org/2026/1/e81745

Comment in: https://cancer.jmir.org/2026/1/e109004

(JMIR Cancer 2026;12:e108419) doi:10.2196/108419

KEYWORDS

auto-netnography; netnography; prostate cancer; sexual minority men; generative AI chatbots

Christiansen and colleagues [1] make a welcome contribution by shifting attention from correctness to interactional framing in a group whose informational needs remain underaddressed. The authors state clearly that they did not set out to assess clinical effectiveness or safety, and they caution in their limitations that the absence of obvious fabrication should not be read as evidence of accuracy. Reading the published outputs from a urological perspective, we offer three observations that may refine future work.

First, the initial prompt states that both radiotherapy and surgery would likely include hormonal therapy. This is a patient utterance, not a claim by the authors, and patients do hold such beliefs. Androgen deprivation is standard in combination with radiotherapy for intermediate- and high-risk disease, but neoadjuvant androgen deprivation around radical prostatectomy is not recommended [2]. What is notable is the response: none of the four systems queried or corrected it, and all four built counseling based on it. Silent acceptance of a premise that is not guideline-concordant is, in our view, among the most consequential failure modes in patient-facing use, and it falls outside the pragmatic accuracy definition used here.

Second, the two treatment arms differ categorically for receptive anal intercourse: after prostatectomy, the gland is absent and prostate-mediated pleasure is permanently lost [3], whereas after radiotherapy, the gland remains in situ and rectal toxicity becomes the limiting factor. Comparative data bear this out: climacturia is far commoner after prostatectomy and anodyspareunia after radiotherapy [4]. Only one output made this explicit. Another described treatment as affecting the sensitivity and function of the prostate, which does not apply once the gland has been removed. The passage quoted twice as an exemplar of compassionate precision proposes a 4‐ to 6-week interval for resuming receptive anal intercourse without reference to anodyspareunia, mucosal fragility, or proctopathy after irradiation.

Third, the statement that no hallucinated content was identified sits alongside a reported support group link misattributed to a Swedish service. We would classify that as a contextual hallucination rather than as a separate category, and the platform responsible is not named. A further unflagged example appears in Multimedia Appendix 2 in Christiansen et al [1], where a hereditary cancer organization is presented as a resource for LGBTQ+ (lesbian, gay, bisexual, transgender, queer) individuals with cancer.

These observations do not undermine the authors' interpretive aim but indicate where a complementary design would add value. We suggest three additions: (1) a prespecified urology-anchored content checklist scored independently by two clinicians, covering the prostatectomy versus radiotherapy distinction, anodyspareunia and rectal toxicity, climacturia, penile length change, and the expected duration and reversibility of androgen deprivation; (2) deliberate inclusion of premises that are not guideline-concordant, with correction rate as an explicit end point; and (3) per-model reporting of every identified error, with exact query dates, session conditions, and repeated generations, as the CHART (Chatbot Health Advice Reporting Tool) statement the authors followed recommends [5].

Communicative warmth and clinical adequacy are separable qualities. This study measures the first with care, and we would welcome a companion study that measures the second.

XSL•FO
RenderX

Funding

The authors declared no financial support was received for this work.

Authors' Contributions

BI conceived the letter, appraised the source article, and drafted the manuscript. AMB critically revised the manuscript. Both authors approved the final version.

Conflicts of Interest

None declared.

References

1. Christiansen M, Eriksson H, Fagerström L. Generative AI chatbots as digital adjuncts for sexual health information after prostate cancer in men who have sex with men: auto-netnographic study. JMIR Cancer 2026 Feb 9;12:e81745. [doi: 10.2196/81745] [Medline: 41662502]
2. Cornford P, van den Bergh RCN, Briers E, et al. EAU-EANM-ESTRO-ESUR-ISUP-SIOG Guidelines on Prostate Cancer-2026 update. Part I: screening, diagnosis, and local treatment with curative intent. Eur Urol 2026 Sep 1:S0302-2838(26)02114-7. [doi: 10.1016/j.eururo.2026.04.013] [Medline: 42457454]
3. Vedovo F, Lonati C, Suardi N, et al. Sexual outcomes in men who have sex with men who underwent radical prostatectomy. Prostate Cancer Prostatic Dis 2025 Mar;28(1):70-80. [doi: 10.1038/s41391-024-00861-9] [Medline: 38918583]
4. Dickstein DR, Bates AJ, Wheldon CW, et al. Treatment choice and sexual health outcomes in gay and bisexual men with prostate cancer. Prostate Cancer Prostatic Dis 2026 Sep;29(3):638-641. [doi: 10.1038/s41391-025-01007-1] [Medline: 40764815]
5. CHART Collaborative, Huo B, Collins GS, et al. Reporting guideline for chatbot health advice studies: the CHART statement. JAMA Netw Open 2025 Aug 1;8(8):e2530220. [doi: 10.1001/jamanetworkopen.2025.30220] [Medline: 40747871]

Abbreviations

CHART: Chatbot Health Advice Reporting Tool
LGBTQ+: lesbian, gay, bisexual, transgender, queer

Edited by M Balcarras; submitted 31.Jul.2026; this is a non–peer-reviewed article;accepted 18.Aug.2026; published 15.Sep.2026.

Please cite as:
Isler B, Bayraktar AM
Urological Content Accuracy as an Unmeasured Dimension in Generative AI Chatbot Communication About Sexual Health After Prostate Cancer
JMIR Cancer 2026;12:e108419
URL: https://cancer.jmir.org/2026/1/e108419
doi:10.2196/108419

© Bilgi Isler, Ahmet Murat Bayraktar. Originally published in JMIR Cancer (https://cancer.jmir.org), 15.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Letter to the Editor

Adapting Interventions for Home Hospice Caregivers Using Digital Health Innovation

Qing Huang[1], PhD

Department of Research & Data Analytics, Singapore Cancer Society, Singapore, Singapore

Corresponding Author:
Qing Huang, PhD
Department of Research & Data Analytics
Singapore Cancer Society
30 Hospital Boulevard, #16-02 NCCS Building
Singapore, 168583
Singapore
Phone: 65 64215863
Email: huang_qing@singaporecancersociety.org.sg

Related Article:

Comment on: https://cancer.jmir.org/2025/1/e71596/

(JMIR Cancer 2026;12:e81589) doi:10.2196/81589

KEYWORDS

palliative; hospice; home care; cancer; meta-synthesis

The recent meta-synthesis by Deng et al [1] on the experiences of family caregivers in home-based cancer hospice care was read with great interest. The authors presented a precise summary of their needs and perceptions and, in doing so, highlighted a critical gap in tailored interventions for supporting this vulnerable group. This is a gap that warrants further consideration.

This gap, while concerning, is understandable. The time and physical confinement inherent in home hospice care limit caregivers' ability to participate in research or access traditional in-person support. Deng et al's [1] review emphasized the dual challenges of this role: the immense burden of being "physically and emotionally present" and the critical need for "sharing responsibility" with palliative care teams, community partners, and family. While this shared support makes caregivers feel less isolated, the home hospice care setting creates logistical barriers to effective collaboration. To make this sharing of responsibility more accessible for home hospice caregivers, we should leverage innovative technology to connect the caregivers with their external support network. This is where we can adopt proven therapies like mindfulness-based stress reduction, delivering them through flexible digital methods that fit the specific constraints of the home hospice setting.

Encouragingly, some recent studies have validated this approach. For example, a pilot study using virtual reality demonstrated that technology can provide homebound caregivers with much-needed individual psychological respite through immersive nature scenes [2]. On a larger scale, a randomized controlled trial of the "Symptom Care at Home" intervention proved that an automated digital coaching system could significantly reduce caregiver burden by connecting them to their care team when symptoms escalated [3]. These studies illustrate the potential of novel ways of delivering support to overcome the practical barriers of home hospice care.

The work of Deng et al [1] should be a catalyst for action. To translate these promising findings into real-world practice, robust collaboration is essential for adapting and testing such interventions. An ideal model would involve researchers and developers co-designing digital tools with caregivers, an approach shown to improve the usability and adoption of digital health interventions [4]. This also calls for a partnership between hospice providers and community organizations to bring these tools into routine care, creating hybrid support models that combine scalable technology for both individual coping and shared care. Through such collaborative efforts, we can deliver interventions that truly empower caregivers to navigate one of life's most profound challenges.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Editorial Notice

The corresponding author of "Family Experiences, Needs, and Perceptions in Home-Based Hospice Care for Patients With Terminal Cancer: Meta-Synthesis and Systematic Review" declined to respond to this letter.

References

1. Deng XM, Hounsri K, Lopez V, Tam WW. Family experiences, needs, and perceptions in home-based hospice care for patients with terminal cancer: meta-synthesis and systematic review. JMIR Cancer 2025 Jun 19;11:e71596 [FREE Full text] [doi: 10.2196/71596] [Medline: 40537084]
2. Patano A, Alanazi M, Lehto R, Goldstein D, Wyatt G. A nature-immersive virtual reality intervention to support hospice family caregivers: qualitative findings from a pilot study. Asia Pac J Oncol Nurs 2024 Dec;11(12):100616 [FREE Full text] [doi: 10.1016/j.apjon.2024.100616] [Medline: 39759956]
3. Mooney K, Whisenant MS, Wilson CM, Coombs LA, Lloyd J, Alekhina N, et al. Technology-assisted mHealth caregiver support to manage cancer patient symptoms: a randomized controlled trial. J Pain Symptom Manage 2023 Jul;66(1):33-43. [doi: 10.1016/j.jpainsymman.2023.02.320] [Medline: 36889453]
4. Duffy A, Boroumandzad N, Sherman AL, Christie G, Riadi I, Moreno S. Examining challenges to co-design digital health interventions with end users: systematic review. J Med Internet Res 2025 Mar 14;27:e50178 [FREE Full text] [doi: 10.2196/50178] [Medline: 40085834]

Edited by N Cahill; submitted 31.Jul.2025; this is a non–peer-reviewed article;accepted 18.Dec.2025; published 14.Jan.2026.

Please cite as:
Huang Q
Adapting Interventions for Home Hospice Caregivers Using Digital Health Innovation
JMIR Cancer 2026;12:e81589
URL: https://cancer.jmir.org/2026/1/e81589
doi:10.2196/81589
PMID:

©Qing Huang. Originally published in JMIR Cancer (https://cancer.jmir.org), 14.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Financial Toxicity of Hematologic Malignancy Therapies, Including Cellular Therapy and Its Impact on Access to Care: Prospective Pilot Study

Lemchukwu Amaeshi[1,2], MD; Lauren Laufer[1,3], BA, MD; Kateryna Fedorov[4,5], MD; Alex Sisto[1,6], MD; David Levitz[4], MD; Rosmi Mathew[1], DNP, APRN, FNP-BC; Karen Wright[1†], RN, MS; Dennis Cooper[4], MD; Mendel Goldfinger[4], MD; Ioannis Mantzaris[4], MD; R Alejandro Sica[4], MD; Kira Gritsman[4], MD, PhD; Marina Konopleva[4,7], MD, PhD; Ridhi Gupta[4], MD; Noah Kornblum[4], MD; Chenxin Zhang[2], MS; Mimi Kim[2], SCD; Amit Verma[4], MD; Bruce Rapkin[1,2], PhD; Aditi Shastri[4], MD; Nishi Shah[4], MD, MPH

[1]Department of Medicine, Montefiore Medical Center, Bronx, NY, United States

[2]Department of Medicine, Albert Einstein College of Medicine, Bronx, NY, United States

[3]Department of Medicine, University of California, Los Angeles, CA, United States

[4]Department of Oncology, Montefiore Medical Center and Albert Einstein College of Medicine, 111 E 210th St, Bronx, NY, United States

[5]Oncology, Vanderbilt University Medical Center, Nashville, TN, United States

[6]Department of Internal Medicine, Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ, United States

[7]Department of Leukemia, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[†]deceased

Corresponding Author:
Nishi Shah, MD, MPH
Department of Oncology, Montefiore Medical Center and Albert Einstein College of Medicine, 111 E 210th St, Bronx, NY, United States

Abstract

Background: Patients with cancer often face significant financial challenges, known as financial toxicity (FT), which is associated with reduced quality of life. Patients with hematologic malignancies (HMs) are especially vulnerable due to intensive and prolonged treatments, frequent hospital visits, and a high risk of complications. While FT affects many in the general population, it is particularly severe among racial and ethnic minorities, especially those below the poverty line. To our knowledge, no studies have specifically examined FT in this vulnerable group in the United States.

Objective: This study aimed to evaluate the severity of FT in patients receiving treatment for HMs in a socioeconomically underserved population, explore sociodemographic factors that may predict the severity of FT, and evaluate the subjective experiences of these patients as they relate to FT.

Methods: We conducted a prospective, observational, longitudinal study at the Montefiore Cancer Center's outpatient department in the Bronx, New York, from October 1, 2022, to October 30, 2023. Participants included either adult patients newly diagnosed (ND) with HMs or those already diagnosed, undergoing cellular therapy (CT). The severity of FT was assessed using the validated Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy (COST-FACIT) questionnaire. Additionally, an investigator-designed questionnaire was developed to gather sociodemographic data and evaluate the subjective effects of financial burden on patient care. Patients in both the ND and CT groups were followed for 90 days. Data collection occurred at their initial presentation, as well as on days 30 and 90.

Results: Ninety patients participated in the study (ND=52 and CT=38). The median age was 59 (IQR 44-66) years, with 27% (n=24) African American and 55% (n=48) Hispanic. Overall, 75% (n=67) of participants experienced some degree of FT, most with mild FT at baseline (day 0, median COST-FACIT score=19.4). In the CT group, FT worsened significantly over time, with a decline in median COST-FACIT scores from 19.9 at day 0 to 15.5 on day 90 (P=.02). In a multivariable linear regression model, race and ethnicity were a significant predictor of FT burden: identifying as African American or Hispanic was associated with a significantly lower COST-FACIT score (ie, higher FT) compared to non-Hispanic White participants (B=−3.08, P=.04, 95% CI −6.05 to −0.12). Additionally, over half of ND and CT participants reported difficulty affording basic necessities (ND: 28/52, 54%; CT: 23/38, 61%) and concerns regarding transportation access and costs (ND: 26/50, 52%; CT: n=18/38, 47%).

Conclusions: FT is prevalent among patients with HMs receiving care in underserved populations, and the burden is significantly higher among African American and Hispanic populations.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e68101) doi:10.2196/68101

KEYWORDS

financial toxicity; hematologic malignancies; cancer; cellular therapy; quality of life; socioeconomically; longitudinal study; questionnaire; cancer care; African American; Hispanic; toxicity; survival rates; survival; medical care; adherence; financial literacy

Introduction

Cancer continues to be a significant public health concern in the United States. According to the American Cancer Society, in 2022, there were nearly 2 million new cancer cases and over 600,000 cancer-related deaths in the United States [1]. Receiving a new cancer diagnosis brings untold physical, emotional, and psychological distress to the patient and their caregivers. Although survival rates of patients with cancer have improved over the years due to advances in diagnosis and treatment, the rising cost of cancer care has become a significant challenge for patients and their providers in the US health care system. Financial toxicity (FT) has emerged in oncologic care to describe the psychological, material, and behavioral hardships arising from the economic burden of cancer [2].

Indeed, several studies have reported a close association between FT and reduced quality of life, delays in seeking medical care, nonadherence with treatment, emotional and psychological distress, and reduced overall survival in patients with cancer [3-5]. The impact of FT is particularly pronounced in patients with hematologic malignancies (HMs) as they must deal with the high cost of therapy, especially with the shift from conventional chemotherapy to immunotherapy, multiple infusion visits, prolonged hospitalizations due to life-threatening presentations, long duration of intensive treatment, and treatment-related complications [6,7]. In 2014, the average cumulative costs of hematologic cancer care in the United States ranged from approximately US $200,000 for chronic leukemias to greater than US $800,000 for acute leukemias within the first 3 years of treatment. In comparison, the cost for lung cancer was around US $250,000 and that for colorectal cancer was approximately US $150,000 [8].

The severity of FT is also determined by the patient's sociodemographic and socioeconomic factors. Extremes of age, Black race, lower income level, limited ability to provide basic household needs, unemployment, and insurance status are associated with worse FT [2,9-11]. The factors associated with worse FT are predominant in households living below the federal poverty line. This population primarily comprises Native American, Black, and Hispanic individuals [11].

Despite growing attention to FT, most studies were conducted predominantly on the White population, who are often insured. To the best of our knowledge, no similar studies in the United States have specifically targeted racial and ethnic minorities with HM in underserved areas in the United States. This study aims to assess the severity of FT, examine relevant sociodemographic factors influencing FT, and explore the subjective experience of FT among patients with HMs.

This study was conducted in a large academic hospital in the Bronx, where more than 25% of the population lives below the federal poverty line [12].

Methods

Study Design and Setting

This prospective observational study was conducted over 12 months, from October 1, 2022, to October 30, 2023, at Montefiore Medical Center in Bronx, New York. The study took place in both outpatient and inpatient hematologic oncologic units of the Montefiore Cancer Center, a quaternary academic center serving a predominantly low-income, racially and ethnically diverse population.

Participant Eligibility and Recruitment

Eligible participants included adults over 18 years old with a HM (ie, acute or chronic leukemia, Hodgkin lymphoma, non-Hodgkin lymphoma, or multiple myeloma) who either received a new diagnosis or were being evaluated for cellular therapy (CT), such as autologous or allogeneic stem cell transplantation (SCT) or chimeric antigen receptor T-cell (CAR-T) therapy.

Patients were excluded if (1) they were being seen for conditions other than HMs, (2) they were seeking a second opinion after treatment at a different institution, (3) they presented after recurrence ineligible for autologous or allogeneic SCT and CAR-T therapy, (4) they were being evaluated for a second autologous SCT as part of tandem autologous transplantation or for recurrent myeloma, or (5) they lacked capacity or were non-English or non-Spanish speakers.

Recruitment took place during routine clinical visits and on the oncology floors. Eligible patients were identified by treating providers or study personnel and were invited to participate on a rolling basis. Participants were divided into two cohorts: (1) newly diagnosed (ND) and (2) undergoing CT (SCT or CAR-T).

Data Collection and Measurements

Data were collected in person at three time points: baseline (day 0), day 30, and day 90. At each time point, the participants completed two instruments:

1. Sociodemographic and Subjective Impact Questionnaire: The research team developed an investigator-designed, bilingual (English and Spanish) questionnaire to gather information on demographics (age, sex, race and ethnicity, income, education, employment, and insurance), cancer type, and subjective experience of financial burden, including effects on basic needs and access to care.
2. Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy: FT was assessed using the validated Comprehensive Score for Financial

XSL•FO
RenderX

Toxicity–Functional Assessment of Chronic Illness Therapy (COST-FACIT) tool [13]. It is categorized into grades 0 to 4, based on the level of FT severity, with scores ranging from 0 to 44. A score of 0 represents grade 0 or *severe toxicity*; 1 - 13, grade 2 or *moderate toxicity*; 14 - 25, grade 3 or *mild toxicity*; and >25, grade 4 or *no toxicity*. Hence, higher scores indicated less severe FT. The patients were assessed at three different time points: day 0, the initial visit time, day 30, and day 90.

Statistical Analysis

Data analysis was performed using SPSS version 29 (IBM Corp). Descriptive statistics summarized demographic characteristics and COST-FACIT scores. Frequencies and percentages were calculated for categorical variables, while medians and IQRs were used for continuous variables. Comparison of the severity of FT across the timelines was determined using the Friedman test. The Mann-Whitney *U* test and Kruskal-Wallis *H* test were used to examine significant differences in FT with respect to sociodemographic groups. Linear regression analysis identified sociodemographic predictors of FT. A two-tailed *P* value of <.05 was considered statistically significant. No adjustment for multiple comparisons was made due to the exploratory nature of the study.

Ethical Considerations

Study Approval

This study was conducted in accordance with ethical standards and received approval from the IRB at Albert Einstein College of Medicine and Montefiore Medical Center (approval number: IRB 2022 - 13798, approval date: 09/13/2022). The research involved human participants and adhered to the principles outlined in the Declaration of Helsinki.

Informed Consent

The provider (MD or nursing practitioner) or study personnel obtained informed consent during the initial visit at the HM clinic or for auto/allo SCT and CAR-T evaluation. All participants gave written informed consent before enrollment. The consent process included explaining the study's objectives, procedures, risks, and benefits and the fact that participation is voluntary. For patients unable to give written consent, verbal consent was obtained in the presence of a trained research coordinator, per institutional policy.

Privacy and Confidentiality

All collected data were deidentified prior to analysis to protect participant confidentiality. Study data were stored in secure, password-protected databases accessible only to authorized personnel. No identifiable personal health information was used in any publication or presentation.

Compensation

Participants were not financially compensated for their involvement in this study.

Results

Sociodemographic Characteristics

Over a period of 12 months, we recruited 90 patients who met eligibility criteria and consented to the study. The sociodemographic characteristics of these participants are shown in Table 1. Fifty-two of 90 (57%) patients were ND, whereas the rest (n=38, 43%) were either currently receiving or preparing to receive CT. The median age was 59 (IQR 44-66) years. There were more male patients (n=56, 62%), 24 (27%) patients were African American, and more than half (n=48, 55%) were Hispanic. Over 60% (n=55, 63%) of the participants were not employed, and over a third (n=33, 37.1%) were on Medicaid. Regarding HMs, most patients had plasma cell dyscrasias (n=33, 37%) compared to other HMs.

XSL•FO
RenderX

Table . Sociodemographic characteristics.

Sociodemographic characteristics	Newly diagnosed group (n=52)	Cellular therapy group (n=38)	Total (n=90)
Age, median (IQR)	59 (42.5 - 66.0)	58 (50.0 - 67.0)	59 (43.5 - 66.0)
Sex, n (%)			
Male	31 (60)	25 (66)	56 (62)
Female	21 (40)	13 (34)	34 (38)
Race, n (%)			
African American	15 (29)	8 (22)	24 (27)
Non-Hispanic White	6 (11)	8 (22)	14 (16)
Others	31 (60)	22 (56)	51 (57)
Ethnicity, n (%)			
Hispanic	28 (54)	20 (57)	48 (55)
Non-Hispanic	24 (46)	15 (43)	40 (45)
Employment status, n (%)			
Employed	21 (41)	11 (31)	32 (37)
Retired	13 (25)	9 (25)	22 (25)
On disability	9 (18)	8 (22)	17 (20)
Unemployed	8 (16)	8 (22)	16 (18)
Health insurance, n (%)			
Private	22 (43)	16 (42)	38 (43)
Medicare	10 (20)	8 (21)	18 (20)
Medicaid	19 (37)	14 (37)	33 (37)
Type of hematologic malignancy, n (%)			
Leukemias	21 (41)	10 (27)	31 (34)
Lymphomas	16 (31)	9 (24)	26 (29)
Plasma cell dyscrasias	14 (28)	18 (49)	33 (37)

Severity of Financial Toxicity Across Time

Overview

Table 2 and Figure 1 depict the median FT scores at various periods. Over 75% (n=67) of patients experienced some degree of FT, with a median COST score of 19.4 at baseline (day 0). While the median COST score indicates mild FT, many patients experienced moderate FT at the individual level. The CT group demonstrated significant changes in FT scores over time (day 0: 19.9; day 30: 19.0; day 90: 15.5; *P*=.02). When comparing FT severity between the ND and CT groups, as shown in Table 3, no statistically significant differences were observed at any period (day 0: *P*=.88; day 30: *P*=.54; day 90: *P*=.75).

XSL•FO
RenderX

Table . Financial toxicity at days 0, 30, and 90.

Financial toxicity	Day 0	Day 30	Day 90	*P* value
Newly diagnosed group				
Number of patients, n	50	39	32	
Median FT[a] (IQR)	16.5 (8.9-27.5)	17.0 (9.5-21.0)	17.5 (9.0-27.0)	.85
Grade category n (%)				
No toxicity	11 (22)	7 (18)	9 (28)	
Mild toxicity	17 (34)	18 (46)	10 (31)	
Moderate toxicity	21 (42)	14 (36)	13 (41)	
Severe toxicity	1 (2)	0 (0)	0 (0)	
Cellular therapy group				
Number of patients, n	36	29	16	
Median FT (IQR)	19.9 (9.0-27.0)	19.0 (13.8-28.0)	15.5 (12.3-32.3)	.02
Grade category n (%)				
No toxicity	10 (28)	8 (28)	5 (31)	
Mild toxicity	11 (31)	14 (48)	4 (25)	
Moderate toxicity	15 (42)	7 (24)	5 (31)	
Severe toxicity	0 (0)	0 (0)	0 (0)	
Total population				
Number of patients, n	88	69	48	
Median FT (IQR)	19.4 (8.9-27.0)	18 (10.0-25.0)	17 (10.0-17.5)	.31
Grade category n (%)				
No toxicity	21 (24)	15 (22)	14 (30)	
Mild toxicity	29 (33)	33 (48)	14 (30)	
Moderate toxicity	36 (41)	21 (30)	18 (39)	
Severe toxicity	2 (2)	0 (0)	0 (0)	

[a]FT: financial toxicity.

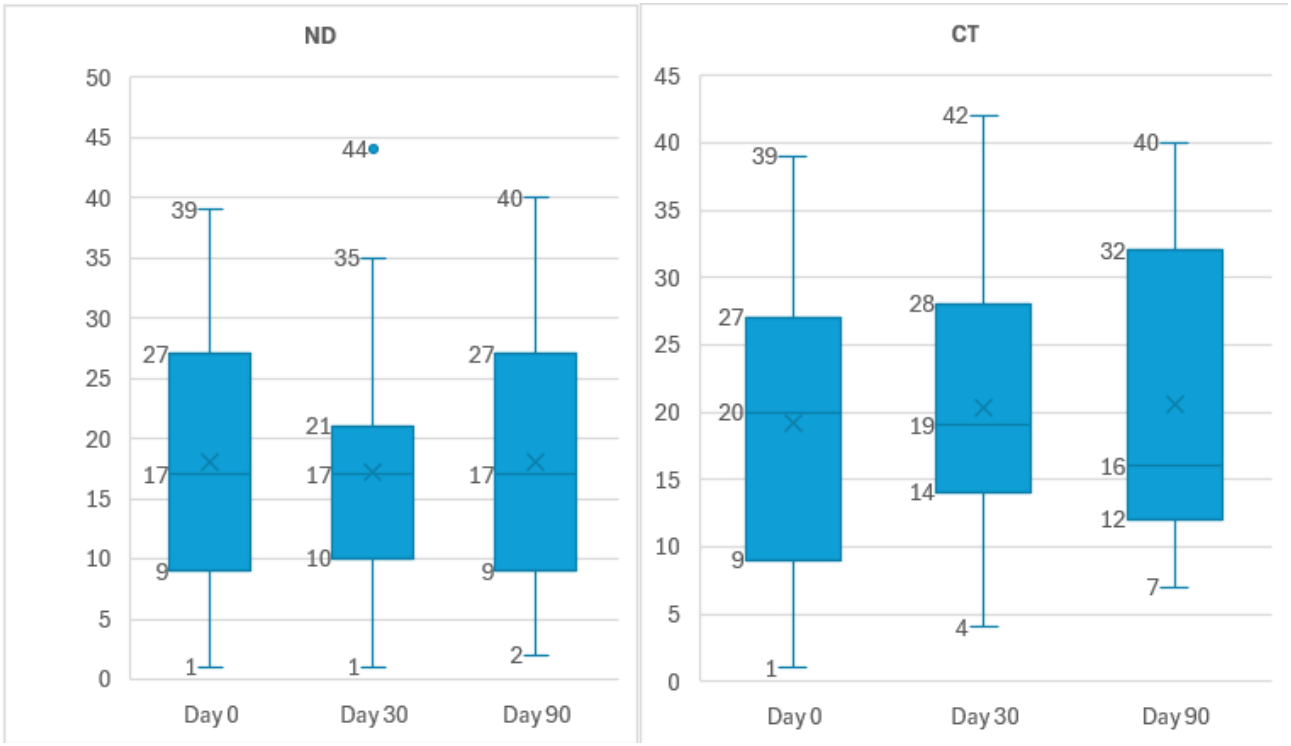
Table . Comparison of financial toxicity between newly diagnosed and cellular therapy groups on days 0, 30, and 90.

Days	ND[a]	CT[b]	*P* value
Day 0			
Number of patients, n	50	37	
Median FT[c] (IQR)	16.5 (8.9-27.5)	19.9 (9.0-27.0)	.88
Day 30			
Number of patients	36	32	
Median FT (IQR)	17.0 (9.5-21.0)	19.0 (13.8-28.0)	.54
Day 90			
Number of patients	31	15	
Median FT (IQR)	17.5 (9.0-27.0)	15.5 (12.3-32.3)	.75

[a]ND: newly diagnosed.
[b]CT: cellular therapy.
[c]FT: financial toxicity.

XSL•FO
RenderX

Figure 1. Box-plot graph comparing COST-FACIT scores across the timelines for both groups. COST-FACIT: Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy; CT: cellular therapy; FT: financial toxicity; FT 0: FT score day 0; FT 30: FT score day 30; FT 90: FT score day 90; ND: newly diagnosed.



Financial Toxicity and Sociodemographic Factors

Table 4 summarizes the difference in FT by race and ethnicity, diagnosis, and employment status at day 0, day 30, and day 90. We examined the association between FT and sociodemographic factors such as race, ethnicity, type of HM, and employment status (employment status was recategorized as employed vs unemployed for ease of analysis). Differences in FT across racial groups were marginally significant, with White patients experiencing relatively milder FT at day 0 (median FT: 27.0 for White patients, 16.5 for African American patients, and 14.0 for other racial groups; P=.08) and at day 90 (median FT: 37.0 for White patients, 13.5 for African American patients, and 15.0 for other racial groups; P=.06).

XSL•FO
RenderX

Table . Difference in financial toxicity on days 0, 30, and 90.

	Financial toxicity, median (IQR)		
Sociodemographics	Day 0	Day 30	Day 90
Total cohort			
Race			
African American	16.5 (8.8-25.2)	18.0 (10.0-25.8)	13.5 (6.0-26.0)
White	27.0 (22.0-36.0)	20.6 (17.0-30.8)	37.0 (19.0-41.0)
Other	14.0 (8.9-24.1)	16.8 (10.5-21.2)	15.0 (10.0-27.5)
P value	.08	.38	.06
Ethnicity			
Hispanic	13.5 (8.9-24.8)	15.5 (8.9-21.2)	14.5 (9.8-28.0)
Non-Hispanic	20.0 (8.5-27.5)	18.0 (14.0-26.5)	19.0 (7.8-31.5)
P value	.45	.17	.49
Diagnosis			
Leukemia	19.4 (10.3-24.5)	18.0 (17.0-22.0)	17.0 (8.8-25.8)
Lymphoma	10.0 (8.0-28.0)	14.4 (8.4-17.6)	15.0 (4.5-27.8)
Plasma cell dyscrasias	20.0 (8.8-28.0)	19.5 (9.8-28.2)	18.0 (10.0-30.0)
P value	.75	.11	.77
Employment status			
Employed	12.0 (8.0-24.0)	19.0 (17.0-22.0)	17.0 (8.0-23.0)
Unemployed	15.0 (10.0-28.0)	18.0 (10.0-28.0)	15.0 (9.0-28.0)
P value	.13	.67	.94
Newly diagnosed cohort			
Race			
African American	15.4 (8.2-20.4)	18.0 (9.0-22.5)	13.5 (6.0-22.0)
White	33.5 (26.7-36.0)	19.8 (18.4-31.9)	22.0 (16.0-40.0)
Other	14.0 (8.8-24.2)	16.5 (9.0-20.0)	19.0 (10.0-28.0)
P value	.04	.46	.16
Ethnicity			
Hispanic	14.0 (9.0-28.0)	15.0 (9.0-20.0)	21.5 (10.0-28.0)
Non-Hispanic	16.5 (7.9-25.5)	18.0 (15.5-23.8)	18.0 (9.0-24.0)
P value	>.99	.29	.38
Diagnosis			
Leukemia	17.0 (8.0-23.0)	18.0 (15.8-20.0)	19.5 (8.5-27.3)
Lymphoma	9.8 (9.0-27.5)	12.0 (8.9-17.3)	11.0 (4.0-27.0)
Plasma cell dyscrasias	22.0 (9.1-28.0)	19.4 (8.0-28.8)	19.0 (10.0-28.0)
P value	.92	.29	.48
Employment status			
Employed	11.0 (7.25-26.0)	18.0 (13.3-20.5)	18.0 (9.0-22.3)
Unemployed	15.0 (9.0-15.0)	12.0 (8.0-28.0)	12.0 (6.0-28.0)
P value	.41	.95	.71
Cellular therapy cohort			
Race			
African American	19.5 (8.8-28.3)	24.0 (14.5-29.2)	33.0 (7.0-44.0)

XSL•FO
RenderX

	Financial toxicity, median (IQR)		
Sociodemographics	Day 0	Day 30	Day 90
White	22.0 (13.8-29.5)	22.8 (10.9-29.4)	39.5 (37.0-42)
Other	19.9 (10.0-24.0)	18.0 (13.9-1.5)	13.0 (8.75-16.25)
P value	.77	.55	<.001
Ethnicity			
Hispanic	16.0 (8.9-22.5)	16.0 (11.3-21.5)	13.0 (8.75-16.3)
Non-Hispanic	24.5 (9.5-31.0)	22.5 (14.5-30.9)	37.0 (20.0-43.0)
P value	.29	.24	.05
Diagnosis			
Leukemia	20.0 (13.0-25.0)	22.0 (19.5-29.5)	13.5 (8.75-25.5)
Lymphoma	13.0 (8.0-28.0)	15.0 (10.7-17.0)	N/A
Plasma cell dyscrasias	19.9 (8.7-27.8)	20.0 (13.5-26.8)	14.0 (8.25-32.25)
P value	.94	.21	.95
Employment status			
Employed	20.0 (5.0-20.0)	22.0 (12.0-29.5)	12.0 (7.5-31.5)
Unemployed	16.5 (10.3-34.0)	20.5 (15.0-29.0)	15.5 (13.3-32.8)
P value	.25	.45	.66

Within the ND cohort, White patients had significantly less severe FT than other racial groups at day 0 (median FT: 33.5 for White patients, 15.4 for African American patients, and 14.0 for other racial groups; *P*=.04). However, the significance was not maintained at day 30 or day 90.

A similar trend was observed in the CT group at day 90 (median FT: 39.5 for White patients, 33.0 for African American patients, and 13.0 for other racial groups; *P*<.001).

Predictors of Financial Toxicity

A linear regression analysis was performed (Table 5) to identify potential predictors of FT, as measured by the continuous COST-FACIT score. The model incorporated sex, race and ethnicity, insurance status, employment, and type of malignancy as independent variables. Among individual predictors, race and ethnicity were the only statistically significant predictors, which were associated with a 3.08-point lower COST-FACIT score, indicating higher FT in this group (B=−3.08, *P*=.04, 95% CI −6.047 to −0.121). The other variables—sex, insurance status, employment, and malignancy type—were not significantly correlated with FT.

Table . Linear regression predicting financial toxicity using the COST-FACIT[a] score.

Predictor	B (95% CI)	SE	β	*t*	*P* value
Sex	3.521 (−1.136 to 8.177)	2.337	.169	1.507	.14
Race/ethnicity	−3.084 (−6.047 to −0.121)	1.487	−.232	−2.074	.04
Health insurance	0.357 (−2.697 to 3.411)	1.533	.027	0.233	.82
Type of malignancy	0.185 (−2.711 to 3.080)	1.453	.014	0.127	.90
Employment	3.724 (−2.711 to 3.080)	2.511	.173	0.127	.90
Intercept	14.317 (−1.272 to 29.906)	7.824	—	1.830	.07

[a]COST-FACIT: Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy.

Subjective Experience of Financial Toxicity

Our social workers evaluated more than 75% of the patients in our survey at the time of diagnosis. Over a third (n=32, 37%) of our patients were employed, while more than 60% were retired, disabled, or unemployed (Table 1). Table 6 and Figure 2 show the results of our investigator-based questionnaire. Most patients reported financial difficulties in carrying out their daily activities, such as paying for food, heating/air-conditioning, or warm clothes over the past 6 months (ND group: n=28, 54%; CT group: n=23, 61%). Additionally, over half of the patients experienced some degree of emotional distress in their daily

XSL•FO
RenderX

lives (ND group: n=29, 57%; CT group: n=18, 48%). About 15% in the ND group and 26% in the CT group delayed seeking medical care due to financial constraints, and only about 10% missed appointments due to caregiver issues.

XSL•FO
RenderX

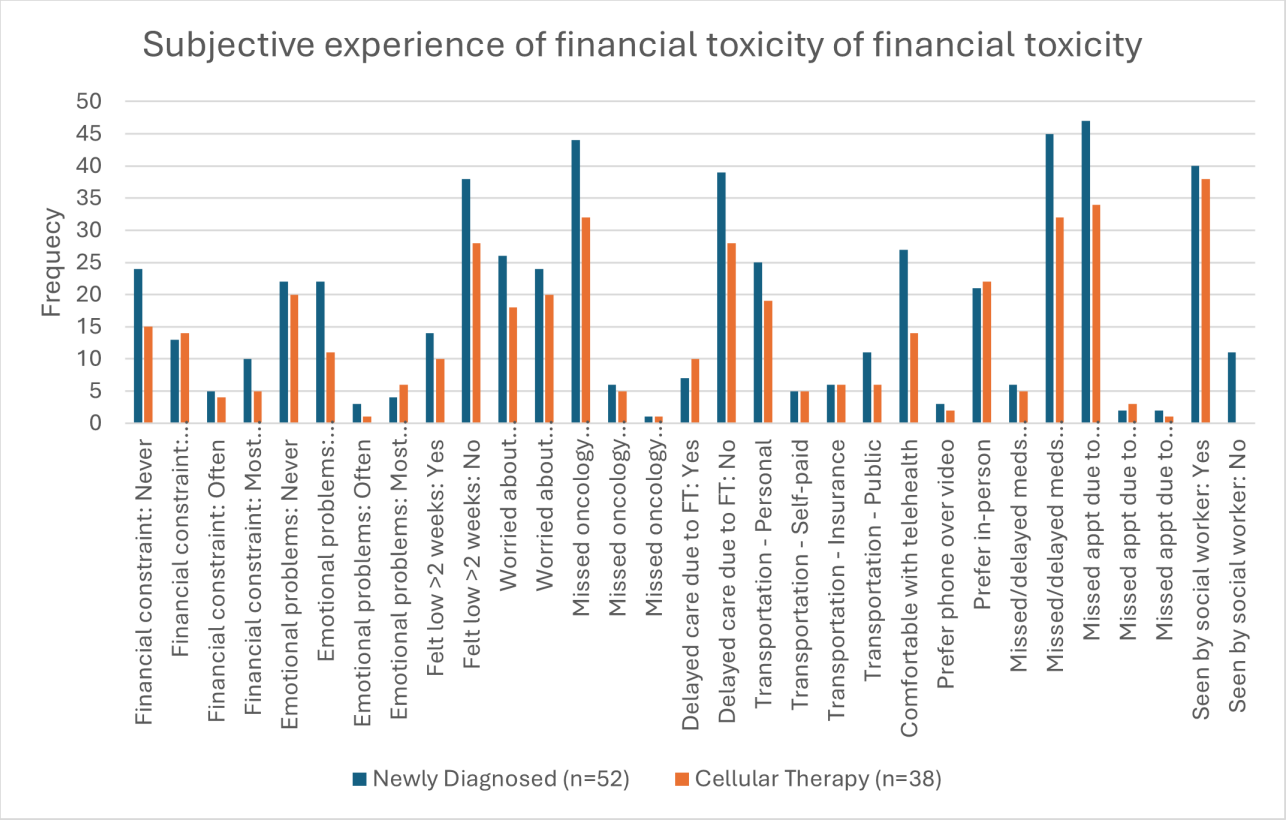
Table . Subjective experience of financial toxicity on patient's well-being and access to care.

	Day 0	
Economic data	Newly diagnosed (n=52)	Cellular therapy (n=38)
Have you felt financially constrained, such as paying for food, heating/cooling in the last 6 months? n (%)		
Never	24 (46)	15 (39)
Sometimes	13 (25)	14 (37)
Often	5 (10)	4 (11)
Most of the time	10 (19)	5 (13)
During the past 4 wk, have you had any problems with your work or daily life due to any emotional problems, such as feeling depressed, sad, or anxious? n (%)		
Never	22 (43)	20 (52)
Sometimes	22 (43)	11 (29)
Often	3 (6)	1 (3)
Most of the time	4 (8)	6 (16)
Over the last 6 months, have you felt particularly low for more than 2 wk? n (%)		
Yes	14 (27)	10 (26)
No	38 (73)	28 (74)
Are you worried about access and cost of transportation for medical appointments? n (%)		
Yes	26 (52)	18 (47)
No	24 (48)	20 (53)
Have you missed an oncology appointment due to lack of transportation? n (%)		
Never	44 (86)	32 (84)
Sometimes	6 (12)	5 (13)
Often	1 (2)	1 (3)
Have you delayed seeking medical care due to financial toxicities? n (%)		
Yes	7 (15)	10 (26)
No	39 (85)	28 (74)
Which mode of transportation do you use to come to medical appointments? n (%)		
Personal	25 (53)	19 (53)
Self-paid transportation	5 (11)	5 (14)
Insurance transportation	6 (13)	6 (17)
Public transportation	11 (23)	6 (17)
Distance from treatment clinic (miles)		
Median (range)	5.8 (0 - 131)	5.8 (0 - 217.1)
How do you feel about telephone or video visits in addition to in-person visits? n (%)		
I feel comfortable with phone/video visits	27 (53)	14 (40)
I prefer phone visits to video visits due to technological challenges	3 (6)	2 (3)

XSL•FO
RenderX

	Day 0	
Economic data	Newly diagnosed (n=52)	Cellular therapy (n=38)
I always prefer in-person visits	21 (41)	22 (57)
Have you missed doses or delayed treatments due to not being able to pay for medications? n (%)		
Yes	6 (12)	5 (14)
No	45 (88)	32 (86)
Have you missed an oncology appointment due to caregiver issues? n (%)		
Never	47 (92)	34 (89)
Sometimes	2 (4)	3 (8)
Often	2 (4)	1 (3)
Evaluated by a social worker, n (%)		
Yes	40 (78)	38 (100)
No	11 (22)	0 (0)

Figure 2. Box-plot graph showing frequencies of parameters assessing subjective experience from financial toxicity by treatment group. CT: cellular therapy; ND: newly diagnosed.



Regarding transportation, the median distance from patients' homes to the treatment center was 5.8 km in both groups. Although fewer than 20% missed appointments due to lack of transportation, many expressed concerns about the cost and accessibility of transportation to their medical appointments (ND group: 52% on day 0; CT group: 47% on day 0). Interestingly, many patients showed interest in telemedicine alongside in-person visits (ND group: 53% and CT group: 40% at day 0). We examined differences in social factors and subjective experiences of FT by ethnicity, and no statistical difference was found at day 0 in either cohort.

Discussion

Principal Findings

Our pilot study highlights a high prevalence of FT within the study population, affecting over 70% of participants. While no significant differences in FT were observed between patients

XSL•FO
RenderX

ND and those undergoing CT, those in the CT group had significantly worse FT at day 90 compared to baseline (day 0). Race and ethnicity emerged as significant predictors of FT, with African American and Hispanic patients experiencing greater FT compared to non-Hispanic White counterparts. Although most patients could attend their oncology appointments and receive treatment, over one-third expressed concerns regarding access to transportation and associated costs.

Comparison to Prior Work

Our findings are consistent with a growing body of literature that identifies FT as a persistent and complex challenge in cancer care [14]. Our patients in the CT group experienced worse FT over time, similar to a previously published study on CAR-T therapy [15]. The study by Maziarz et al [16] on health care costs in patients receiving allogeneic transplants reported that the median cost of health care for a patient receiving a transplant was over US $300,000. Apart from the high cost of SCT and CAR-T, prolonged hospital stays and multiple hospitalizations from treatment-related toxicities add to the burden of FT in these patients [17,18].

In addition to direct medical costs, indirect expenses—such as transportation, loss of employment from poor productivity due to illness, or switching to low-income jobs because they are less physically demanding—also add to the burden of FT [19,20]. Our study confirms these findings. In our study, although most patients made their oncology appointments, over a third were worried about access and the cost of transportation to attend them. Patients undergoing scheduled chemotherapy infusions need to make multiple visits to the clinic, which can be burdensome, especially for those in the CT group who must travel long distances to a FACT (Foundation for Accreditation of Cellular Therapy)-accredited center. This increases gas and public transportation costs, ultimately increasing the overall treatment cost burden [21,22]. Insurance coverage, while protective, may be insufficient to offset the rising costs. Increased deductibles, co-payments, and coinsurance have shifted much of the burden to patients, leading to substantial out-of-pocket expenses for their cancer care. Our study did not assess these out-of-pocket costs directly, nor did it quantify medical debt, areas that merit further investigation.

We found that race and ethnicity are significant predictors of FT, with African American and Hispanic patients experiencing worse FT than non-Hispanic White patients. Our results do not differ from studies on race and cancer-related FT in the literature [23-25]. Most of our patients are essentially racial and ethnic minorities and immigrants, and some are undocumented. They are likely to have lower incomes compared to their White counterparts, to be in the lowest socioeconomic tier, and have access only to public health insurance with limited coverage or even be uninsured [26-28]. Additionally, these patients are likely to have lower health literacy, leading to late disease presentation, increasing the intensity and cost of treatment. These findings underscore the structural inequities that contribute to FT among racial and ethnic minorities and immigrant populations. The convergence of social, economic, and health care–related factors amplifies financial strain and perpetuates disparities in outcomes. Addressing FT in these populations will require more than individual-level solutions; it demands systems-level change.

Beyond the direct effects of FT, it also encompasses the subjective financial distress—the emotional and psychological toll of economic strain [2,29]. This includes the depletion of household wealth and nonmedical budgets, as well as worries about the effectiveness of coping strategies [30]. In our study, over half of the patients experienced financial difficulty in paying for basic needs such as food, and more than 50% experienced emotional distress in their daily life, aligning with findings from Yu et al [31]. Although our results indicate a substantial emotional burden, further research is needed to determine whether this stems from receiving devastating news of a cancer diagnosis, the financial burden of treatment, or a combination of both. In our study, where over 50% of patients are unemployed, most of them will likely have to adjust their monthly budget and spend less on basic needs to pay for their cancer treatment. As their cancer treatment progresses, which is often prolonged in HM, they may face increasing financial burdens, leading them to adopt coping strategies such as relying on retirement savings, selling valuables, or borrowing from friends, family, or financial institutions [32]. In some cases, patients may resort to maladaptive strategies such as missing hospital appointments, medication nonadherence, or even stopping treatment entirely [33]. In our study, though a few patients missed or delayed treatments due to financial constraints, we see a decrease in the number of patients when followed up on day 30 and day 90. Although only a few patients in our study reported missing or delaying treatment due to financial constraints, we observed a decline in patient follow-up at 30 and 90 days. While the reasons for this attrition are not fully elucidated, the inability to afford ongoing treatment remains a plausible factor. These strategies may temporarily mitigate the financial impact of cancer care but ultimately lead to reduced quality of life, emotional distress, and devastating clinical outcomes.

Limitations

Our study certainly has limitations. First, we had a low sample size, especially at the 90-day time point, as many of our patients were lost to follow-up, deceased, or chose not to continue with the study. Second, although our study was observational, the follow-up period may not have been long enough to detect a significant change in FT across time. Third, we did not collect data on the participants' monthly income and the number of cycles of chemotherapy received at each survey time. This may have provided additional insights into assessing the severity of FT in our patient population. With our pilot study, we aimed to capture some of the barriers to cancer care in a unique patient population.

Potential Solutions and Future Directions

Mitigating the burden of FT will involve developing and implementing intervention strategies at multiple levels, from the state/national level to the health insurance/payer, hospital, and provider levels [34]. In our study, many patients were worried about the transportation cost to their oncology appointments but were open to telemedicine visits. Therefore, incorporating telemedicine visits as an option in the care of our

XSL•FO
RenderX

patient population could offer more flexibility for patients, reducing the travel burden and lost income from missed work [35]. Most of our patients were assessed by a social worker. While they play a role in identifying patients at high risk of financial distress, their assistance may not be sufficient, and they may not have the expertise to provide solutions to mitigate the severity of FT. A dedicated financial navigator, especially in a quaternary academic medical center, is required to help patients understand the economic aspects of their cancer care, budget appropriately, and maximize their employment and disability benefits in the context of ongoing financial commitment [36]. A larger cancer center–wide study is underway to evaluate the social determinants of health and better understand their implications on patient outcomes.

Conclusions

This study highlights the significant and far-reaching impact of FT experienced by patients with cancer, particularly those from socioeconomically disadvantaged and ethnic and racial minority backgrounds. Our findings underscore the need for early FT screening and multilevel interventions to protect vulnerable populations from economic harm during cancer care.

Acknowledgments

This manuscript is dedicated to the exceptional life and work of Karen Wright, an oncology nurse and a dedicated patient advocate for patients at Montefiore Medical Center.

Part of this study was presented at the American Society of Hematology Annual Meeting in San Diego, California, 2023 (abstract presentation and online publication) and the American Society of Clinical Oncology Annual Meeting in Chicago, 2023 (online publication).

Generative artificial intelligence (ChatGPT-5, OpenAI, 2025) was used to (1) generate Figure 2 based on the data presented in Table 6, (2) assist with the interpretation of the linear regression analysis performed using SPSS version 29, and (3) improve grammar and clarity of the text.

Data Availability

The datasets underpinning this study are securely stored on a restricted institutional server and are not publicly accessible, safeguarding participant confidentiality. Access may be granted upon formal request to the corresponding author, contingent upon ethical and legal approvals.

Authors' Contributions

Conceptualization: KF, NS, AS, KW
Data curation: LA, LL, KF, AS, DL, RM
Formal analysis: CZ, M Kim
Writing-original draft: LA, NS.
Writing -review and editing: DC, MG, IM, RAS, KG, M Konopleva, RG, NK, AV, BR, AS.

Conflicts of Interest

None declared.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin 2022 Jan;72(1):7-33. [doi: 10.3322/caac.21708] [Medline: 35020204]
2. Zafar SY, Peppercorn JM, Schrag D, et al. The financial toxicity of cancer treatment: a pilot study assessing out-of-pocket expenses and the insured cancer patient's experience. Oncologist 2013;18(4):381-390. [doi: 10.1634/theoncologist.2012-0279] [Medline: 23442307]
3. Pitis A, Diamantopoulou M, Panagiotou A, Papageorgiou D, Tzavella F. Financial toxicity and its association with the quality of life of Greek patients with cancer: a cross-sectional study. Nurs Rep 2025 Feb 13;15(2):67. [doi: 10.3390/nursrep15020067] [Medline: 39997803]
4. McLean L, Hong W, McLachlan SA. Financial toxicity in patients with cancer attending a public Australian tertiary hospital: a pilot study. Asia Pac J Clin Oncol 2021 Jun;17(3):245-252. [doi: 10.1111/ajco.13448] [Medline: 32894819]
5. Oshiro NN, Nogueira LDA, Santos YHD, Guimarães PRB, Kalinke LP. Quality of life and financial toxicity of hematopoietic stem cell transplant recipients in COVID-19. Rev Lat Am Enfermagem 2023;31:e3995. [doi: 10.1590/1518-8345.6688.3995] [Medline: 37729248]
6. Fitch K, Ferro C, Pittinger S. The cost burden of blood cancer care in medicare. : The Leukemia & Lymphoma Society; 2019 URL: https://us.milliman.com/en/insight/the-cost-burden-of-blood-cancer-care-in-medicare [accessed 2024-01-10]
7. Bestvina CM, Zullig LL, Yousuf Zafar S. The implications of out-of-pocket cost of cancer treatment in the USA: a critical appraisal of the literature. Future Oncol 2014 Nov;10(14):2189-2199. [doi: 10.2217/fon.14.130] [Medline: 25471033]

XSL•FO
RenderX

8. Dieguez G, Ferro C, Rotter D. The cost burden of blood cancer care. : The Leukemia & Lymphoma Society; 2018 URL: https://us.milliman.com/en/insight/the-cost-burden-of-blood-cancer-care [accessed 2024-01-10]
9. Ouchveridze E, Banerjee R, Desai A, et al. Financial toxicity in hematological malignancies: a systematic review. Blood Cancer J 2022 Apr 22;12(4):74. [doi: 10.1038/s41408-022-00671-z] [Medline: 35459862]
10. Gordon LG, Merollini KMD, Lowe A, Chan RJ. A systematic review of financial toxicity among cancer survivors: we can't pay the co-pay. Patient 2017 Jun;10(3):295-309. [doi: 10.1007/s40271-016-0204-x] [Medline: 27798816]
11. Lentz R, Benson AB, Kircher S. Financial toxicity in cancer care: prevalence, causes, consequences, and reduction strategies. J Surg Oncol 2019 Jul;120(1):85-92. [doi: 10.1002/jso.25374] [Medline: 30650186]
12. The Bronx neighborhood profile. NYU Furman Center. URL: https://furmancenter.org/neighborhoods/view/the-bronx [accessed 2023-09-01]
13. de Souza JA, Yap BJ, Wroblewski K, et al. Measuring financial toxicity as a clinically relevant patient-reported outcome: the validation of the Comprehensive Score for Financial Toxicity (COST). Cancer 2017 Feb 1;123(3):476-484. [doi: 10.1002/cncr.30369] [Medline: 27716900]
14. Jain T, Litzow MR. No free rides: management of toxicities of novel immunotherapies in ALL, including financial. Blood Adv 2018 Nov 27;2(22):3393-3403. [doi: 10.1182/bloodadvances.2018020198] [Medline: 30482769]
15. Cusatis R, Tan I, Piehowski C, et al. Worsening financial toxicity among patients receiving chimeric antigen receptor t-cell (CAR-T) therapy: a mixed methods longitudinal study. Blood 2021 Nov 5;138(Supplement 1):567-567. [doi: 10.1182/blood-2021-146032]
16. Maziarz RT, Gergis U, Edwards ML, et al. Health care costs among patients with hematologic malignancies receiving allogeneic transplants: a US payer perspective. Blood Adv 2024 Mar 12;8(5):1200-1208. [doi: 10.1182/bloodadvances.2023011033] [Medline: 38055922]
17. McDiarmid S, Hutton B, Atkins H, et al. Performing allogeneic and autologous hematopoietic SCT in the outpatient setting: effects on infectious complications and early transplant outcomes. Bone Marrow Transplant 2010 Jul;45(7):1220-1226. [doi: 10.1038/bmt.2009.330] [Medline: 19946343]
18. Ringdén O, Remberger M, Holmberg K, et al. Many days at home during neutropenia after allogeneic hematopoietic stem cell transplantation correlates with low incidence of acute graft-versus-host disease. Biol Blood Marrow Transplant 2013 Feb;19(2):314-320. [doi: 10.1016/j.bbmt.2012.10.011] [Medline: 23089563]
19. Abrams HR, Durbin S, Huang CX, et al. Financial toxicity in cancer care: origins, impact, and solutions. Transl Behav Med 2021 Nov 30;11(11):2043-2054. [doi: 10.1093/tbm/ibab091] [Medline: 34850932]
20. Ehsan AN, Wu CA, Minasian A, et al. Financial toxicity among patients with breast cancer worldwide. JAMA Netw Open 2023 Feb 1;6(2):e2255388. [doi: 10.1001/jamanetworkopen.2022.55388]
21. Ambroggi M, Biasini C, Del Giovane C, Fornari F, Cavanna L. Distance as a barrier to cancer diagnosis and treatment: review of the literature. Oncologist 2015 Dec;20(12):1378-1385. [doi: 10.1634/theoncologist.2015-0110] [Medline: 26512045]
22. Wercholuk AN, Parikh AA, Snyder RA. The road less traveled: transportation barriers to cancer care delivery in the rural patient population. JCO Oncol Pract 2022 Sep;18(9):652-662. [doi: 10.1200/OP.22.00122] [Medline: 35834768]
23. Knight TG, Deal AM, Dusetzina SB, et al. Financial toxicity in adults with cancer: adverse outcomes and noncompliance. J Oncol Pract 2018 Oct 24:JOP1800120. [doi: 10.1200/JOP.18.00120] [Medline: 30355027]
24. Panzone J, Welch C, Morgans A, et al. Association of race with cancer-related financial toxicity. JCO Oncol Pract 2022 Feb;18(2):e271-e283. [doi: 10.1200/OP.21.00440] [Medline: 34752150]
25. Yabroff KR, Dowling EC, Guy GP Jr, et al. Financial hardship associated with cancer in the United States: findings from a population-based sample of adult cancer survivors. J Clin Oncol 2016 Jan 20;34(3):259-267. [doi: 10.1200/JCO.2015.62.0468] [Medline: 26644532]
26. Pisu M, Kenzik KM, Oster RA, et al. Economic hardship of minority and non-minority cancer survivors 1 year after diagnosis: another long-term effect of cancer? Cancer 2015 Apr 15;121(8):1257-1264. [doi: 10.1002/cncr.29206] [Medline: 25564986]
27. Bronx borough health equity report 2021. County B. 2021. URL: https://www.cdc.gov/minorityhealth/index.html [accessed 2025-11-20]
28. Frieden TR. Health care access among adults in New York city. The importance of having insurance and a regular care provider. : New York City Department of Health and Mental Hygiene; 2007 URL: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_resources_2007_new_report_examines_new_yorkers_access_to_care_access_dohmh_for_web_pdf.pdf [accessed 2024-01-10]
29. Zafar SY, Abernethy AP. Financial toxicity, Part I: a new name for a growing problem. Oncology (Williston Park) 2013 Feb;27(2):80-81. [Medline: 23530397]
30. Davis MP. Cancer pain. In: The MASCC Textbook of Cancer Supportive Care and Survivorship: Springer International Publishing; 2018:17-35. [doi: 10.1007/978-3-319-90990-5_2]
31. Yu H, Li H, Zuo T, et al. Financial toxicity and psychological distress in adults with cancer: a treatment-based analysis. Asia Pac J Oncol Nurs 2022 Sep;9(9):100069. [doi: 10.1016/j.apjon.2022.04.008] [Medline: 35669288]

XSL•FO
RenderX

32. Carrera PM, Kantarjian HM, Blinder VS. The financial burden and distress of patients with cancer: understanding and stepping-up action on the financial toxicity of cancer treatment. CA Cancer J Clin 2018 Mar;68(2):153-165. [doi: 10.3322/caac.21443] [Medline: 29338071]
33. Altice CK, Banegas MP, Tucker-Seeley RD, Yabroff KR. Financial hardships experienced by cancer survivors: a systematic review. J Natl Cancer Inst 2017 Feb;109(2):djw205. [doi: 10.1093/jnci/djw205] [Medline: 27754926]
34. Yabroff KR, Bradley C, Shih YCT. Understanding financial hardship among cancer survivors in the United States: strategies for prevention and mitigation. J Clin Oncol 2020 Feb 1;38(4):292-301. [doi: 10.1200/JCO.19.01564] [Medline: 31804869]
35. Patel KB, Turner K, Alishahi Tabriz A, et al. Estimated indirect cost savings of using telehealth among nonelderly patients with cancer. JAMA Netw Open 2023 Jan 3;6(1):e2250211. [doi: 10.1001/jamanetworkopen.2022.50211] [Medline: 36626174]
36. Offodile AC, Gallagher K, Angove R, Tucker-Seeley RD, Balch A, Shankaran V. Financial navigation in cancer care delivery: state of the evidence, opportunities for research, and future directions. J Clin Oncol 2022 Jul 20;40(21):2291-2294. [doi: 10.1200/JCO.21.02184]

Abbreviations

CAR-T: chimeric antigen receptor T-cell therapy
COST-FACIT: Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy
CT: cellular therapy
FACT: Foundation for Accreditation of Cellular Therapy
FT: financial toxicity
HM: hematologic malignancies
ND: newly diagnosed
SCT: stem cell transplantation

Edited by N Cahill; submitted 28.Oct.2024; peer-reviewed by L Wreyford, S Vengadassalapathy; revised version received 31.Aug.2025; accepted 29.Sep.2025; published 06.Jan.2026.

Please cite as:
Amaeshi L, Laufer L, Fedorov K, Sisto A, Levitz D, Mathew R, Wright K, Cooper D, Goldfinger M, Mantzaris I, Sica RA, Gritsman K, Konopleva M, Gupta R, Kornblum N, Zhang C, Kim M, Verma A, Rapkin B, Shastri A, Shah N
Financial Toxicity of Hematologic Malignancy Therapies, Including Cellular Therapy and Its Impact on Access to Care: Prospective Pilot Study
JMIR Cancer 2026;12:e68101
URL: https://cancer.jmir.org/2026/1/e68101
doi:10.2196/68101

© Lemchukwu Amaeshi, Lauren Laufer, Kateryna Fedorov, Alex Sisto, David Levitz, Rosmi Mathew, Karen Wright, Dennis Cooper, Mendel Goldfinger, Ioannis Mantzaris, R Alejandro Sica, Kira Gritsman, Marina Konopleva, Ridhi Gupta, Noah Kornblum, Chenxin Zhang, Mimi Kim, Amit Verma, Bruce Rapkin, Aditi Shastri, Nishi Shah. Originally published in JMIR Cancer (https://cancer.jmir.org), 6.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient and Clinician Perspectives on Expanding Telehealth Use for Older Adults Across the Cancer Control Continuum: Mixed Methods Study

Robin T Higashi[1,2], PhD; Bella Etingen[3,4], PhD; Jessica Lee[5], MPH; Suzanne Cole[2,6], MD; John C Mansour[2,7], MD; Alice Zhao[5], MPH; Timothy P Hogan[4,8], PhD

[1]Department of Social and Behavioral Sciences, O'Donnell School of Public Health, The University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX, United States

[2]Harold C. Simmons Comprehensive Cancer Center, Dallas, TX, United States

[3]Research and Development Service, Dallas VA Medical Center, Dallas, TX, United States

[4]Department of Health Economics, Systems, and Policy, O'Donnell School of Public Health, The University of Texas Southwestern Medical Center, Dallas, TX, United States

[5]School of Medicine, The University of Texas Southwestern Medical Center, Dallas, TX, United States

[6]Division of Hematology and Oncology, Department of Internal Medicine, The University of Texas Southwestern Medical Center, Dallas, TX, United States

[7]Department of Surgery, The University of Texas Southwestern Medical Center, Dallas, TX, United States

[8]Center for Health Optimization and Implementation Research (CHOIR), VA Bedford Healthcare System, Bedford, MA, United States

Corresponding Author:
Robin T Higashi, PhD
Department of Social and Behavioral Sciences, O'Donnell School of Public Health, The University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX, United States

Abstract

Background: Reliance on telehealth increased dramatically during the COVID-19 pandemic, introducing new opportunities to consider the use of telehealth across the cancer control continuum. However, patient, clinician, and staff perspectives about the types of cancer care appointments that are considered appropriate and the clinical care needs to support expanded remote care services are limited. Understanding older adults' diverse technology needs and perspectives is especially important given that they comprise a large and growing proportion of patients with cancer.

Objective: This study aimed to describe the perceptions and experiences of older patients with cancer and their clinical care team members regarding the expansion of telehealth use across the cancer control continuum and to solicit suggestions about how to support telehealth use for cancer care delivery.

Methods: Using a convergent mixed methods design, we surveyed and interviewed patients aged ≥60 years, clinicians, and staff at a comprehensive cancer center in the southern United States between December 2020 and November 2021. Interview questions were rooted in the sociotechnical model, which proposes 8 interrelated dimensions representing factors influencing the design, use, and outcomes associated with health information technologies. Patient survey domains included telehealth experience and satisfaction and factors affecting telehealth perceptions and use; clinician survey domains included contexts of telehealth appropriateness, training, and barriers and facilitators to telehealth service provision. Survey data were analyzed using descriptive statistics. Qualitative data were thematically analyzed using a combined deductive and inductive approach.

Results: We received completed surveys from 128 patients (567 invited) and 106 clinicians and staff (146 invited). We completed 14 patient (29 invited) and 20 clinician and staff (22 invited) interviews. Across all participants, most agreed or strongly agreed that multiple cancer care appointment types should be offered via telehealth, including discussing treatment side effects (75/102, 73.5% of patients and 66/94, 70.2% of clinicians and staff), results communication (71/102, 69.6% of patients and 65/94, 69.1% of clinicians and staff), and treatment follow-up (67/102, 65.7% of patients and 52/93, 55.9% of clinicians and staff). In interviews, participants elaborated on factors influencing the appropriateness of telehealth versus in-person appointments, including symptom severity, type of cancer, and purpose of the appointment. Many patient and staff suggestions focused on ways to address digital literacy gaps, while clinicians recommended improving clinic workflows, infrastructure, and training.

Conclusions: Overall, clinicians, staff, and older patients with cancer all responded positively toward expanding telehealth use across multiple cancer and appointment types across the cancer control continuum. Older adults with cancer are generally interested in telehealth for cancer care, especially if strategies to address digital literacy gaps are incorporated. Clinicians and staff members expressed specialized training and infrastructure needs to optimize telehealth uptake and service delivery.

XSL•FO
RenderX

(JMIR Cancer 2026;12:e73058) doi:10.2196/73058

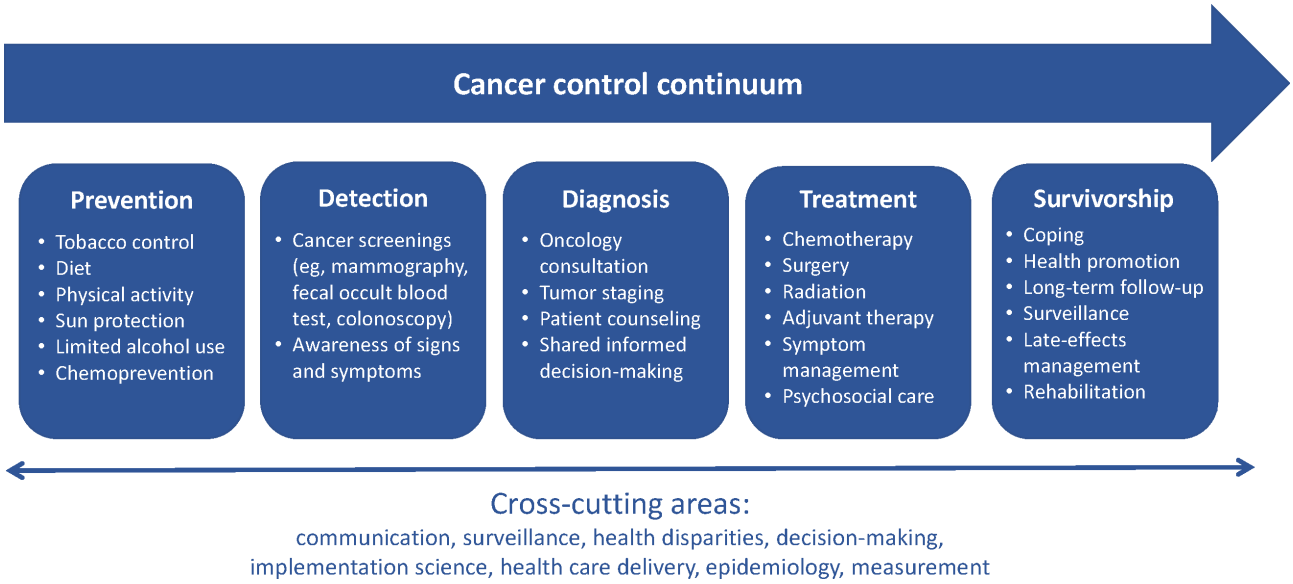
KEYWORDS

telemedicine; video visits; remote technology; older adults; remote cancer care; cancer prevention; cancer treatment

Introduction

Synchronous video visits, or telehealth, have transformed the way cancer care is delivered, accelerated in large part due to the COVID-19 pandemic [1-4]. Avoiding infection due to COVID-19 was particularly important for patients with cancer, who were twice as likely to contract the SARS-CoV-2 virus and 8 times as likely to die from it compared to individuals without cancer [5,6]. In the pre-COVID era, telehealth services for patients with cancer primarily focused on the survivorship phase of the cancer control continuum [7-10] (see Figure 1), addressing issues like long-term monitoring and rehabilitation [11,12]. Among cancer survivors, telehealth had been shown to increase independence, convenience, quality of life, and patient satisfaction with care [13-16]. However, patient and clinician attitudes toward broader implementation of telehealth across other phases of the cancer control continuum, detection, diagnosis, and treatment, for example [17,18] are limited.

Figure 1. Cancer control continuum, adapted from the National Cancer Institute [17].



During the COVID pandemic, health care systems rapidly developed and deployed new remote care protocols and processes for patients across the cancer control continuum, with little time to assess patient and clinician needs and preferences. Cancer centers reported between 33% and 50% of patient visits occurring via telehealth, compared to rates formerly in the single digits [19]. Given that approximately 65% of cancer cases occur in patients aged 65 years and older [20,21], as well as the potential for this age demographic to face disparities in technology use, we had anticipated the need to examine the impact of the rapid expansion of telehealth on older patients specifically. Before the pandemic, a growing body of literature had begun to assess digital health literacy and technology preferences among older adults. Attitudes were diverse and multiple factors affected inequities in health literacy in addition to age, including socioeconomic status, level of education, and rural residence, particularly in the delivery of health care services [22-27]. The transition to telehealth use beyond the survivorship phase created an opportunity to understand shareholder perspectives on both the types of appointments and clinical needs that would be appropriate for remote care, and what could be done to bridge the digital health disparities facing older adults.

The objective of this study was to explore patients' and clinicians' pandemic-era experiences and perceptions of the appropriateness of telehealth expansion for different encounter types across the cancer control continuum. We also solicited suggestions for enhancing access to telehealth in the current era of increasing remote cancer care delivery.

Methods

Study Design

Using a convergent mixed methods design [28], we conducted surveys and semistructured interviews with patients, clinicians, and staff at a National Cancer Institute-designated comprehensive cancer center ("Cancer Center") located at a large academic medical center in the southern United States. We conducted a small number of interviews as a preliminary step to inform the question items and domains included on subsequent survey instruments and to capture participant perceptions and experiences in more depth than could be captured through surveys alone. Concurrently, we used surveys to measure the robustness of experiences and perceptions with a larger sample than could be accomplished in interviews.

XSL•FO
RenderX

Instrument Design

We selected survey domains by drawing on the literature and the sociotechnical model [29], which proposes 8 interrelated dimensions that represent the range of factors that can influence the design, use, and outcomes associated with health information technologies. The patient survey addressed a range of domains relevant to perceptions of and experiences with telehealth, including access and use, challenges encountered, perceived usefulness, and satisfaction. The clinician and staff survey similarly addressed a range of domains, including extent of experience with telehealth, perceived appropriateness of telehealth in different clinical contexts, training, and barriers and facilitators to use. In all, the patient survey comprised 89 unique questions and the clinician staff survey 124 unique questions, including Likert scale ratings, categorical items, and open-ended responses. We pilot tested the survey and interview guides with clients, caregivers of patients with cancer, and staff members of a local community-based organization for seniors in 4 focus groups prior to launching data collection. Their insights helped us optimize the content, wording, and length of instruments.

In addition to exploring some of the survey domains in greater depth, interview guides for clinicians and staff reflected their unique roles (eg, patient care, administrative leadership, and front-line clinical and administrative staff). Additionally, interviews with patients, clinicians, and staff included questions soliciting suggestions about how to improve telehealth services and the role of telehealth in cancer care in the future.

Sampling

Surveys

We sampled patients for surveys from the Adult Cancer Patient Registry at the Cancer Center (n=4037) (Multimedia Appendix 1). Eligible patients were ≥60 years of age, had completed at least one telehealth encounter since March 15, 2020 (approximately the start of the COVID-19 pandemic), were English- or Spanish-speaking, were open to contact for research through the portal, and were diagnosed and received treatment for breast, prostate, colorectal, and/or hematological cancer (in order to achieve a diverse sample by cancer type; n=2464). Of these individuals, we identified n=580 patients for survey recruitment, first stratified by cancer organ site in proportion with the Adult Cancer Patient Registry, then randomly selected within each organ category by date of telehealth appointment, age, gender, and race and ethnicity to achieve a diverse sample. We invited all employees across the Cancer Center to participate in the electronic survey.

Interviews

Patients recruited for semistructured interviews followed the same eligibility criteria as those for the survey but additionally were identified as being in active cancer treatment (defined as receiving chemotherapy, radiation, or being within 6 months postsurgery) at the time of recruitment. We also purposefully recruited 3 patients for interviews who met all criteria except that they had not completed a telehealth encounter in an effort to understand possible reasons for nonengagement or barriers to telehealth.

The primary investigators (RTH and TPH), who are nonphysicians with expertise in qualitative methods, worked with Cancer Center leaders and clinical co-investigators (SC and JM) to create a purposive sample of clinicians and staff given the diversity of cancer care services available. We jointly identified individuals with expertise across various organ sites (breast, colorectal, prostate, and heme), roles (medical oncologists, surgical oncologists, radiation oncologists, clinical staff, and administrative staff), and levels of telehealth proficiency.

Recruitment

Surveys

Because our objective was to survey patients who had completed telehealth appointments, and because telehealth appointments require the use of a patient portal account at the Cancer Center where we conducted our work, we first contacted patients through a messaging function within the patient portal with an invitation to complete the survey. Patients who did not respond to the portal message within one week were sent the same message to their personal email on file in the electronic health record. Those who did not respond to the first email inviting them to complete the survey, or who did not decline to participate, were sent a reminder email one week later, and those who did not respond to both emails within 2 weeks were sent an opt-out invitation letter by postal mail before they were called twice and offered the option to complete the survey by telephone, online through REDCap (Research Electronic Data Capture; Vanderbilt University; RRID:SCR_003445), or on paper via postal mail.

We recruited Cancer Center clinicians and staff via email for both surveys and interviews. The survey invitation included a link to take the survey in REDCap [30]. Clinicians and staff invited to participate in interviews were provided the option to schedule the interview via video conferencing or telephone. Two additional emails were sent if a survey response was not logged or if no reply was received to the interview invitation.

Interviews

Patients who completed a survey and who indicated in the survey a willingness to be contacted for interview participation were in turn recruited for interviews. We conducted semistructured telephone interviews with 14 patients, 11 of whom completed at least one telehealth appointment since March 15, 2020; we also purposefully recruited 3 patients who had never completed a telehealth appointment to provide contrasting perspectives. Patients who had completed telehealth appointments were recruited the same way as survey participants. The 3 patients who did not complete a telehealth appointment were mailed an opt-out letter, then contacted by phone. Up to two attempts were made to contact each of these participants.

Data Collection

Surveys

We conducted patient surveys between July and October 2021 and clinician and staff surveys between April and June 2021. Patient surveys took approximately 20 minutes to complete,

XSL•FO
RenderX

while clinician and staff surveys took approximately 10 minutes to complete.

Interviews

We conducted patient interviews between August and November 2021 and clinician and staff interviews between December 2020 and September 2021. Patient interviews lasted 24 - 78 minutes (average 51 minutes) and were conducted via telephone in English or Spanish by bilingual, qualitatively trained nonphysician research staff (RTH and Ana Belen Conrado). Clinician and staff interviews lasted 29 - 54 minutes (average 41 minutes) and were conducted via telephone or videoconferencing per individual preference by nonphysician investigators (RTH and TPH). Interviewers recorded brief field notes following each interview, which were reviewed and discussed in weekly meetings with members of the research team to determine when thematic saturation had been reached and thus data collection ended.

Data Analysis

Survey data were analyzed using descriptive statistics. Audio-recorded interviews were transcribed by a professional vendor and data were managed using NVivo 12.0 (QSR International; RRID:SCR_014802). Interviews were analyzed using a thematic content analysis approach. Two qualitatively trained researchers (RTH) first developed an initial, deductively driven codebook corresponding to the semistructured interview guide domains and sociotechnical model. They jointly coded the first 20% of transcripts, sampling from a diverse set of clinician and staff roles to refine codebook definitions and add and define codes as needed based on emergent findings. The codebook was then finalized, and researchers independently coded the remaining 80% of transcripts, meeting weekly to resolve discrepancies. The lead qualitative analyst (RTH) then reviewed thematic code reports and selected representative quotes and findings for discussion and interpretation with the research team. In line with our mixed methods design, to assess areas of convergence between the survey and interview findings, the team used a joint display approach [31] to systematically integrate findings from the interviews corresponding to survey questions that we had chosen to explore in greater depth (eg, preferences for appointment type, perceptions of the telehealth implementation process, and training or preparedness for telehealth use).

We engaged the Cancer Center leaders and clinical co-investigators in 2 small group meetings to review the data and discuss their interpretations of the findings.

Ethical Considerations

This study was approved by the Institutional Review Board at the University of Texas Southwestern Medical Center. All participants provided verbal informed consent to participate in interviews in accordance with the approved protocol (STU 2020 - 0919). Participants who completed the survey received a $10 gift card. All patients who completed an interview received a $20 gift card. Clinicians and staff did not receive remuneration per institutional policy. Consent included recognition that identifiable information about participants would not be included in any publications or reports resulting from the research.

Results

Overview

Of 567 patients and 146 clinicians and staff invited to participate in the survey, 128 patients (23% recruitment rate) and 106 clinicians and staff (73% recruitment rate) completed the survey. For interviews, we contacted 29 patients who completed the survey and had indicated they were willing to be contacted for interviews; 3 people were unavailable after the maximum 3 attempted calls, 12 declined to participate ("not interested after all"), and 14 patients consented and were interviewed. We invited 22 clinicians and staff to participate in interviews; 2 did not respond, and 20 consented and participated in an interview. Demographic characteristics of survey and interview participants are shown in Table 1.

The qualitative and quantitative findings converged across multiple domains. In this analysis, we focus on two principal thematic findings: (1) agreement in patients' and clinicians' openness toward expanding cancer care via telehealth beyond the survivorship phase of the cancer control continuum, and (2) comparison between patients' and clinicians' recommendations about how to enhance telehealth in the future. Representative quotes from semistructured interviews are documented by theme in Multimedia Appendix 2.

XSL•FO
RenderX

Table . Demographics of patient and clinician and staff survey and interview participants.

Total	Patient surveys (n=131), n (%)	Clinician and staff surveys (n=106), n (%)	Patient interviews (n=14), n (%)	Clinician and staff interviews (n=20), n (%)
Gender				
Women	55 (42)	57 (54)	7 (50)	10 (50)
Men	61 (47)	14 (13)	7 (50)	10 (50)
No response	15 (11)	35 (33)	0 (0)	0 (0)
Age (years)				
60 - 69	55 (42)	__[a]	__[b]	__[a]
70 - 79	50 (38)	__[a]	__[b]	__[a]
≥80	12 (9)	__[a]	__[b]	__[a]
≤35	__[a]	23 (22)	__[b]	2 (10)
36 - 45	__[a]	25 (24)	__[b]	7 (35)
46 - 55	__[a]	11 (10)	__[b]	4 (20)
56 - 65	__[a]	11 (10)	__[b]	3 (15)
≥66	__[a]	2 (2)	__[b]	3 (15)
No response	14 (11)	34 (32)	__[b]	1 (5)
Race or ethnicity				
Asian	3 (2)	8 (8)	0 (0)	4 (20)
Black	10 (8)	4 (4)	2 (14)	1 (5)
Hispanic	9 (7)	11 (10)	0 (0)	2 (10)
White	93 (71)	43 (41)	12 (86)	12 (60)
Two or more race or ethnicity	2 (2)	3 (3)	0 (0)	1 (5)
Other	0 (0)	1 (1)	0 (0)	0 (0)
No response	14 (11)	36 (34)	0 (0)	0 (0)
Language				
English	117 (89)	106 (100)	14 (100)	20 (100)
Spanish	1 (1)	0 (0)	0 (0)	0 (0)
No response	13 (10)	0 (0)	0 (0)	0 (0)
Driving time from home to Cancer Center				
<15 minutes	4 (3)	__[a]	__[a]	__[a]
15 - 30 minutes	38 (29)	__[a]	__[a]	__[a]
31 - 59 minutes	49 (37)	__[a]	__[a]	__[a]
1 - 2 hours	18 (14)	__[a]	__[a]	__[a]
>2 hours	8 (6)	__[a]	__[a]	__[a]
No response	14 (11)	__[a]	__[a]	__[a]
Role				
Administrative managers	__[a]	3 (3)	__[a]	3 (15)
Administrative staff (eg, Medical Office Assistant)	__[a]	8 (8)	__[a]	2 (10)

XSL•FO
RenderX

Total	Patient surveys (n=131), n (%)	Clinician and staff surveys (n=106), n (%)	Patient interviews (n=14), n (%)	Clinician and staff interviews (n=20), n (%)
Clinical staff (eg, Clinical Staff Assistant, Dietician)	—[a]	6 (6)	—[a]	1 (5)
Clinicians (eg, MD, NP, PA, Psychologist)	—[a]	26 (25)	—[a]	13 (65)
Nurse (RN, LVN)	—[a]	32 (30)	—[a]	1 (5)
No response	—[a]	31 (29)	—[a]	0 (0)

[a]Not applicable.

[b]Not available.

Theme #1: Perceptions Toward Expanding Telehealth to Other Phases of the Cancer Control Continuum

Patients

Both patient and clinician and staff participants responded positively toward questions about the future use of telehealth during other phases of the cancer control continuum. A majority of surveyed patients indicated they would be "willing to have a video visit with a member of their cancer care team" for discussing treatment side effects (75/102, 73.5%), results communication (71/102, 69.6%), and treatment follow-up (67/102, 65.7%), as shown in Table 2. Of note, less than 6% (6/102) of our sample indicated that telehealth should not be used for any types of cancer care appointments.

Table . Patient perceptions regarding appointment types for which telehealth should be offered (n=102).

Appointment type (check all that apply)	Value, n (%)
Discussing treatment side effects	75 (73.5)
Results communication	71 (69.6)
Treatment follow-up	67 (65.7)
New treatment planning	48 (47.1)
Routine check-up	32 (31.4)
Planning the end of cancer treatment	32 (31.4)
Postoperative surgical follow-up	28 (27.5)
Genetic counseling	26 (25.5)
Physical evaluation	13 (12.7)
None	6 (5.9)
Other appointments	3 (2.9)
Total	102 (100)

In semistructured interviews, patients cited various reasons why they would feel comfortable having certain types of appointments via telehealth versus in-person, including how they were feeling physically, whether the appointment was routine, whether they had extensive questions, or whether the appointment was with a new clinician. Patients also mentioned that, regardless of appointment type, telehealth was often a convenient option for purely logistical reasons. Some had small children, others had mobility impairments, and still others noted that an easy public transportation system was unavailable to them.

Clinicians and Staff

Clinician and staff survey findings mostly mirrored patients', but there was recognition of the potential role for telehealth in an even greater range of appointment types. A majority of clinicians and staff felt telehealth should "definitely" be offered for discussing treatment side effects (66/94, 70.2%), results communication (65/94, 69.1%), genetic counseling (63/94, 66.7%), routine check-ups (61/94, 64.9%), and treatment follow-up (52/93, 55.9%), as shown in Figure 2.

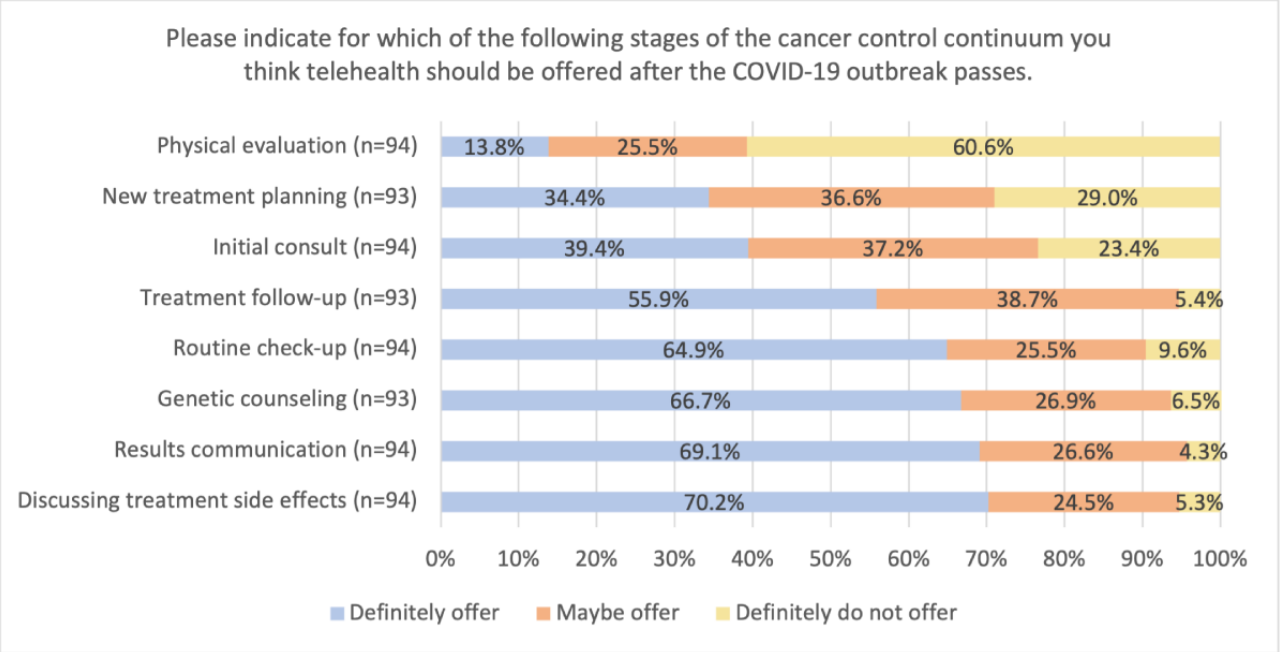
In semistructured interviews, clinicians elaborated on the multiple considerations that would determine whether they would offer a telehealth versus in-person appointment. One clinician felt that appointments scheduled to discuss laboratory results were an appropriate opportunity to use telehealth because the primary purpose of such appointments is to support conversation as opposed to gathering more laboratory data or providing other treatment. Similarly, another clinician felt that patients on infusion could be followed adequately via telehealth appointments, provided they are in a stable place with their treatment.

XSL•FO
RenderX

Some clinicians indicated that the appropriateness of telehealth depended on the type of cancer the patient has. A urology oncologist commented that telehealth was appropriate for his patients with prostate cancer, whose treatment comprised only hormone shots every 6 months. A hematologist noted there were varying circumstances in which patients with leukemia required in-person visits; lymph node checks required a physical exam, whereas discussing lab results did not.

Other factors included the complexity of a patient's care, with one clinician noting that the immediate perioperative period for surgical patients could influence the appropriateness of the modality of care delivery. Further, the stage of a patient's disease was also an important consideration in the eyes of some clinicians, with advanced-stage disease requiring more aggressive monitoring and treatment. Not surprisingly, a majority of clinicians felt physical exams still required in-person visits. One clinician, for example, noted that an in-person exam would be crucial when determining what type of breast surgery was most appropriate for a patient with breast cancer because the breast size, shape, tumor size, and tumor location all factored into her surgical strategy.

Figure 2. Clinician and staff perceptions regarding appointment types for which telehealth should be offered.



Theme #2: Suggestions to Enhance Remote Cancer Care in the Future

Patients

The Cancer Center dedicated significant resources to patient education and technical assistance during the COVID-19 pandemic to facilitate rapid conversion of patient appointments to telehealth. Surveyed approximately one year after the height of the pandemic, fewer than 20% of patients reported experiencing specific challenges with initiating or preparing for telehealth appointments. Specifically, 18% did not understand the video visit setup process, 19% did not feel it was easy to join the video visit session, and 17% did not have enough technical support. But while the vast majority (88%) of the surveyed patients felt sufficiently prepared for their telehealth appointment, less than a third (31%) of surveyed clinicians and staff agreed or strongly agreed that the Cancer Center did a good job of preparing patients for their first telehealth encounter.

Importantly, in semistructured interviews, patients still offered various suggestions about how telehealth care could be enhanced in the future. One participant recommended improving the quality of educational materials pertaining to telehealth, noting they often assume patients have better digital literacy than they actually do. Another patient felt that in-person education for older patients, particularly when initiating or preparing to use telehealth, would be helpful. One individual, who was initially opposed to having telehealth appointments before the pandemic but became a strong supporter of the technology after completing telehealth appointments during the pandemic, suggested offering testimonials to convince patients who may be apprehensive about trying telehealth.

Other patients indicated that telehealth appointments needed to be offered to patients on a more regular basis, acknowledging that they sometimes ask if telehealth is possible when scheduling.

Clinicians and Staff

Administrative leaders, clinicians, and front-line staff offered various suggestions based on the different challenges they faced and the telehealth workarounds they devised. During the early part of the pandemic, clinical and administrative staff were assigned to assist patients with initiating their telehealth appointments. This process involved first logging on through a patient's portal account, and during this step, some patients discovered problems with their portal login credentials. In such cases, knowing what level of familiarity patients already had with the portal was key to triaging a high volume of technical assistance requests. Other staff suggestions to assist patients with limited digital literacy included establishing "fake

XSL•FO
RenderX

appointments" for patients to practice online connections and hiring a dedicated employee like a digital health navigator to provide on-call technical assistance and setup of the online portal.

Whereas front-line staff suggestions tended to focus on digital literacy assistance for patients, clinician, and administrator suggestions tended to focus on the need to enhance workflows and infrastructure to improve privacy, efficiency, and clinician training. Multiple clinicians expressed the need for dedicated physical spaces with sufficient privacy for them to conduct telehealth appointments. From a technical standpoint, some clinicians noted that it would also greatly improve efficiency to implement consenting for surgical procedures and clinical trials via telehealth. Some clinicians felt that additional training was needed, for example, to show patients images over telehealth using the "share screen" function.

Despite the challenges they experienced, a majority of both patients (65.7%) and clinician and staff (76.9%) survey respondents reported that they intended to continue using telehealth in the postpandemic era.

Discussion

Principal Findings

In this mixed methods study, patients, clinicians, and staff reported an openness toward expanding telehealth-based cancer care across the cancer control continuum. Participants recognized that some situations still favored in-person visits, including appointments in which a physical exam would be necessary and appointments for patients who required more aggressive surveillance due to the stage of their disease. However, both patients and clinicians agreed that telehealth would be convenient and efficient for appointments to discuss treatment side effects, lab results, and treatment follow-up. Clinicians also noted that genetic counseling and consenting patients for surgery or clinical trials could be effectively performed over telehealth. Participants also offered multiple recommendations for health care systems to establish the infrastructure to support these specific opportunities for telehealth-based cancer care delivery.

Comparison to Prior Work

Despite the reported willingness of older patients to engage in telehealth, literature indicates that older adults are often offered telehealth appointments and patient portal registration at lower rates than younger patients, partly due to health care clinicians' assumptions that older patients may be less interested in or capable of using digital health tools [32-34]. Such assumptions can inadvertently limit older patients' access to telehealth, even though many are both willing and able to engage with these technologies when provided adequate support. Such ageist biases highlight the need for intentional strategies to encourage telehealth use among older adults, ensuring that they are not inadvertently excluded from the potential benefits of the technology. Further, given that approximately 65% of cancer cases occur among patients aged 65 years and older [20,21], addressing such inequities among older patients could expand the role of telehealth in cancer care.

Several strategies to improve digital engagement for older adults may serve to address inequities in telehealth service delivery and outcomes for older adults [22-24]. These include technical modifications in portal interfaces to simplify the sign-in process, changes to size or color of text or icons, and incorporation of more pictures as opposed to text, which may improve user experience and capability [35,36]. Other recommendations include patient portal adoption campaigns targeting older adults, task-specific training, support for proxy users, challenging assumptions about older adults and technology, and alternative workflows to allow for communication between patients and personnel to review personal health information within patient portals [22,37]. Clinical and administrative staff may benefit from having policies and procedural guidance detailing when and for whom telehealth appointments should be offered.

Our findings contribute evidence that educational materials must be tailored to older adults or other patients with lower digital health literacy, such as through in-person educational sessions and testimonials from those who have benefited from telehealth. Clinicians and staff on our study also suggested the idea of establishing "fake appointments" for patients to practice using telehealth interfaces and hiring a dedicated employee such as a digital health navigator for technical assistance. The use of digital navigators—trained individuals who assist patients in accessing and using digital health resources—has been found to be effective in enhancing digital literacy and reducing barriers to telehealth access, especially for socioeconomically disadvantaged populations [38-41]. These programs could be particularly beneficial for older adults with cancer, who may face additional challenges due to their complex care needs.

Increasing the use of telehealth for older adults with cancer has the potential to increase independence, patient satisfaction, and quality of life [13-16,42-44]. Studies have found that older adults who learned to use new technology were over twice as likely to adopt telehealth during the pandemic compared to those who did not [45]. In light of the "gray tsunami" phenomenon, where an increasing number of older adults are anticipated to require health care services, and the overwhelming burden of cancer among patients aged 65 years and older [20,21], it is important to recognize the urgency of addressing digital health literacy to minimize health disparities.

Strength and Limitations

This study has several limitations that impact the potential transferability of our findings to other settings. First, the sample was drawn from a single Cancer Center, with most participants being White and English-speaking, thereby limiting applicability to more diverse populations. Second, data collection for this study occurred during the latter half of 2021, and as such, should be considered within the context of the COVID-19 pandemic. Third, attitudes toward patient portals and telehealth use were evolving during the pandemic, and remote cancer care delivery was a necessity for some patients with cancer during this time. Fourth, by recruiting participants through the portal to sample patients who could adequately comment on telehealth experiences, we may have introduced a bias toward patients who are already more technologically proficient or feel positively about telehealth. This may have limited insights into

XSL•FO
RenderX

the experiences of older adults with lower digital literacy, potentially underrepresenting the experiences of those with significant barriers or negative perceptions. Anticipating this limitation, we purposefully recruited 3 individuals without telehealth experience to participate in an interview; however, the small overall patient sample size may also be a limiting factor. Similarly, by recruiting our patient interview sample from our survey sample, we may have introduced a potential response bias toward the inclusion of individuals who felt strongly toward telehealth (positively or negatively). Our recruitment for clinician and staff interviews could also have introduced bias given our reliance on suggestions by our clinician co-investigators. Strengths of this study include its use of the sociotechnical model as a guiding framework for the instrument design and analysis plan, and our inclusion of both patients and clinicians and staff. Finally, our approach of conducting a small number of targeted, preliminary clinician and staff interviews enabled us to assess the appropriateness and representativeness of our survey question items prior to launching our full surveys with patients and other clinicians and staff.

Future Directions

If telehealth expands across the cancer control continuum, it will be critical that health care organizations prioritize digital equity and inclusion of populations with lower digital literacy. This includes racial and ethnic minorities and older adults and their care partners, to ensure that the needs of these growing populations are adequately met and to provide a more balanced view of the challenges they may face. Future research would benefit from exploring the relative impacts of various interventions, such as digital navigator programs and tailored educational materials, on telehealth engagement and health outcomes.

Conclusions

This study contributes new evidence that older patients and their clinical care members welcome the expansion of telehealth use across the cancer control continuum. This is an important finding given the fact that digital health technologies, including but not limited to synchronous telehealth visits, are already integral components of health care delivery. It suggests that, as the population ages, health care systems will have the potential to reach a greater share of their patients by enhancing access to cancer services through telehealth and other digital health technologies.

Acknowledgments

The authors thank Ms Suzanne Ouyang and Ms Jessica L Lee for coordinating the study, and Ms Ana Belen Conrado for conducting interviews.

Funding

This study was funded by a Community-Engaged Research Pilot Award from the Office of Community Outreach, Equity & Engagement at Simmons Comprehensive Cancer Center, 2020-2022. The study was also supported by the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health (UL1 TR003163). Funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

Data are presented in the main manuscript; additional data may be available upon reasonable request from the authors.

Authors' Contributions

RTH and TPH contributed to conceptualization and methodology. RTH, BE, SC, JCM, and TPH conducted the investigation. RTH, JL, and AZ drafted the original manuscript. All authors contributed to manuscript review and editing. RTH and TPH were responsible for funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey recruitment consort diagram.
[PDF File, 94 KB - cancer_v12i1e73058_app1.pdf]

Multimedia Appendix 2
Illustrative quotes from patients, clinicians, and staff.
[DOCX File, 18 KB - cancer_v12i1e73058_app2.docx]

XSL•FO
RenderX

Checklist 1
GRAMMS checklist.
[DOCX File, 13 KB - cancer_v12i1e73058_app3.docx]

References

1. Shih KK, Arechiga AB, Chen X, et al. Telehealth preferences among patients with advanced cancer in the post COVID-19 vaccine era. J Pain Symptom Manage 2024 Jun;67(6):525-534. [doi: 10.1016/j.jpainsymman.2024.02.572] [Medline: 38467349]
2. Mahoney ME, Sridhar SS. Clinical trial reform in the post-COVID era. Ther Adv Med Oncol 2023;15:17588359231183676. [doi: 10.1177/17588359231183676] [Medline: 37431430]
3. Sim AJ, Redler G, Peacock J, et al. Harnessing COVID-driven technical innovations for improved multi-disciplinary cancer care in the post-COVID era: the virtual patient room. Cancer Control 2020;27(1):1073274820964800. [doi: 10.1177/1073274820964800] [Medline: 33023342]
4. Marks VA, Hsiang WR, Nie J, et al. Telehealth availability for cancer care during the COVID-19 pandemic: cross-sectional study. JMIR Cancer 2023 Nov 2;9:e45518. [doi: 10.2196/45518] [Medline: 37917149]
5. Al-Shamsi HO, Alhazzani W, Alhuraiji A, et al. A practical approach to the management of cancer patients during the novel coronavirus disease 2019 (COVID-19) pandemic: an international collaborative group. Oncologist 2020 Jun;25(6):e936-e945. [doi: 10.1634/theoncologist.2020-0213] [Medline: 32243668]
6. Report of the WHO-china joint mission on coronavirus disease 2019 (COVID-19). World Health Organization. 2020. URL: https://www.who.int/publications/i/item/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19) [accessed 2024-04-14]
7. Kim J, Linos E, Fishman DA, Dove MS, Hoch JS, Keegan TH. Factors associated with online patient-provider communications among cancer survivors in the United States during COVID-19: cross-sectional study. JMIR Cancer 2023 May 22;9:e44339. [doi: 10.2196/44339] [Medline: 37074951]
8. Oppong BA, Lustberg MB, Nolan TS, et al. Utilization of cancer survivorship services during the COVID-19 pandemic in a tertiary referral center. J Cancer Surviv 2023 Dec;17(6):1708-1714. [doi: 10.1007/s11764-022-01231-x] [Medline: 35895236]
9. Wu QL, Street RL Jr. Factors affecting cancer patients' electronic communication with providers: implications for COVID-19 induced transitions to telehealth. Patient Educ Couns 2020 Sep 28;103(12):2583-2587. [doi: 10.1016/j.pec.2020.09.036] [Medline: 33036815]
10. Islam JY, Camacho-Rivera M, Vidot DC. Examining COVID-19 preventive behaviors among cancer survivors in the United States: an analysis of the COVID-19 impact survey. Cancer Epidemiol Biomarkers Prev 2020 Dec;29(12):2583-2590. [doi: 10.1158/1055-9965.EPI-20-0801] [Medline: 32978173]
11. Goncalves Leite Rocco P, Reategui-Rivera CM, Finkelstein J. Telemedicine applications for cancer rehabilitation: scoping review. JMIR Cancer 2024 Aug 21;10:e56969. [doi: 10.2196/56969] [Medline: 39079103]
12. O'Neill L, Brennan L, Sheill G, Connolly D, Guinan E, Hussey J. Moving forward with telehealth in cancer rehabilitation: patient perspectives from a mixed methods study. JMIR Cancer 2023 Nov 9;9:e46077. [doi: 10.2196/46077] [Medline: 37943595]
13. Cox A, Lucas G, Marcu A, et al. Cancer survivors' experience with telehealth: a systematic review and thematic synthesis. J Med Internet Res 2017 Jan 9;19(1):e11. [doi: 10.2196/jmir.6575] [Medline: 28069561]
14. Agochukwu NQ, Skolarus TA, Wittmann D. Telemedicine and prostate cancer survivorship: a narrative review. Mhealth 2018;4:45. [doi: 10.21037/mhealth.2018.09.08] [Medline: 30505843]
15. Pang L, Liu Z, Lin S, et al. The effects of telemedicine on the quality of life of patients with lung cancer: a systematic review and meta-analysis. Ther Adv Chronic Dis 2020;11:2040622320961597. [doi: 10.1177/2040622320961597] [Medline: 33101621]
16. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 3;7(8):e016242. [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
17. Cancer control continuum. National Cancer Institute. 2020. URL: https://cancercontrol.cancer.gov/about-dccps/about-cc/cancer-control-continuum [accessed 2020-09-24]
18. Steffen LE, Boucher KM, Damron BH, et al. Efficacy of a telehealth intervention on colonoscopy uptake when cost is a barrier: the family CARE cluster randomized controlled trial. Cancer Epidemiol Biomarkers Prev 2015 Sep;24(9):1311-1318. [doi: 10.1158/1055-9965.EPI-15-0150] [Medline: 26101306]
19. Knudsen KE, Willman C, Winn R. Optimizing the use of telemedicine in oncology care: postpandemic opportunities. Clin Cancer Res 2021 Feb 15;27(4):933-936. [doi: 10.1158/1078-0432.CCR-20-3758] [Medline: 33229457]
20. Cinar D, Tas D. Cancer in the elderly. North Clin Istanb 2015;2(1):73-80. [doi: 10.14744/nci.2015.72691] [Medline: 28058345]
21. Given B, Given CW. Older adults and cancer treatment. Cancer 2008 Dec 15;113(12 Suppl):3505-3511. [doi: 10.1002/cncr.23939] [Medline: 19058145]

XSL•FO
RenderX

22. Irizarry T, Shoemake J, Nilsen ML, Czaja S, Beach S, DeVito Dabbs A. Patient portals as a tool for health care engagement: a mixed-method study of older adults with varying levels of health literacy and prior patient portal use. J Med Internet Res 2017 Mar 30;19(3):e99. [doi: 10.2196/jmir.7099] [Medline: 28360022]
23. Chien L, Roberts E, Soto-Perez-de-Celis E, et al. Telehealth in geriatric oncology: a novel approach to deliver multidisciplinary care for older adults with cancer. J Geriatr Oncol 2020 Mar;11(2):197-199. [doi: 10.1016/j.jgo.2019.06.011] [Medline: 31272897]
24. Ware P, Bartlett SJ, Paré G, et al. Using eHealth technologies: interests, preferences, and concerns of older adults. Interact J Med Res 2017 Mar 23;6(1):e3. [doi: 10.2196/ijmr.4447] [Medline: 28336506]
25. Aoun B, Ebbert JO, Ramar P, Roellinger DL, Philpot LM. Assessing the ability to use eHealth resources among older adults: cross-sectional survey study. JMIR Form Res 2025 Aug 6;9:e70672. [doi: 10.2196/70672] [Medline: 40768764]
26. Bogic M, Garcia JK, Morgan E, Louise-Hsu YC, Boyd A. Misinformation and digital health literacy among American Indian and Alaska Native people. Digit Health 2025;11:20552076251338893. [doi: 10.1177/20552076251338893] [Medline: 40386091]
27. Datta P, Eiland L, Samson K, Donovan A, Anzalone AJ, McAdam-Marx C. Telemedicine and health access inequalities during the COVID-19 pandemic. J Glob Health 2022 Dec 3;12:05051. [doi: 10.7189/jogh.12.05051] [Medline: 36462207]
28. Pérez Bentancur V, Tiscornia L. Iteration in mixed-methods research designs combining experiments and fieldwork. Sociol Methods Res 2024 May;53(2):729-759. [doi: 10.1177/00491241221082595]
29. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. Qual Saf Health Care 2010 Oct;19 Suppl 3(Suppl 3):i68-i74. [doi: 10.1136/qshc.2010.042085] [Medline: 20959322]
30. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
31. Creswell JW, Inoue M. A process for conducting mixed methods data analysis. J Gen Fam Med 2025 Jan;26(1):4-11. [doi: 10.1002/jgf2.736] [Medline: 39776872]
32. Mannheim I, Wouters EJM, van Boekel LC, van Zaalen Y. Attitudes of health care professionals toward older adults' abilities to use digital technology: questionnaire study. J Med Internet Res 2021 Apr 21;23(4):e26232. [doi: 10.2196/26232] [Medline: 33881408]
33. Jarva E, Oikarinen A, Andersson J, et al. Healthcare professionals' perceptions of digital health competence: a qualitative descriptive study. Nurs Open 2022 Mar;9(2):1379-1393. [doi: 10.1002/nop2.1184] [Medline: 35094493]
34. Kalicki AV, Moody KA, Franzosa E, Gliatto PM, Ornstein KA. Barriers to telehealth access among homebound older adults. J Am Geriatr Soc 2021 Sep;69(9):2404-2411. [doi: 10.1111/jgs.17163] [Medline: 33848360]
35. Zoorob D, Hasbini Y, Chen K, et al. Ageism in healthcare technology: the older patients' aspirations for improved online accessibility. JAMIA Open 2022 Oct;5(3):ooac061. [doi: 10.1093/jamiaopen/ooac061] [Medline: 35855421]
36. Sakaguchi-Tang DK, Turner AM, Taylor JO, Kientz JA. Connected personas: translating the complexity of older adult personal health information management for designers of health information technologies. AMIA Annu Symp Proc 2019;2019:1177-1186. [Medline: 32308915]
37. Mackwood M, Butcher R, Vaclavik D, et al. Adoption of telemedicine in a rural US cancer center amid the COVID-19 pandemic: qualitative study. JMIR Cancer 2022 Aug 16;8(3):e33768. [doi: 10.2196/33768] [Medline: 35895904]
38. Doueiri ZN, Bajra R, Srinivasan M, Schillinger E, Cuan N. Bridging the telehealth digital divide with collegiate navigators: mixed methods evaluation study of a service-learning health disparities course. JMIR Med Educ 2024 Oct 1;10:e57077. [doi: 10.2196/57077] [Medline: 39353186]
39. Jaso-Yim B, Eyllon M, Sah P, et al. Evaluation of the impact of a digital care navigator on increasing patient registration with digital mental health interventions in routine care. Internet Interv 2024 Dec;38:100777. [doi: 10.1016/j.invent.2024.100777] [Medline: 39410952]
40. Rodriguez JA, Charles JP, Bates DW, Lyles C, Southworth B, Samal L. Digital healthcare equity in primary care: implementing an integrated digital health navigator. J Am Med Inform Assoc 2023 Apr 19;30(5):965-970. [doi: 10.1093/jamia/ocad015] [Medline: 36795062]
41. Alpert JM, Taylor G, Hampton CN, Paige S, Markham MJ, Bylund CL. Clinicians' perceptions of the benefits and challenges of teleoncology as experienced through the COVID-19 pandemic: qualitative study. JMIR Cancer 2022 Feb 24;8(1):e34895. [doi: 10.2196/34895] [Medline: 35142622]
42. Berry DL, Hong F, Halpenny B, et al. Electronic self-report assessment for cancer and self-care support: results of a multicenter randomized trial. JCO 2014 Jan 20;32(3):199-205. [doi: 10.1200/JCO.2013.48.6662]
43. Gajarawala SN, Pelkowski JN. Telehealth benefits and barriers. J Nurse Pract 2021 Feb;17(2):218-221. [doi: 10.1016/j.nurpra.2020.09.013] [Medline: 33106751]
44. Rutledge CM, Kott K, Schweickert PA, Poston R, Fowler C, Haney TS. Telehealth and eHealth in nurse practitioner training: current perspectives. Adv Med Educ Pract 2017;8:399-409. [doi: 10.2147/AMEP.S116071] [Medline: 28721113]
45. Qin W. Technology learning and the adoption of telehealth among community-dwelling older adults during the COVID-19 outbreak. J Appl Gerontol 2022 Jul;41(7):1651-1656. [doi: 10.1177/07334648221085473] [Medline: 35255744]

XSL•FO
RenderX

Abbreviations

NCATS: National Center for Advancing Translational Sciences
REDCap: Research Electronic Data Capture

Edited by M Balcarras, N Cahill; submitted 24.Feb.2025; peer-reviewed by O Nguyen, Q Wang; revised version received 02.Dec.2025; accepted 30.Dec.2025; published 09.Feb.2026.

Please cite as:
Higashi RT, Etingen B, Lee J, Cole S, Mansour JC, Zhao A, Hogan TP
Patient and Clinician Perspectives on Expanding Telehealth Use for Older Adults Across the Cancer Control Continuum: Mixed Methods Study
JMIR Cancer 2026;12:e73058
URL: https://cancer.jmir.org/2026/1/e73058
doi:10.2196/73058

© Robin T Higashi, Bella Etingen, Jessica Lee, Suzanne Cole, John C Mansour, Alice Zhao, Timothy P Hogan. Originally published in JMIR Cancer (https://cancer.jmir.org), 9.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding Social Media Strategies to Improve Engagement in and Equitable Dissemination of Online Cancer Nutrition Information: Cross-Sectional Study

Ellie Harrison[1], MPH; Sara Buzali-Soto[2], MA, PhD; Eileen Rillamas-Sun[2], MPH, PhD; Heather Greenlee[2], ND, PhD; Echo L Warner[3], MPH, PhD

[1]Department of Epidemiology, School of Public Health, University of Washington, Seattle, United States

[2]Fred Hutch Cancer Center, Seattle, United States

[3]Department of Pediatrics, Huntsman Cancer Institute, University of Utah, 2000 Circle of Hope, Salt Lake City, UT, United States

Corresponding Author:
Echo L Warner, MPH, PhD
Department of Pediatrics, Huntsman Cancer Institute, University of Utah, 2000 Circle of Hope, Salt Lake City, UT, United States

Abstract

Background: Social media is a common source of cancer information for patients, caregivers, and survivors. Factors influencing engagement with oncology nutrition content on social media are not well-defined, despite the high prevalence of social media use for health information. The impact of bilingual content strategies on that engagement is also poorly understood.

Objective: This study aims to describe changes in postlevel engagement on a Facebook page used to disseminate evidence-based cancer nutrition information during the period before and after the implementation of a bilingual, multiformat content strategy and to identify creative and linguistic features associated with audience engagement at the post level.

Methods: This cross-sectional study analyzed 306 Facebook posts from Cook for Your Life over 12 months, comprising 9 months before the strategy change (T1) and 3 months after (T2). Postlevel engagement metrics (likes, comments, shares, and link clicks) and aggregated user demographics were obtained from Meta Business Suite. Trained coders annotated visual and creative features, including branding, post shape, and infographics, as well as topical categories. Linguistic Inquiry and Word Count characterized language features. Group differences across time periods were tested using *t* tests and chi-square tests. Generalized linear models estimated associations between post characteristics and engagement, adjusting for the time period.

Results: Users who engaged with oncology nutrition Facebook content (N=10,196) were mostly female (n=9500, 93.2%) participants, over 35 years of age (n=9311, 91.4%), based in the United States (n=6708, 71.4%), and English-language users (n=9419, 93.4%). Engagement tripled from T1 (mean 5.3, SD 4.2) to T2 (mean 17.7, SD 37.9; *P*<.001). Creative features such as branding and format were positively associated with engagement. Linguistically, there was a shift toward longer sentence structures, simplified vocabulary, increased analytical language, and a decline in emotionally charged and cognitive language in posts during T2. In models controlling for time period, posts that used bilingual content had 1.72-fold higher odds of audience engagement than English-only posts (95% CI 1.05 - 2.80).

Conclusions: Implementing a bilingual, multiformat content strategy may promote engagement with oncology nutrition information on Facebook and enhance the equitable dissemination of health information. The use of creative visual design elements in combination with language that is more analytical and emotionally neutral was associated with higher engagement with educational public health content. Future research should test bilingual language presentation more directly and examine whether exposure to such content influences health beliefs and downstream behaviors across additional platforms.

(JMIR Cancer 2026;12:e69886) doi:10.2196/69886

KEYWORDS

Facebook; content analysis; social media; engagement; cancer prevention; nutrition education; bilingual

Introduction

Emerging Health Communication on Social Media

The pervasive use of the internet, driven by the emergence of social media platforms, has revolutionized communication patterns, particularly within health-related contexts [1]. Social media, ranking as the third most frequent online activity, has become an integral part of human interaction, facilitating the exchange of ideas, information sharing, and the creation of online communities [2,3]. With the increasing prevalence of personal wireless devices, the popularity of social media applications continues to grow worldwide [4-6]. This

XSL•FO
RenderX

observation has prompted health communication experts to explore innovative opportunities for using social media to enhance population health [3,7-11].

Social media platforms are powerful tools for health communication interventions and offer opportunities to improve outreach and impact [12-14]. Notably, cancer organizations have increasingly embraced social media for outreach and patient support and have promoted technology-based interventions that use social media to support patients and caregivers in the period after a cancer diagnosis [15]. Studies across multiple scientific disciplines have highlighted the potential of social media to improve health through the efficient exchange of user-generated health information, increased perceived social support, and the facilitation of public health programs such as dietary interventions [16-23].

It is valuable to study the language organizations use in online cancer communities related to oncology nutrition education, considering the rapid transformations in the communication landscape brought on by participative internet use and social media alongside the potential for social media to amplify misinformation [24]. Research studying the association between engagement and the language used within social media aligns with the National Cancer Institute's focus on bringing cancer research studies to individuals in their communities [25]. The first step in this effort is to identify the characteristics of current social media users. *Facebook* is a social media platform that allows users to share and interact with stories, images, and videos. In the United States, Facebook is widely used within the adult population, more commonly by females than males [26]. On this platform, *engagement* is defined as any action users take on a Facebook page or post [27]. The most common examples of Facebook engagement actions are reactions (eg, likes), comments, and shares, but engagement can also include video views, link clicks, or post saves. Understanding social media usage and how people engage with information online will support health communication efforts aiming to use social media effectively to promote the dissemination of trusted, evidence-based oncology nutrition content online. This may inform the work of health science institutions to design effective population health campaigns that use digital technology.

Cook for Your Life

Cook for Your Life (CFYL), a website based at Fred Hutchinson Cancer Center (FHCC), stands out as a pioneer in promoting healthy cooking for individuals with cancer [28]. Founded in 2007 by a 2-time cancer survivor as a community-based nonprofit in New York City teaching in-person culinary and nutrition courses to cancer survivors, CFYL launched its website in 2012 to address the unmet need for culinary and nutrition education in the field of oncology. In 2019, the website was transferred to FHCC to continue serving as a bilingual (English/Spanish) research platform and community resource under their institutional support. The content produced by CFYL, which is grounded in science and supported by 2 advisory boards composed of FHCC leaders and experts—Scientific & Medical Advisors and Clinical Oncology Nutrition Advisors—caters to patients with cancer, survivors, individuals interested in cancer prevention, as well as health care professionals and oncology societies. Disseminated through social media platforms such as Facebook, X (formerly known as Twitter), Pinterest, YouTube, and Instagram, CFYL's library comprises a rich collection of science-based recipes, articles, and videos.

In a strategic move in September 2021, CFYL implemented a bilingual multiformat social media strategy, producing content in both English and Spanish. This targeted approach, grounded in evidence-based communication techniques, aimed to enhance engagement and broaden the audience for oncology nutrition content. This strategy was executed on the CFYL Facebook page, and the content produced from the strategy is publicly available for review online. The evidence for implementing this content strategy drew on visual communication and health communication research that links visuals and branding to higher social sharing, greater engagement, and improved persuasive impact for health information. The incorporation of bilingual content into CFYL's social media strategy serves as a unique opportunity to investigate linguistic changes in Facebook content and their influence on engagement.

With that being said, the importance of using social media to communicate science-based information to people affected by cancer warrants further research on this topic, and few studies have examined which bilingual, visual, and linguistic features of oncology nutrition Facebook posts are associated with engagement. We hope to unravel the characteristics of social media posts that promote engagement with the CFYL content on Facebook, focusing on nutrition and cancer, before and after integrating targeted bilingual, multiformat content into their social media approach. Our objective is to describe changes in postlevel engagement on the CFYL Facebook page in the period before and after the implementation of this content strategy and to identify creative and linguistic features associated with audience engagement at the post level. We aim to deepen our understanding of the factors influencing user engagement with health information via social media platforms, specifically within online cancer communities.

Methods

Study Timeline

This descriptive cross-sectional study uses linguistic content analysis to examine textual data from Facebook posts about nutrition and cancer prevention and to describe (1) indicators of social engagement, (2) valence (ie, positive/negative terms), and (3) linguistic content (eg, emotional terms/pronouns). The study analyzes CFYL Facebook content over a 12-month period during the COVID-19 pandemic. We divided the 12-month observation window into 2 prespecified periods based on the implementation of a revised social media strategy in September 2021. Time point 1 (T1) represented 9 months before implementation (January 1-September 30, 2021; n=202 posts), and time point 2 (T2) represented the first 3 months after implementation (October 1-December 31, 2021; n=104 posts). T2 was intentionally restricted to the immediate postimplementation period to capture early changes in posting practices while minimizing contamination from later programmatic changes or cultural events. This temporal division

XSL•FO
RenderX

facilitated the examination of post characteristics associated with engagement, linguistic content, and other factors in the introduction of a bilingual, multiformat social media strategy by CFYL.

Facebook Metadata and Sampling

The analytic sample comprised 306 original posts published by the official CFYL Facebook page during the 12-month observation period from January 1, 2021, through December 31, 2021. Each post was created and published by CFYL page administrators as part of the organization's dissemination of evidence-based oncology nutrition content, including links to CFYL recipes, articles, videos, and branded educational graphics. The unit of analysis was the individual CFYL-authored Facebook post. User-generated comments and posts from non-CFYL accounts were not included as analytic units; engagement with CFYL posts was captured through platform-generated metadata.

Postlevel reach, impressions, engagement, and other key metrics (post type: photo, link, video, or albums) were exported for each CFYL-authored post from Meta Business Suite, the standard analytics interface for Facebook pages. Reach reflects the number of unique accounts that saw a post; impressions reflect total on-screen displays, and engagement aggregates user interactions such as reactions/likes, comments, shares, and link clicks. These are widely used reporting metrics and serve as a data source for digital health research. Meta Business Suite also provided aggregate demographic information describing the audience that interacted with the CFYL Facebook page. Demographic variables extracted from Meta Business Suite include gender, age group, language spoken, and country. Access to this Facebook Metadata is limited to the administrators of the CFYL Facebook account and is not publicly available.

Qualitative Content Analysis

Certain variables of interest were not available directly from Meta Business Suite (branding, post shape, image content, and topic). Content codes were manually collected from the CFYL Facebook page by a research assistant. This manual extraction included the post text in both English and Spanish, the number of comments and emojis, and image content data. Each post was assigned a unique identifier to enable linkage to Meta Business Suite variables. Additional data included postpublication dates and hyperlinks. These manual content codes were used to construct independent variables for analysis to contextualize observed pre-post shifts in creative strategy and topical focus. To ensure data accuracy, another investigator quality-checked this manually collected data. Next, to analyze how the content of posts was associated with engagement, we performed qualitative content analysis. A codebook from prior social media research was adapted and implemented to code the visual and topical content of posts [29]. Discrepancies and inconsistencies were discussed among the research team, and the codebook was refined over time (Multimedia Appendix 1). User data were grouped according to categorical coding units with adequate variation for statistical comparisons. Manual content codes (branding, post shape, and topic) were used to construct independent variables for quantitative analyses and to contextualize observed pre-post shifts in creative strategy and topical focus.

Linguistic Variables

To analyze the linguistic content of CFYL Facebook posts, Linguistic Inquiry and Word Count (LIWC) was used. LIWC is a widely recognized content analysis platform that provides valuable insights into the linguistic characteristics of textual content [30]. It includes various dimensions such as emotional tone, positivity, negativity, pronouns, and numerous other linguistic attributes. The selection of specific LIWC dimensions was based on the research questions, hypotheses, and preliminary findings, aligning with the study's objectives. LIWC's "Analytic" summary dimension reflects formal, logical, and hierarchical thinking styles, with higher values indicating more analytic language. "Tone" summarizes overall emotional tone, and "Big Words" reflects the use of longer words (commonly operationalized as words with 6 or more letters).

There are also more specific index variables used to assess distinct linguistic processes. For example, the "Cognitive" dimension indicates reasoning and stance-taking, "Authentic" indicates the use of a more personal, self-disclosing, first-person style, and "Perceptual" notes the presence of sensory language. Informal language indicators such as "Conversational" and "Netspeak" capture a more chat-like style and internet shorthand. Social orientation is captured by "Social processes" and related subcategories (eg, "Social referents," "Family," and gender references), whereas "Affect," "Emotion," "Positive emotion," and "Negative emotion" quantify the proportion of emotion-laden words. Topic-oriented categories such as "Health" and "Illness" index the extent to which posts reference health and disease experiences. The "Netspeak" dimension describes internet-specific language such as emojis and internet acronyms (eg, LOL [laughing out loud] and OMG [oh my god]).

A stop list was applied to exclude noncontent words, ensuring that the analysis focused on meaningful language. Textual data from hashtags were also omitted from the analysis, as they were considered irrelevant and repetitive content. Then, LIWC was applied to estimate the percentage of terms in each post containing different language types (with scores ranging from 0 to 100). Higher values indicate a greater proportion of words in the post belonging to that category (or higher scores on LIWC summary dimensions), and lower values indicate less frequent use of that linguistic feature.

Engagement Outcomes

Our primary outcome variables were indicators of social engagement within online cancer communities, defined as the number of likes, comments, and shares on each CFYL Facebook post. This definition using postlevel engagement was selected based on its prior use in digital health research and the presence of these features across other social media platforms [23,31]. We also extracted postlevel engaged user metrics from Meta Business Suite, which reflect the number of users who performed one or more engagement actions on the post. Outcomes of engagement included postlevel summary statistics (mean, SD) for each time period. High engagement indicated that posts were relevant to the target audience. Reach and

XSL•FO
RenderX

impressions were also extracted to characterize post exposure. We integrated platform analytics (Meta Business Suite), manual content coding of post features and topics, and LIWC text analysis to characterize how changes in post design, topic focus, and linguistic style aligned with changes in engagement across time periods.

Statistical Analysis

To facilitate statistical comparisons, textual data were grouped based on categorical coding units with sufficient variation. We descriptively summarized post characteristics, engagement metrics, and aggregate audience demographics obtained from Meta Business Suite using both total counts and postlevel distributional summaries. Bivariate analyses were used to compare postlevel characteristics between the preimplementation (T1) and postimplementation (T2) periods. Independent-samples *t* tests were used for continuous variables, including engagement metrics and LIWC summary measures, whereas chi-square tests were used for categorical variables such as post format, branding, bilingual presentation, and topic categories. To examine associations between post characteristics and engagement while accounting for differences between time periods, we fit generalized linear models, treating the number of total engaged users as a continuous postlevel outcome, and reported odds ratios (ORs) with 95% CIs. In these models, the time period was included as a covariate to account for temporal differences between T1 and T2. All statistical tests were 2-sided, and $P<.05$ was considered statistically significant. This regression model provides insights into the dynamics of engagement and linguistic content on the CFYL Facebook page while controlling for temporality.

Ethical Considerations

This study analyzed CFYL's public Facebook posts and aggregated, deidentified engagement metrics exported from Meta Business Suite at the post level. Manual coding was limited to CFYL-owned post text and images, not user comments or profiles. No direct identifiers (eg, names, usernames, and profile links) were collected, and no individual user content from private accounts or direct messages was accessed. Analyses were conducted on deidentified, aggregated data; demographic information (eg, age group, country, and language) was available only in platform-provided summary form. The FHCC determined the project exempt from institutional review board approval, and no individual consent was required. FHCC provides funding for CFYL and supported research activities. Procedures for data use complied with ethical expectations for internet-based research and did not involve automated scraping or circumvention of platform protections.

Results

Characteristics of Individuals Who Engaged With CFYL Content on Facebook

The final analytic dataset included 306 original CFYL-authored Facebook posts published between January 1 and December 31, 2021. Of these, 202 posts were published during T1 (preimplementation) and 104 posts during T2 (postimplementation). Posts disseminated CFYL oncology nutrition content, including recipes, educational articles, visual graphics, and links to CFYL web content. According to aggregate demographic summaries provided by Meta Business Suite, users who engaged with CFYL content during the 12-month monitoring period were predominantly female (n=9500, 93.2%) and over 35 years of age (n=9311, 91.4%), with the highest engagement present in the group of people aged 55 to 64 years (n=2708, 26.6%). These findings are consistent with broader platform usage patterns for Facebook among midlife adults. The majority of the CFYL audience is based in the United States (n=6708, 71.4%) and identified by the platform as English-language users (n=9419, 93.4%). A smaller proportion of engaged users were identified as Spanish-language users (n=311, 3.1%) with an even lesser number of users residing in a Spanish-speaking country (n=101 from Mexico, 1.1%), indicating that there is also a community of Spanish-speaking individuals residing in the United States who engaged with the CFYL content during this time period (Table 1). Because these demographic data were available only in aggregate platform-provided categories, they describe the engaged audience at the page level rather than post-specific audiences.

XSL•FO
RenderX

Table . Demographics of users who engaged with Cook for Your Life (CFYL) content during the 12-month monitoring period (N=10,196).

Demographics	n (%)
Gender	
Female	9500 (93.2)
Male	626 (6.1)
Other	70 (0.7)
Location	
United States	6708 (71.4)
Canada	1360 (14.5)
United Kingdom	855 (9.1)
Australia	127 (1.4)
Mexico	102 (1.1)
South Africa	99 (1.1)
Ireland	85 (0.9)
India	53 (0.6)
Age group (y)	
13 - 17	4 (0.04)
18 - 24	86 (0.1)
25 - 34	795 (7.8)
35 - 44	1835 (18.0)
45 - 54	2486 (24.4)
55 - 64	2708 (26.6)
≥65	2282 (22.4)
Language	
English	9419 (93.4)
Spanish	311 (3.1)
French	224 (2.2)
Portuguese	44 (0.4)
Italian	22 (0.2)
Arabic	14 (0.1)
Greek	13 (0.1)
Other	41 (0.4)

Post Characteristics Associated With Engagement With CFYL Content on Facebook

The overall engagement with posts increased from T1 (mean 5.3, SD 4.2) to T2 (mean 17.7, SD 37.9; *P*<.001; Table 2). Similar increases were observed across engagement components, including likes (T1: mean 2.5, SD 2.0 vs T2: mean 12.5, SD 33.5; *P*<.01), shares (T1: mean 0.4, SD 0.7 vs T2: mean 0.8, SD 1.1; *P*<.01), comments (T1: mean 0.1, SD 0.6 vs T2: mean 0.5, SD 1.3; *P*<.01), and link clicks (T1: mean 2.2, SD 2.6 vs T2: mean 3.9, SD 7.5; *P*=.003; Table 2), underscoring a surge in user interaction with the content. The magnitude of change in user engagement metrics can be visualized in Figure 1. Post format and creative features shifted over time, with a decrease in photography posts (T1: 98.5% vs T2: 67.3%) and increases in text-dominant posts (T1: 1.5% vs T2: 32.7%), infographics (T1: 0.0% vs T2: 7.7%), and branding (T1: 19.8% vs T2: 94.2%; all *P*<.01; Table 2). Similarly, the qualitative content analysis results indicate that the presence of specific content topics within the posts exhibited significant changes during T2. For instance, there were higher proportions of posts focused on cancer (T1: 25.7% vs T2: 58.6%; *P*<.01), nutrition and diet (T1: 64.4% vs T2: 84.6%; *P*<.01), health care support (T1: 18.3% vs T2: 40.4%; *P*<.01), and breast cancer–specific content (T1: 0.5% vs T2: 9.6%; *P*<.01), suggesting a strategic focus on these themes to enhance engagement and promote actionable oncology nutrition information online (Table 2). Bilingual content presentation was also rare in the preimplementation phase but nearly universal after applying the new content strategy (T1: 0.5%, T2: 99.0%; *P*<0.1, Table 2).

XSL•FO
RenderX

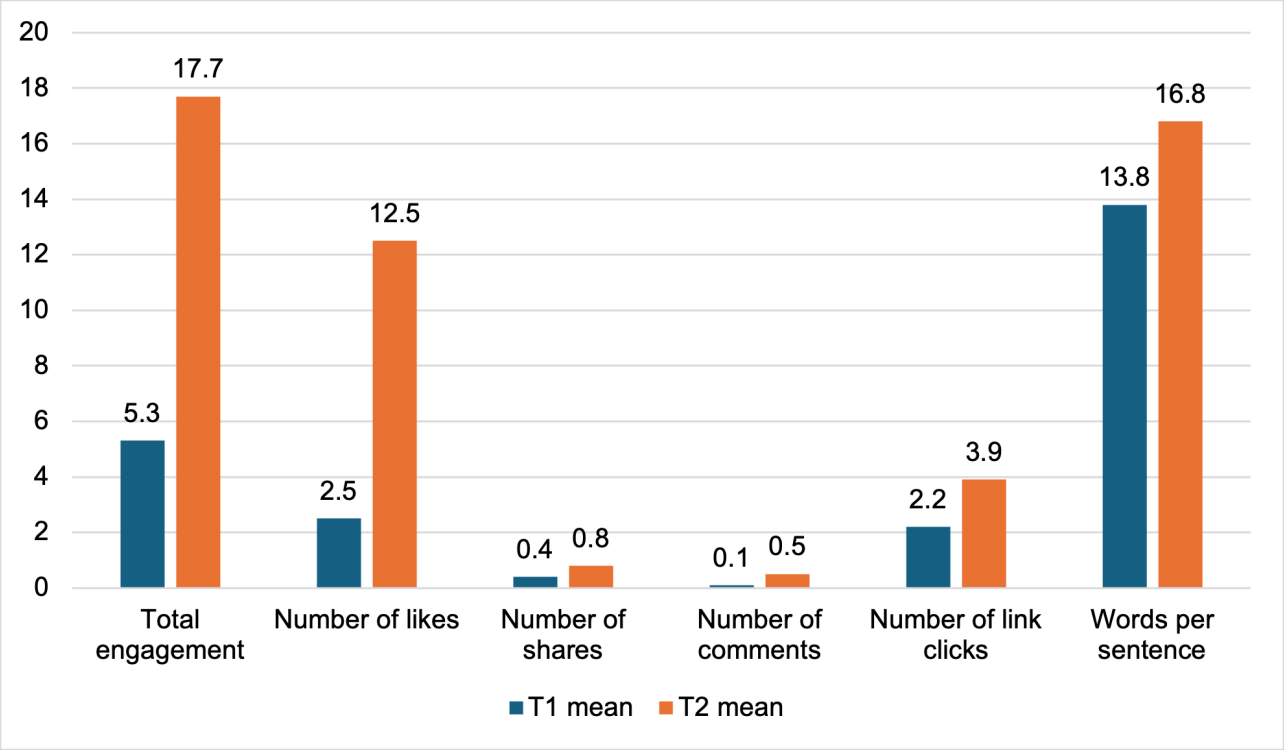
Table . Engagement with posts during 12-month monitoring period by characteristics of posts (N=306).

Engagement type	T1 (n=202)	T2 (n=104)	*P* value
Total engagement, mean (SD)	5.3 (4.2)	17.7 (37.9)	<.001[a]
Number of likes, mean (SD)	2.5 (2.0)	12.5 (33.5)	<.001[a]
Number of shares, mean (SD)	0.4 (0.7)	0.8 (1.1)	<.001[a]
Number of comments, mean (SD)	0.1 (0.6)	0.5 (1.3)	<.001[a]
Number of link clicks, mean (SD)	2.2 (2.6)	3.9 (7.5)	.003[a]
Words per sentence, mean (SD)	13.8 (3.8)	16.8 (3.8)	<.001[a]
Image content, n (%)			
Text	3 (1.5)	34 (32.7)	<.001[a]
Photography	199 (98.5)	70 (67.3)	<.001[a]
Infographic	0 (0.0)	8 (7.7)	<.001[a]
GIF	0 (0.0)	3 (2.9)	.02[a]
Branding	40 (19.8)	98 (94.2)	<.001[a]
Square shape	12 (5.9)	94 (90.4)	<.001[a]
Rectangular shape	190 (94.1)	10 (9.6)	<.001[a]
Bilingual post	1 (0.5)	103 (99.0)	<.001[a]
Content topic, n (%)			
Immune system	3 (1.5)	7 (6.7)	.01[a]
Chemotherapy	1 (1.5)	4 (3.9)	.03[a]
Radiation	1 (0.5)	2 (1.9)	.23
Cancer	52 (25.7)	61 (58.6)	<.001[a]
Survivorship	4 (2.0)	1 (1.0)	.50
Caregiver	0 (0.0)	2 (1.9)	.05[a]
Health care support	37 (18.3)	42 (40.4)	<.001[a]
Prevention	41 (20.3)	30 (28.8)	.09
Nutrition and diet	130 (64.4)	88 (84.6)	<.001[a]
Heart health	0 (0.0)	1 (1.0)	.16
Cancer type			
Gynecological	1 (0.5)	0 (0.0)	.47
Colorectal	1 (0.5)	1 (1.0)	.63
Breast	1 (0.5)	10 (9.6)	<.001[a]
Blood	0 (0.0)	2 (1.9)	.048[a]
Bladder	1 (0.5)	0 (0.0)	.47

[a]Indicates statistical significance at $P<.05$.

XSL•FO
RenderX

Figure 1. Changes in user engagement with Cook for Your Life (CFYL) across the implementation period.



Engagement by Lexical Content

The results from the natural language processing analysis using LIWC produced a large number of linguistic and semantic dimensions, which may be interpreted as an exploratory descriptive comparison of language style across time periods before and after the implementation of the targeted social media content strategy. T2 posts showed higher Analytic scores, longer sentences, and lower rates of affect- and cognition-related terms, indicating a shift toward more formal, informational language. The average number of words per sentence increased from mean 13.83 (SD 3.84) at T1 to mean 16.81 (SD 3.85) at T2 (P<.01; Table 2), demonstrating a shift toward longer sentence structures. The linguistic analysis revealed that posts in the latter time period (T2) exhibited a substantial increase in the use of "analytical" language from mean 80.04 (SD 18.05) at T1 to mean 87.70 (SD 5.66) at T2 (P<.001; Table 3), along with greater use of "Big Words" from mean 32.85 (SD 7.48) at T1 to mean 39.66 (SD 5.27) at T2 (P<.01; Table 3), suggesting a shift toward evidence-based logic and the use of complex medical terminology related to health science, oncology, and nutrition. Conversely, the use of words included in the LIWC dictionary declined from mean 64.92 (SD 8.34) at T1 to mean 37.23 (SD 5.76) at T2 (P<.001; Table 3), suggesting the use of a scientific vocabulary with uncommon words appearing more frequently in the text of T2 posts. Other linguistic style indices such as Linguistic (mean 44.84, SD 8.49 to mean 19.10, SD 3.49; P<.001) and Function words (mean 31.18, SD 8.48 to mean 12.36, SD 3.02; P<.01) also declined from T1 to T2 (Table 3).

XSL•FO
RenderX

Table . Lexical content by time period (N=306 posts).

Lexical category	T1, mean (SD)	T2, mean (SD)	*P* value
Analytical thinking	80.04 (18.05)	87.70 (5.66)	<.001
Dictionary	64.92 (8.34)	37.23 (5.76)	<.001
Linguistic	44.84 (8.49)	19.10 (3.49)	<.001
Tone	39.95 (34.60)	28.86 (19.79)	.003
Big words	32.85 (7.48)	39.66 (5.27)	<.001
Function	31.18 (8.48)	12.36 (3.02)	<.001
Authentic	19.03 (21.64)	1.12 (0.60)	<.001
Drives			
Affiliation	1.77 (2.41)	1.12 (1.20)	.009
Cognition			
Cognition	7.33 (4.6)	3.0 (1.8)	<.001
Cognitive processes	6.65 (4.47)	2.64 (1.59)	<.001
Differentiation	3.08 (2.70)	0.85 (0.80)	<.001
Tentative	2.38 (2.22)	0.50 (0.61)	<.001
Insight	1.01 (1.92)	0.60 (0.75)	.02
Discrepancy	1.00 (1.48)	0.49 (0.76)	.001
Certitude	0.43 (1.00)	0.08 (0.25)	<.001
Affect			
Affect	6.11 (3.83)	3.76 (1.50)	<.001
Positive tone	3.59 (3.25)	2.07 (1.25)	<.001
Emotion	2.88 (2.08)	1.62 (0.93)	<.001
Negative tone	2.44 (1.54)	1.65 (0.83)	<.001
Negative emotion	2.15 (1.26)	1.34 (0.64)	<.001
Positive emotion	0.59 (1.38)	0.24 (0.42)	.01
Social processes			
Social processes	5.96 (4.05)	4.17 (2.54)	<.001
Social referents	4.58 (3.41)	2.66 (1.50)	<.001
Family	0.05 (0.35)	0.24 (0.42)	<.001
Conflict	0.04 (0.36)	0.14 (0.27)	.01
Male references	0.01 (0.14)	0.23 (0.43)	<.001
Culture			
Culture	2.23 (1.37)	1.45 (0.67)	<.001
Technology	2.06 (1.23)	1.31 (0.63)	<.001
Politics	0.01 (0.12)	0.07 (0.22)	.001
Physical			
Health	1.96 (3.57)	6.69 (3.29)	<.001
Food	6.57 (4.44)	3.81 (2.39)	<.001
Illness	0.83 (2.08)	5.64 (3.09)	<.001
States			
Acquire	1.12 (1.52)	0.37 (0.51)	<.001
Motives			
Allure	4.74 (3.48)	1.87 (1.10)	<.001

XSL•FO
RenderX

Lexical category	T1, mean (SD)	T2, mean (SD)	*P* value
Perception			
Physical	8.63 (5.33)	10.84 (3.57)	<.001
Perception	7.26 (4.30)	2.09 (1.46)	<.001
Space	4.87 (3.15)	1.06 (0.97)	<.001
Present focus	2.85 (2.42)	1.08 (0.88)	<.001
Time	2.64 (2.81)	1.06 (1.03)	<.001
Conversational	2.18 (1.36)	1.47 (0.69)	<.001
Netspeak	2.14 (1.33)	1.45 (0.68)	<.001
Future focus	0.82 (1.59)	0.31 (0.48)	.001
Motion	0.71 (1.36)	0.19 (0.34)	<.001
Visual	0.63 (1.34)	0.34 (0.58)	.04
Feeling	0.59 (1.34)	0.19 (0.42)	.003
Fulfill	0.44 (1.09)	0.14 (0.34)	.005
Want	0.27 (0.80)	0.07 (0.31)	.02

Informal language markers such as the use of conversational language (mean 2.18, SD 1.36 to mean 1.47, SD 0.69; *P*<.001) and netspeak (mean 2.14, SD 1.33 to mean 1.45, SD 0.68; *P*<.001) decreased significantly from T1 to T2 (, Table 3), along with a decrease in the Authentic summary score (mean 19.03, SD 21.64 to mean 1.12, SD 0.60; *P*<.001; Table 3). Affect-related categories revealed significant changes in emotional tone, including affect (mean 6.11, SD 3.83 to mean 3.76, SD 1.50; *P*<.001), positive emotion (mean 0.59, SD 1.38 to mean 0.24, SD 0.42; *P*=.01), and negative emotion (mean 2.15, SD 1.26 to mean 1.34, SD 0.64; *P*<.001), with decreases in both positive tone (mean 3.59, SD 3.25 to mean 2.07, SD 1.25; *P*<.001) and negative tone (mean 2.44, SD 1.54 to mean 1.65, SD 0.83; *P*<.001; Table 3). These findings signify a shift toward less emotionally charged language in T2 posts.

Cognition-related categories were lower at T2, including cognitive processes (mean 6.65, SD 4.47 to mean 2.64, SD 1.59; *P*<.001), tentative language (mean 2.38, SD 2.22 to mean 0.50, SD 0.61; *P*<.001), discrepancy (mean 1.00, SD 1.48 to mean 0.49, SD 0.76; *P*<.001), and certitude (mean 0.43, SD 1.00 to mean 0.08, SD 0.25; *P*<.001; Table 3), suggesting a shift toward certainty and scientific reasoning. There was also less perception-related language used in T2 posts, indicating that the content shifted away from describing subjective experiences with sensory terms and toward promoting objective scientific evidence (mean 7.26, SD 4.30 to mean 2.09, SD 1.46; *P*<.001). Additionally, the use of physical language related to health (mean 1.96, SD 3.57 to mean 6.69, SD 3.29; *P*<.001) and illness terms (mean 0.83, SD 2.08 to mean 5.64, SD 3.09; *P*<.001) was greater at T2, suggesting a shift in content topics toward cancer-centered language.

Several characteristics of posts were associated with engagement in univariate regression analyses controlling for the time period (Table 4). First, posts that used the bilingual multiformat content strategy were associated with higher engagement (OR 1.72, 95% CI 1.05 - 2.80; *P*=.03). The presence of CFYL branding (OR 1.08, 95% CI 1.04 - 1.13; *P*<.001) and square-formatted images (OR 1.11, 95% CI 1.06 - 1.17; *P*<.001) was also associated with higher engagement, whereas rectangular-formatted posts were associated with lower engagement (OR 0.90, 95% CI 0.86 - 0.94; *P*<.001; Table 4). Only bilingual content was associated with increased odds of engagement after adjusting for the time point (data not shown). These findings underscore the influence of linguistic and thematic elements on user engagement with CFYL content on Facebook, even when adjusting for variations in the time period of the posts.

XSL•FO
RenderX

Table . Characteristics of posts and lexical content associated with engagement across all posts controlling for time period (N=306).

Post characteristics	OR (95% CI)	*P* value
Post structure		
Mostly text	1.00 (0.99 - 1.01)	.61
Mostly image	1.00 (0.98-1.01)	.61
Infographic	0.99 (0.93 - 1.06)	.78
With branding	1.08 (1.04 - 1.13)	<.001[a]
GIF	1.00 (0.97 - 1.04)	.80
Square shape	1.11 (1.06 - 1.17)	<.001[a]
Rectangular shape	0.90 (0.86 - 0.94)	<.001[a]
Bilingual	1.72 (1.05 - 2.80)	.03[a]
Topic		
Immune system	0.99 (0.95 - 1.04)	.83
Chemotherapy	0.97 (0.84 - 1.13)	.74
Radiation	0.97 (0.79 - 1.19)	.77
Cancer	1.00 (0.99 - 1.01)	.33
Survivorship	0.98 (0.96 - 1.12)	.77

[a]Indicates statistical significance at $P<.05$.

Changes observed in engagement metrics during T2 coincided with co-occurring shifts in post design and language. Specifically, the increase in branded and square-formatted creative (manual coding; Table 2) aligned with positive associations between these features and engagement in regression analyses (Table 4). Similarly, increases in posts coded as cancer, health care support, and nutrition-focused (Table 2) aligned with LIWC shifts showing more health and illness language and less food language (Table 3), suggesting that posts increasingly framed nutrition within cancer and treatment contexts during the strategy period.

Discussion

Principal Findings

This study analyzed 306 original CFYL-authored Facebook posts published over 12 months and described both overall engagement volume and postlevel engagement patterns before and after the implementation of a bilingual, multiformat content strategy. Across the observation window, engagement with CFYL content was concentrated among women, adults aged 35 years and older, and predominantly English-language users based in the United States, according to aggregate platform demographic summaries. These demographic results are in contrast to the epidemiology of cancer in America, where there is a much broader distribution of demographics across patients and survivors. One reason for this finding may be that younger adults are more likely to engage with social media platforms and have demonstrated high rates of digital health literacy. This distinct cohort could also be interpreted as the population of female caregivers who provide the invisible labor of seeking and implementing health information within their home and community. The demographic data might suggest that women act as the primary drivers of health-related caregiving and nutrition management within their social support networks.

In analyses controlling for the time period, bilingual posts had higher odds of engagement, and creative features, including CFYL branding and formatted images, were also positively associated with engagement. Linguistically, posts produced with this content strategy used more analytical language and longer sentences, with fewer conversational, authentic, affective, and cognitive terms in T2 than T1. These results amount to more formal, scientific language and less emotionally charged vocabulary used in T2, aligning with best practices for organizational evidence-based messaging. Content-related lexical categories also shifted, with increases in health and illness terms, indicating a greater emphasis on nutrition in the context of cancer and treatment-related concerns. Metrics of engagements per post increased over time, and in models that adjusted for the time period, bilingual posts had higher odds of engagement.

Creative features such as institutional branding and purpose-designed visual content were also positively associated with engagement. Visual branding implementation associated with CFYL and the Fred Hutch Cancer Center may have the potential to drive higher credibility among the target content audience. Given the observational design and co-occurring changes to content and format, these results should be interpreted as associations rather than causal effects. However, these findings emphasize the importance of strategic social media approaches in increasing user interaction with high-quality cancer information on social media platforms [32-34].

XSL•FO
RenderX

Comparison to Prior Work

Widespread access to the internet has drastically shifted the sources of health information from health care providers to user-generated online content. Simultaneously, access to high-quality cancer information online may not be equitable. This is especially true in a global online community where the source and credibility of cancer information are not always transparent. Incorporating bilingual cancer content on the CFYL Facebook page extended accessibility for Spanish-speaking individuals. Bilingual social media strategies have been documented to promote user engagement with online health promotion content in other settings as well, with Latino Facebook users tending to show more engagement with either Spanish or bilingual posts compared to English posts [35]. This suggests that cancer centers, hospitals, and cancer advocacy groups that intend to promote accessible cancer information online should strongly consider a bilingual approach, matching the linguistic needs of the target population.

Differences in health literacy on a global scale also emphasize the need for cancer information to be delivered at an appropriate literacy level, particularly given our current digital age where people worldwide have access to social media content [36]. The observed rise in the use of complex language elements represents a conscious effort to make the content more informative and scientifically accurate. However, the longer sentence structure and use of more complex, analytical language may present comprehension challenges for diverse audiences, such as individuals with varying levels of health literacy. Simplifying language in health communication is a recognized strategy to ensure that information is easily understood and accessible to a wider spectrum of readers. It may be valuable for CFYL to assess the health literacy of their audience and tailor social media content in order to make critical information about cancer nutrition more accessible and understandable to a broader public.

Moreover, the shift toward less emotionally charged language, as evidenced by the decrease in language that uses both positive and negative tones, aligns with the expected findings observed in organizational communication. Previous research has indicated that organizations, especially those in the health domain, tend to use fewer emotionally charged words compared to individual users [37]. This strategic approach aims to maintain a professional and informative tone, avoiding potential misinterpretations that emotionally charged language might introduce. It is also plausible that less emotional, more objective language is interpreted as being more credible and scientifically accurate. We also observed declines in perceptual processes (see/hear/feel), which likely reflect fewer sensory descriptors and a more didactic, informational tone used in the new content strategy. The decline of conversational and internet-specific language suggests fewer informal linguistic markers, which aligns with the educational content strategy used by cancer research institutes. This can increase credibility for an audience seeking trustworthy nutrition guidance.

We found that visual content such as photo media, branding, and infographics are useful tools to improve engagement with cancer nutrition information on social media, which is supported by prior literature on the topic within the field of health information and cancer research [32-34]. Visual content is commonly used across social media platforms, and social media posts with visual content are more frequently shared with other users and are associated with engagement metrics (including likes and comments) [34]. In addition to its impact on user engagement, visual content improves persuasive impact [38]. Other factors associated with higher engagement across time points were branding, square-shaped posts, and bilingual content. Square-shaped posts represent purposively selected photos that were identified by the CFYL content creators as opposed to rectangular-shaped posts, which represent preselected images that are input automatically when adding an external link into a Facebook post. Therefore, this finding does not imply an attraction to a specific geometrical shape but it suggests a preference for intentional photos over simple links. Institutions and researchers aiming to improve engagement with their social media content may use these strategies.

Strengths and Limitations

Strengths of this study design include the integration of platform analytics with manual content coding and the addition of linguistic analysis to characterize message features. There are also limitations of this approach that should be considered in the interpretation of our results. First, we cannot attribute engagement changes solely to bilingual content due to the multiple changes that occurred around the same time as the bilingual strategy (eg, more branding, more square images, and topic shifts). Because exposure varied and multiple changes occurred alongside the strategy, causality cannot be inferred. We attempted to mitigate this by comparing post characteristics across periods and by including the time period in regression models; however, residual confounding by unmeasured time-varying factors likely remains. Future studies should use longer, balanced observation windows, staggered rollouts, or interrupted time-series designs to strengthen causal inference.

Second, we could not directly compare engagement by language because nearly all posts after implementation were bilingual. This limits inferences about the relative contribution of Spanish vs English content. Although we examined overall engagement changes and the presence of bilingual formatting, this approach cannot isolate language-specific effects. Future work could stratify analyses by audience language segments when platform analytics allow it. Another factor to consider is that engagement counts are influenced by audience exposure (reach and impressions), which can vary by post and over time. Our primary analyses did not explicitly model exposure as an offset to normalize engagement rates for delivery. As a result, posts with wider distribution of engagement may appear to perform better independent of their content.

Third, there are threats to the generalizability of these findings in other settings and online communities. The data reflect a single oncology nutrition platform and an engaged audience that was predominantly female and largely based in the United States, although demographic granularity was limited to platform-provided aggregates. These findings are limited to Facebook, and characteristics may not represent other populations or platforms with distinct content formats and algorithms. Replication across additional cancer centers,

XSL•FO
RenderX

community organizations, and other social networks (eg, Instagram, YouTube, and TikTok) would strengthen the external validity of these findings.

We acknowledge that the asymmetric time windows to assess preimplementation and postimplementation of this content strategy may also introduce seasonality and exposure differences into our results. Additionally, changing temporal, cultural, and algorithmic dynamics may have altered engagement patterns independent of content during the observation period. We partially addressed this by incorporating time period comparisons, but more granular controls (eg, calendar month, day of the week, or platform policy change markers) would better account for temporal and algorithmic shifts. Extending monitoring across multiple years in future studies may also help separate strategy effects from seasonality.

Future Directions

In the future, novel digital media and technological tools (eg, eye-tracking and web-monitoring) may be useful for determining how individuals evaluate the quality of cancer nutrition information and the features that promote online engagement with cancer nutrition information. This technology has been previously implemented in research to record users' viewing patterns on web-delivered information and provides a physiological measure linked to cognitive processing that may indicate the presence of active or passive engagement [39,40]. Data from eye-tracking can assess how users process online information such as social media content [41,42]. In the context of public health research, eye-tracking technology improves the understanding of how online health promotion content attracts attention from users [39,41,43].

The association between social media use and oncology nutrition knowledge is another topic that should be explored further, particularly because engaged social media users are more likely to be aware of other prevention measures, such as cancer screening [44,45]. Similarly, improving engagement with oncology nutrition content on social media may promote healthy dietary choices for cancer prevention and patients with cancer. The dietary needs of patients with cancer undergoing treatment such as chemotherapy are different from those of cancer survivors or caregivers aiming to reduce cancer risk. Therefore, oncology nutrition interventions through social media may benefit from addressing these individualized needs through targeted social media content. Importantly, future work should assess how exposure to online cancer information influences users' health beliefs, perceived cancer risk, trust in nutrition recommendations, and subsequent health behaviors.

It is also relevant to examine how these findings translate to other social media platforms with distinct audiences, content formats, and engagement algorithms such as Instagram, TikTok, and Snapchat. Each of these platforms reaches different demographic groups and prioritizes different types of content, particularly short-form video, visual storytelling, and creator-driven posts. Studying oncology nutrition messaging across these environments may help determine whether bilingual strategies, visual elements, or linguistic elements operate similarly on platforms that emphasize rapid, algorithm-driven content exposure. Understanding these mechanisms will be essential for identifying which platform features most effectively support evidence-based cancer prevention and nutrition education and for ensuring that such content reaches diverse communities that may not engage with Facebook.

Going forward, providers and staff can leverage cancer nutrition websites and social media to guide patients to high-quality cancer information by incorporating visual content such as photos, branding, and infographics, which improve social engagement, as well as including more analytical language and a neutral emotional tone within a multilanguage content strategy. It is important for providers to be aware of disparities in accessing information and to make cancer information available in different languages and at varying literacy levels. Another area with clinical relevance is to consider the specific dietary needs of patients with cancer undergoing treatment and provide targeted social media content to address these dietary needs both during treatment and across the cancer control continuum.

Conclusion

This study describes how CFYL, an online cancer nutrition community, disseminates information about nutrition education and cancer prevention. The results of this analysis demonstrate that implementing a dedicated bilingual, multiformat social media content strategy is associated with higher audience engagement with oncology nutrition posts. We describe strategies to increase engagement with online social media content disseminating nutrition-related cancer prevention and survivorship information. In models adjusted for the time period, posts that were bilingual, branded, and formatted performed better, and the language profile shifted toward more analytical, emotionally neutral text.

These associations were observed alongside concurrent changes in creative elements and content delivery, and exposure varied across posts. As such, causal inferences cannot be made. Even so, these findings suggest practical signals for designing bilingual, branded, multiformat content to support the equitable dissemination of evidence-based cancer nutrition information. Studying online language use may help cancer organizations strengthen the rationale for investing in bilingual social media cancer communication initiatives tailored to the language needs of the communities they serve. Future work should test language presentation more directly and evaluate how exposure to such content influences health beliefs and downstream behaviors across multiple social media platforms. Overall, these findings provide an initial context for communicating about nutrition with patients with cancer, survivors, and caregivers on Facebook and other social media platforms. These results are a step toward implementing more supportive interventions for the quality dissemination of cancer nutrition information on social media.

XSL•FO
RenderX

Acknowledgments

The authors acknowledge the involvement of the Cook for Your Life community in making this research possible. Generative AI was not used to prepare any portion of the manuscript.

Funding

The authors declared no financial support was received for this work. Dr Warner was supported, in part, by the National Cancer Institute under award T32CA078447. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: EH, SB-S, ER-S, HG, ELW
Data curation: SB-S, ER-S, HG
Formal analysis: EH, ER-S, ELW
Investigation: EH, SB-S, ER-S, HG, ELW
Methodology: EH, SB-S, ER-S, HG, ELW
Project administration: HG, ELW
Resources: HG, ELW
Software: HG, ELW
Supervision: HG, ELW
Validation: EH, SB-S, ER-S, HG, ELW
Visualization: EH, SB-S, ER-S, HG, ELW
Writing – original draft: EH, SB-S, ER-S, HG, ELW
Writing – review and editing: EH, SB-S, ER-S, HG, ELW

Conflicts of Interest

Dr Warner previously received grant funding as a co-investigator through Merck to study HPV vaccination for cancer prevention and online social media discourse about HPV vaccination by geography.

Multimedia Appendix 1
Qualitative codebook.
[DOCX File, 15 KB - cancer_v12i1e69886_app1.docx]

References

1. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917. [doi: 10.2196/17917] [Medline: 33978589]
2. Social media “likes” healthcare: from marketing to social business. : PricewaterhouseCoopers LLP (PwC); 2012 URL: https://www.pwc.com/us/en/health-industries/health-research-institute/publications/pdf/health-care-social-media-report.pdf [accessed 2023-04-10]
3. Chou WS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009 Nov 27;11(4):e48. [doi: 10.2196/jmir.1249] [Medline: 19945947]
4. Chen J, Hou X, Zhao W. Research on the model of consumer health information seeking behavior via social media. Int J Commun Netw Syst Sci 2016;09(08):326-337. [doi: 10.4236/ijcns.2016.98029]
5. Use of smartphones and social media is common across most emerging economies. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/2019/03/07/use-of-smartphones-and-social-media-is-common-across-most-emerging-economies/ [accessed 2023-11-26]
6. Della LJ, Eroglu D, Bernhardt JM, Edgerton E, Nall J. Looking to the future of new media in health marketing: deriving propositions based on traditional theories. Health Mark Q 2008;25(1-2):147-174. [doi: 10.1080/07359680802126210] [Medline: 18935883]
7. Thackeray R, Neiger BL, Hanson CL, McKenzie JF. Enhancing promotional strategies within social marketing programs: use of Web 2.0 social media. Health Promot Pract 2008 Oct;9(4):338-343. [doi: 10.1177/1524839908325335] [Medline: 18936268]
8. CDC’s social media channels. Centers for Disease Control and Prevention. 2023. URL: https://www.cdc.gov/digital-social-media-tools/Social-Media-Channels.html [accessed 2023-04-10]

XSL•FO
RenderX

9. Norman CD, McIntosh S, Selby P, Eysenbach G. Web-assisted tobacco interventions: empowering change in the global fight for the public's (e)Health. J Med Internet Res 2008 Nov 25;10(5):e48. [doi: 10.2196/jmir.1171] [Medline: 19033147]
10. Wangberg SC, Andreassen HK, Prokosch HU, Santana SMV, Sørensen T, Chronaki CE. Relations between internet use, socio-economic status (SES), social support and subjective health. Health Promot Int 2008 Mar;23(1):70-77. [doi: 10.1093/heapro/dam039] [Medline: 18083686]
11. Dooley JA, Jones SC, Iverson D. Using Web 2.0 for health promotion and social marketing efforts: lessons learned from Web 2.0 experts. Health Mark Q 2014;31(2):178-196. [doi: 10.1080/07359683.2014.907204] [Medline: 24878406]
12. Schleyer T, Spallek H, Butler BS, et al. Facebook for scientists: requirements and services for optimizing how scientific collaborations are established. J Med Internet Res 2008 Aug 13;10(3):e24. [doi: 10.2196/jmir.1047] [Medline: 18701421]
13. Bedrick SD, Sittig DF. A scientific collaboration tool built on the facebook platform. AMIA Annu Symp Proc 2008 Nov 6;2008:41-45. [Medline: 18999247]
14. Lewis K, Kaufman J, Gonzalez M, Wimmer A, Christakis NA. Tastes, ties, and time: a new social network dataset using Facebook.com. Soc Networks 2008 Oct;30(4):330-342. [doi: 10.1016/j.socnet.2008.07.002]
15. Kent EE, Rowland JH, Northouse L, et al. Caring for caregivers and patients: research and clinical priorities for informal cancer caregiving. Cancer 2016 Jul 1;122(13):1987-1995. [doi: 10.1002/cncr.29939] [Medline: 26991807]
16. Khanna PM. Icyou: how social media is the new resource for online health information. Medscape J Med 2008 May 9;10(5):113. [Medline: 18596952]
17. Gaysynsky A, Romansky-Poulin K, Arpadi S. "My YAP Family": analysis of a Facebook group for young adults living with HIV. AIDS Behav 2015 Jun;19(6):947-962. [doi: 10.1007/s10461-014-0887-8] [Medline: 25186783]
18. Idriss SZ, Kvedar JC, Watson AJ. The role of online support communities: benefits of expanded social networks to patients with psoriasis. Arch Dermatol 2009 Jan;145(1):46-51. [doi: 10.1001/archdermatol.2008.529] [Medline: 19153342]
19. Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff (Millwood) 2009 Mar;28(2):361-368. [doi: 10.1377/hlthaff.28.2.361]
20. Khan SA, McFarlane DJ, Li J, et al. Healthy Harlem: empowering health consumers through social networking, tailoring and Web 2.0 technologies. AMIA Annu Symp Proc 2007 Oct 11:1007. [Medline: 18694106]
21. Block G, Sternfeld B, Block CH, et al. Development of Alive! (a lifestyle intervention via email), and its effect on health-related quality of life, presenteeism, and other behavioral outcomes: randomized controlled trial. J Med Internet Res 2008 Nov 19;10(4):e43. [doi: 10.2196/jmir.1112] [Medline: 19019818]
22. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008 Aug 25;10(3):e22. [doi: 10.2196/jmir.1030] [Medline: 18725354]
23. Kim SJ, Marsch LA, Brunette MF, Dallery J. Harnessing Facebook for smoking reduction and cessation interventions: Facebook user engagement and social support predict smoking reduction. J Med Internet Res 2017 May 23;19(5):e168. [doi: 10.2196/jmir.6681] [Medline: 28536096]
24. Tobey LN, Manore MM. Social media and nutrition education: the Food Hero experience. J Nutr Educ Behav 2014;46(2):128-133. [doi: 10.1016/j.jneb.2013.09.013] [Medline: 24220043]
25. The NCI Community Oncology Research Program (NCORP). URL: https://ncorp.cancer.gov/ [accessed 2023-05-05]
26. Social media use in 2021. : Pew Research Center; 2021 URL: https://www.pewresearch.org/wp-content/uploads/sites/20/2021/04/PI_2021.04.07_Social-Media-Use_FINAL.pdf [accessed 2026-07-14]
27. Post engagement in Meta Ads. Meta. URL: https://www.facebook.com/business/help/735720159834389 [accessed 2023-05-05]
28. About us. Cook for Your Life. URL: https://www.cookforyourlife.org/about-us/ [accessed 2026-07-14]
29. Merten JW, Roberts KJ, King JL, McKenzie LB. Pinterest homemade sunscreens: a recipe for sunburn. Health Commun 2020 Aug;35(9):1123-1128. [doi: 10.1080/10410236.2019.1616442] [Medline: 31111735]
30. Pennebaker JW, Boyd RL, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015. : The University of Texas at Austin; 2015 URL: https://www.liwc.app/static/documents/LIWC2015%20Manual%20-%20Development%20and%20Psychometrics.pdf [accessed 2026-07-14]
31. Hayes RA, Carr CT, Wohn DY. It's the audience: differences in social support across social media. Soc Media Soc 2016 Oct;2(4). [doi: 10.1177/2056305116678894]
32. Josephson S, Kelly J, Smith K, editors. Handbook of Visual Communication: Theory, Methods, and Media, 2nd edition: Routledge; 2020.
33. Heley K, Gaysynsky A, King AJ. Missing the bigger picture: the need for more research on visual health misinformation. Sci Commun 2022 Aug;44(4):514-527. [doi: 10.1177/10755470221113833] [Medline: 36082150]
34. Strekalova YA, Krieger JL. A picture really is worth a thousand words: public engagement with the National Cancer Institute on social media. J Cancer Educ 2017 Mar;32(1):155-157. [doi: 10.1007/s13187-015-0901-5] [Medline: 26351003]
35. Andrade EL, Evans WD, Barrett N, Edberg MC, Cleary SD. Strategies to increase Latino immigrant youth engagement in health promotion using social media: mixed-methods study. JMIR Public Health Surveill 2018 Dec 19;4(4):e71. [doi: 10.2196/publichealth.9332] [Medline: 30567689]

XSL•FO
RenderX

36. Rivera YM, Moran MB, Thrul J, Joshu C, Smith KC. Contextualizing engagement with health information on Facebook: using the social media content and context elicitation method. J Med Internet Res 2022 Mar 4;24(3):e25243. [doi: 10.2196/25243] [Medline: 35254266]
37. Warner EL, Ellington L, Kirchhoff AC, Cloyes KG. Acquisition of social support and linguistic characteristics of social media posts about young adult cancer. J Adolesc Young Adult Oncol 2018 Apr;7(2):196-203. [doi: 10.1089/jayao.2017.0076] [Medline: 29099664]
38. Seo K. Meta-analysis on visual persuasion—does adding images to texts influence persuasion? Athens J Mass Media Commun 2020;6(3):177-190. [doi: 10.30958/ajmmc.6-3-3]
39. Fox RJ, Krugman DM, Fletcher JE, Fischer PM. Adolescents' attention to beer and cigarette print ads and associated product warnings. J Advert 1998;27(3):57-68. [doi: 10.1080/00913367.1998.10673563]
40. Alley S, Jennings C, Persaud N, Plotnikoff RC, Horsley M, Vandelanotte C. Do personally tailored videos in a web-based physical activity intervention lead to higher attention and recall?—an eye-tracking study. Front Public Health 2014;2:13. [doi: 10.3389/fpubh.2014.00013] [Medline: 24575398]
41. Chou WYS, Trivedi N, Peterson E, Gaysynsky A, Krakow M, Vraga E. How do social media users process cancer prevention messages on Facebook? An eye-tracking study. Patient Educ Couns 2020 Jun;103(6):1161-1167. [doi: 10.1016/j.pec.2020.01.013] [Medline: 32044193]
42. Kim SC, Vraga EK, Cook J. An eye tracking approach to understanding misinformation and correction strategies on social media: the mediating role of attention and credibility to reduce HPV vaccine misperceptions. Health Commun 2021 Nov;36(13):1687-1696. [doi: 10.1080/10410236.2020.1787933] [Medline: 32633151]
43. Liu J, Phua J, Krugman D, Xu L, Nowak G, Popova L. Do young adults attend to health warnings in the first IQOS advertisement in the U.S.? An eye-tracking approach. Nicotine Tob Res 2021 May 4;23(5):815-822. [doi: 10.1093/ntr/ntaa243] [Medline: 33245339]
44. Qin L, Zhang X, Wu A, et al. Association between social media use and cancer screening awareness and behavior for people without a cancer diagnosis: matched cohort study. J Med Internet Res 2021 Aug 27;23(8):e26395. [doi: 10.2196/26395] [Medline: 34448708]
45. Nour M, Yeung SH, Partridge S, Allman-Farinelli M. A narrative review of social media and game-based nutrition interventions targeted at young adults. J Acad Nutr Diet 2017 May;117(5):735-752.e10. [doi: 10.1016/j.jand.2016.12.014] [Medline: 28238894]

Abbreviations

CFYL: Cook For Your Life
FHCC: Fred Hutchinson Cancer Center
LIWC: Linguistic Inquiry and Word Count
LOL: laughing out loud
OMG: oh my god
OR: odds ratio

Edited by M Balcarras; submitted 10.Dec.2024; peer-reviewed by D Patel, P Corr; revised version received 02.Jul.2026; accepted 03.Jul.2026; published 14.Aug.2026.

Please cite as:
Harrison E, Buzali-Soto S, Rillamas-Sun E, Greenlee H, Warner EL
Understanding Social Media Strategies to Improve Engagement in and Equitable Dissemination of Online Cancer Nutrition Information: Cross-Sectional Study
JMIR Cancer 2026;12:e69886
URL: https://cancer.jmir.org/2026/1/e69886
doi:10.2196/69886

© Ellie Harrison, Sara Buzali-Soto, Eileen Rillamas-Sun, Heather Greenlee, Echo L Warner. Originally published in JMIR Cancer (https://cancer.jmir.org), 14.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient and Provider Experiences With Precision Oncology: Qualitative Descriptive Study at the Department of Veterans Affairs

Daniel Becker[1,2], MS, MD; Kenneth Csehak[2], MD; Alexander Barbaro[3], MD; Christian Miller[2], MD; Antoinette Vo[1], MA; Stefanie Roman[1], BA; Danil Makarov[1,2], MD; Scott Sherman[1,2], MD, MPH; Allison Squires[2], PhD

[1]VA NYHHS (Veterans Affairs New York Harbor Healthcare System), New York, NY, United States

[2]NYU Langone, 550 First Avenue, New York, NY, United States

[3]Northwell Health, New York, NY, United States

Corresponding Author:
Christian Miller, MD
NYU Langone, 550 First Avenue, New York, NY, United States

Abstract

Background: Precision oncology (PO) improves and extends the lives of patients living with cancer, but multiple studies have documented its underuse in practice. Specifically, studies note a significant lack of PO use within the Veterans Affairs (VA) medical system. A paucity of implementation of PO in oncologic practice poses a significant barrier to providing the most up-to-date guideline-based care.

Objective: While several studies have explored determinants of PO use, we sought to contribute to the body of knowledge by additionally focusing on the unique perspectives of patients, as well as conducting a comprehensive study within the VA medical system, the United States' largest single-payer health care system. We conducted interviews with both patients and providers at multiple VA sites to identify and characterize barriers and facilitators of PO use in clinical care.

Methods: Using a qualitative descriptive approach, we conducted semistructured interviews with 17 patients with cancer and 16 oncology providers recruited from multiple VA sites. Cancer types included prostate, gastrointestinal, and lung. Data were analyzed via a team-based coding approach using directed content analysis. Data were coded and then aggregated into themes and mapped to the Theoretical Domains Framework (TDF) and Behavior Change Wheel sources of behavior (Capability, Opportunity, and Motivation) based on the consensus of the study team.

Results: The patient sample consisted of 17 all-male veterans seen at VA oncology clinics in 2022. Participants predominantly self-identified as White (n=9, 52.9%) or Black (n=6, 35.3%), and the majority (n=11, 64.7%) held a high school degree or a higher level of education. The provider sample consisted of 16 physicians, all of whom held MD degrees and practiced oncology. The provider sample represented 6 states, was 50% (8/16) female, and participants averaged 14 years in their current position. The overarching theme was the "Precision Oncology Feedback Loop," which captured the essence of the complex processes involved in facilitating PO care in the VA system. The TDF and Behavior Change Wheel helped categorize findings to identify where issues in the feedback loop could facilitate or generate barriers to care.

Conclusions: Our findings expand on the current literature by highlighting both patient and provider experiences across key TDF domains (Environmental Context and Resources, Knowledge, Memory, and Attention). The conceptual model produced by the analysis illustrates the complexities associated with the implementation. Our findings support the design of multilevel interventions that target increased knowledge or education, improved workflow, and ease of communication to enhance PO delivery.

(JMIR Cancer 2026;12:e84858) doi:10.2196/84858

KEYWORDS

oncology; precision oncology; health services; qualitative; health equity

Introduction

Precision oncology (PO), the identification of tumor-specific abnormalities as targets for personalized therapy, has revolutionized cancer care over the last several decades by extending survival and minimizing treatment-related toxicities [1]. Despite the benefits established in clinical trials, extensive studies by our group and others document that up to 50% of patients do not receive guideline-recommended PO treatment. For example, molecular testing of colon cancer biopsy tissue for clinically targetable alterations in patients with advanced colon cancer has been recommended by expert consensus guidelines since 2010 [2]. Studies of this population in the Veterans Affairs (VA) health system established that such

XSL•FO
RenderX

testing was completed in only 28% of eligible patients with advanced colon cancer by the year 2015 [3]. This hinders the ability of clinicians to use the cancer mutational profile to determine therapy, which is standard care for patients with advanced colon cancer. Similarly, among patients with non-VA colorectal cancer included in the Surveillance Epidemiology and End Results database, molecular testing was completed for only 31% of eligible patients [4,5].

Consistent with ecological models of health behaviors, barriers and facilitators to PO exist at many levels (eg, patient, provider, and system) and steps in the process of care [6,7]. At the patient level, older patients and those with Medicaid insurance are less likely to receive PO [3,8]. Furthermore, Black and Latino patients have been consistently underrepresented in the practice and study of PO [9]. Furthermore, at the provider level, quantitative and chart review data suggest that providers do not always act upon PO testing results [10]. Facility and systems-level barriers have also been identified in multiple studies, including lower testing rates in lower-volume cancer centers, difficulty coordinating testing, financial barriers, and lack of clear protocols for test indications and use [10,11].

A recent review of qualitative research on barriers and facilitators to PO suggests that barriers often involve insurance or reimbursement as well as a lack of certainty with regard to clinical utility, while facilitators include the incorporation of PO into clinical guidelines. Notably, this review included only 5 qualitative studies, with very limited representation of oncology clinicians and patients with cancer [12]. One study included no clinicians, another included only 1 clinician, and a third included only 3 clinicians [13-15]. The final 2 studies included 10 and 14 clinicians, respectively, but did not include clinicians in the VA [16,17]. Only one study specified the inclusion of patients and included 16 [16]. While this suggests that there is active investigation in the space, there are relatively few studies, with small sample sizes, and no prior VA representation.

The Veterans Health Administration is the largest integrated health system in the United States, and the VA National Oncology Program cares for approximately 43,000 veterans newly diagnosed with cancer annually [18]. We designed a qualitative study of patients with cancer and providers based on the Theoretical Domains Framework (TDF) [19] to explore barriers and facilitators to PO care in this population. We know that molecular testing, essential for the standard of care in cancer treatment, is being underused. The reasons for the underuse of molecular testing remain unclear, and we are conducting a qualitative investigation to facilitate an in-depth exploration of the patient and provider experiences of PO and identify areas in which interventions to increase guideline-concordant PO can be effectively used.

Methods

Design

We used a qualitative descriptive approach with directed content analysis to structure the study. The goal of qualitative descriptive studies is to describe the experiences of those participating in the health-related behavior or process under consideration [20]. When used for health services research, this approach can identify and differentiate common and unique factors experienced by patients and clinicians when engaging with health services.

Theoretical Underpinning

The TDF is a widely used behavior change framework, consisting of 14 domains that are used in the identification of determinants of change [21]. TDF categories are then mapped to the Behavior Change Wheel (BCW) components of Capability, Opportunity, and Motivation. The BCW is an evidence-based framework for understanding drivers of behavior and translating that understanding into actionable change [22]. Both TDF and BCW were chosen for this analysis over other frameworks, such as the Consolidated Framework for Implementation Research. The TDF can be more useful for the design of qualitative research interview guides by specifying discrete domains, which are applicable at the individual level as well as at the systems level. Our focus was predominantly on the behaviors and experiences of the patients and providers, which is best suited to TDF and BCW over the more process-oriented Consolidated Framework for Implementation Research.

Reflexivity

The research team comprised 4 researchers: one senior qualitative PhD researcher who is a nurse (AS) and three MD researchers board-certified in internal medicine (KC, DB, and AB). Of the MD researchers, one was a practicing oncologist, and two were oncology fellows in the third year of their training. The nurse researcher had no oncology background and brought methodological expertise. All identify as having White European heritage.

Sample and Setting

Adult patients seen in oncology clinics between January and June 2022 at participating VA medical centers (VAMCs) in Pittsburgh and New York with a histologically confirmed prostate, gastrointestinal, or lung cancer were included in the study. We chose these 3 cancers as they together represent more than 60% of all patients with cancer within the VA, and PO is well established in all 3 cancers. We chose common cancers to understand practice patterns in patients with relatively common oncologic diseases. Patients with other cancers were excluded. The US Department of Veterans Affairs includes both VAMCs with inpatient and outpatient care as well as VA Community–Based Outpatient Care Centers with outpatient care only. Our study was conducted only at VAMCs. The VA is a single-payer health care system that serves 9 million American veterans and veteran family members.

Recruitment

Prospective patient participants were recruited by either letter or flyer. A letter was mailed to patients at the Veterans Affairs–New York Harbor Health Care System (VA-NYHHS), whereas a flyer was used at the Pittsburgh VAMC. At the Pittsburgh VAMC, flyers were placed in highly visible locations in the oncology clinics, such as the front desk at registration and in patient examination rooms. Providers were encouraged

XSL•FO
RenderX

to inform patients that the study was available. The recruitment of provider participants was email-based. Emails were sent to oncology physicians at the VA-NYHHS, Greater Los Angeles Area (GLA), Pittsburgh, Michigan, and Seattle VAMCs. Informed consent was obtained from all participants, with all study methods approved by the local VA institutional review board (IRB). Patients were eligible for this study if they were receiving medical care at one of the participating VAMCs and were receiving treatment for prostate, lung, or gastrointestinal cancer. Only English-speaking patients were eligible, as US military enlistment requires English proficiency. Providers were eligible if they provided care for patients with cancer in one of the participating VAMCs.

Data Collection

The semistructured interview guide was constructed through an expert review process by a multidisciplinary team of oncologists, qualitative health researchers, and patient representatives. All 14 TDF domains were represented by at least 1 question in the interview tool. Three patient and three provider interviews, not included in the sample population, were used to refine the interview guide and construct a draft code list. This was done to provide test samples for assessing the feasibility of the interview guide.

Two interviewers then conducted semistructured interviews with patients and providers that lasted between 30 and 60 minutes. All interviews were conducted via audio using patients' telephone numbers through Microsoft Teams. The interviews were transcribed using Microsoft Teams' automated transcription feature, with verification by study personnel. Interviews continued until data saturation was achieved.

Data Analysis

Directed content analysis using the TDF categories guided the analytical process. Directed content analysis allows for codes dictated by a preanalysis identified list, as well as purely descriptive codes, to emerge during the process [23].

AS provided coding training to the physician team and reviewed findings as they emerged. A team-based coding approach structured the consensus-based process [24]. Based on guidance from Cascio et al [24], we drew from principles of team-based coding to generate intercoder consensus throughout the analytic process. The consensus process was documented throughout the study to record the process of data review, idea generation, and analytical technique. The consensus process was conducted in person through meetings with the entire coding group.

Atlas TI.web (version 9.24.3; Lumivero) was used to code interviews. Each interview, whether from a patient or clinician, was coded independently by 3 research staff members (DB, KC, and AB). Initial codes came from the theoretical framework to start the process, and others emerged iteratively during the coding process. All team members met after 3 interviews were coded to harmonize code names and conceptual definitions. The codes were then compared and reviewed to achieve consensus on any coded data about which study staff disagreed. Disagreements regarding code application were discussed among the study team until a consensus was reached. A common codebook was then generated, and the team members checked in periodically to continue coding harmonization throughout the process.

The team then met to finalize categories and themes that emerged independently from the data as well as to evaluate how coding aligned with the TDF theoretical framework. Similar code groups were aggregated and mapped to related BCW sources of behavior (Capability, Opportunity, and Motivation) based on the consensus of the study team.

Ethical Considerations

A determination of minimal risk and approval was granted by the VA-NYHHS IRB for our protocol (1605219‐1). We received subsequent approval from the Research and Development Committee and the Research Safety and Security Subcommittee. Consent forms were reviewed and approved by each committee and subcommittee. Informed consent was obtained from each participant in accordance with IRB-approved documents. Privacy and confidentiality were protected in accordance with data security specifications outlined in the IRB-approved protocol and were further reviewed by the institutional data security officer. Participant compensation consisted of US $25 gift cards for each participant.

Results

Overview

At the VA-NYHHS, letters were sent to 136 patients, of whom 10 (7.4%) enrolled. At the Pittsburgh VAMC, flyers were posted for patients attending the oncology clinic. Of these patients, 7 (1.8%) enrolled, from approximately 400 patients who passed through the clinic. A total of 17 patients and 16 providers (n=33) were interviewed.

Patient participants were predominantly aged 70 to 80 years, entirely male, and mostly White. Twenty-four providers received the email, of whom 16 (67%) participated. Of the 16 providers who were interviewed, 1 provider requested to reply to demographic questions by email but did not subsequently respond. Providers were mostly New York–based, predominantly aged 40 to 50 or 50 to 70 years, and predominantly White. Providers were mainly MDs. Patient and provider characteristics are further outlined in Tables 1 and 2 and include self-reported age, gender, race or ethnicity, as well as the state of the site from which the participant was recruited. For providers, additional information was collected regarding the number of years in position, type of professional degree, and leadership role status.

XSL•FO
RenderX

Table . Provider characteristics collected for the study of barriers and facilitators to precision oncology.

Characteristic	Participants
Number of providers	
Completed interviews	16
State site, n (%)	
California	1 (6.3)
Michigan	2 (12.5)
North Carolina	1 (6.3)
New York	8 (50)
Pennsylvania	2 (12.5)
Washington	1 (6.2)
Not specified	1 (6.2)
Age range (y)	
30 - 40	1 (6.3)
40 - 50	5 (31.2)
50 - 60	3 (18.8)
60 - 70	6 (37.5)
Unknown	1 (6.2)
Gender, n (%)	
Male	7 (43.7)
Female	8 (50)
Other	1 (6.3)
Race, n (%)	
White	10 (62.5)
Asian or East Indian	4 (25)
Other	1 (6.3)
Unknown	1 (6.2)
Ethnicity, n (%)	
Hispanic	0 (0)
Not Hispanic	14 (87.5)
Other	1 (6.3)
Unknown	1 (6.2)
Professional degrees, n (%)	
MD	11 (68.7)
MD and MDA	1 (6.3)
MD and MBBS	2 (12.5)
MD and PhD	1 (6.3)
Unknown	1 (6.2)
Years in current position, mean (range)	13.6 (1 - 29)
Leadership role, n (%)	
Yes	10 (62.5)
No	5 (31.2)
Unknown	1 (6.3)

XSL•FO
RenderX

Table . Patient characteristics collected for the study of barriers and facilitators to precision oncology.

Characteristic	Participants, n (%)
Patients	
Completed interviews	17 (100)
State site	
New York	10 (58.8)
Pennsylvania	7 (41.2)
Age range (y)	
50 - 60	1 (5.9)
60 - 70	2 (11.8)
70 - 80	12 (70.6)
80 - 90	2 (11.7)
Gender	
Male	17 (100)
Female	0 (0)
Other	0 (0)
Race	
Black	6 (35.3)
White	9 (52.9)
Native American	1 (5.9)
Other	1 (5.9)
Ethnicity	
Hispanic	1 (5.9)
Non-Hispanic	14 (82.4)
Unknown or none	2 (11.7)
Highest education	
Some high school	2 (11.8)
High school graduate	4 (23.5)
GED[a]	1 (5.9)
Some college	2 (11.8)
College graduate	5 (29.4)
Unknown	3 (17.6)

[a]GED: graduate equivalency degree.

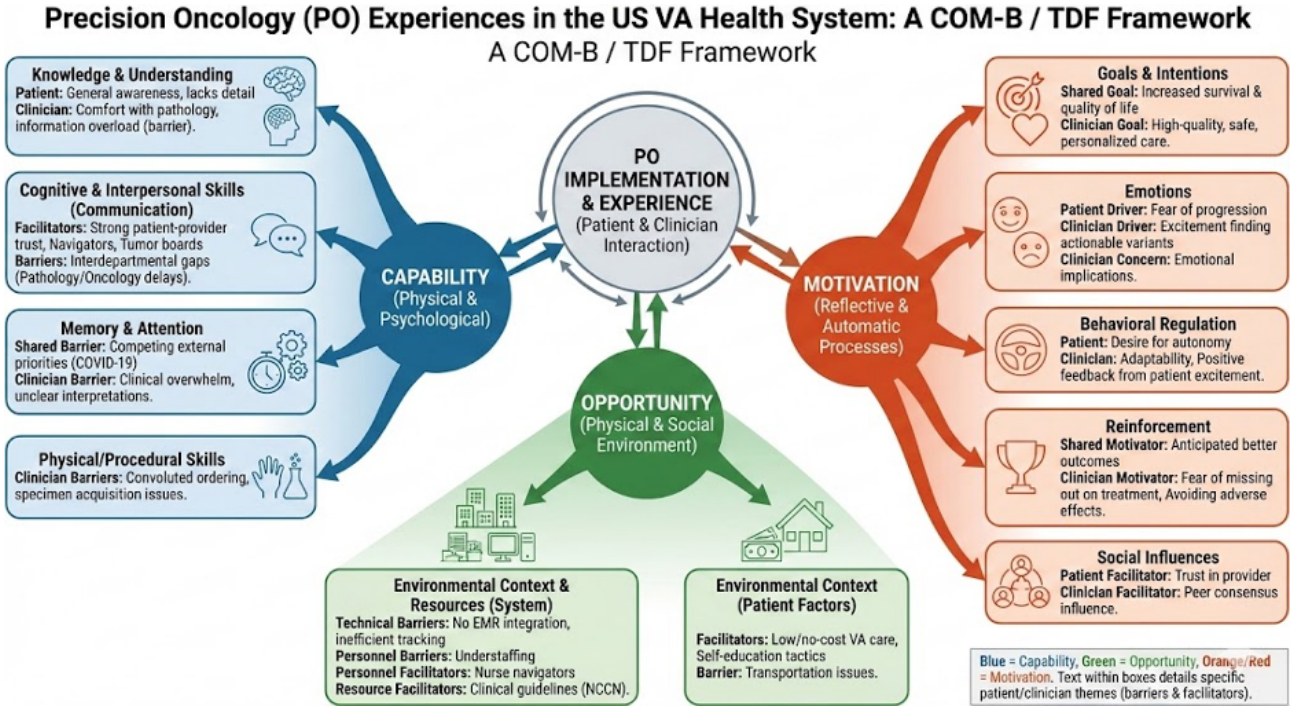
Overarching Theme: The Precision Oncology "Feedback Loop"

Our findings suggest that there is a cyclical feedback loop related to the delivery of PO. The feedback loop is driven by the opportunity for this kind of care, as dictated by the resources of the system where care is delivered. Satisfaction from the clinician and patient side is the result of technical process outcomes associated with the PO testing process. Motivation to engage with PO services is a product of clinician learning and success with the tool and the resulting success with patient treatments, which also generates happiness for the patient.

The following sections illustrate how the feedback loop occurs, with the TDF providing structure for the categories that shape it. Figure 1 illustrates the feedback loop of Capability, Opportunity, and Motivation on the part of both patients and providers, all of which contribute to patient and provider engagement leading to PO. Sections are organized by the framework's components.

XSL•FO
RenderX

Figure 1. Precision oncology feedback loop. EMR: electronic medical record; NCCN: National Comprehensive Cancer Network; TDF: Theoretical Domains Framework; VA: Veterans Affairs.



Behavioral Regulation (TDF)–Motivation (BCW)

In the TDF, behavioral regulation is defined as the components aimed at managing or changing objectively observed or measured actions, while motivation aligned with the BCW is defined as brain processes that energize and direct behavior [21,22].

In our analysis, statements that referred to self-moderation of actions on the part of patients and providers fell into this category. Patients reflected on the importance of a sense of motivation and autonomy in navigating their care, while providers identified motivation and adaptability as components of effective action. Furthermore, positive feedback emerged as providers sensed that patients were motivated to pursue testing. As one clinician described:

> *I think most of the guys get it pretty good and are excited to hear that there's something beyond standard chemotherapy that they've heard all the horror stories about. They're excited about it so as you move from standard chemotherapy into immunotherapy, now into gene therapy, people are happy to hear that there's more knowledge, more options, more thinking, we're getting somewhere with cancer treatment.* [Provider 13]

Cognitive and Interpersonal Skills Capability: The Impact of Communication

This category captures the way in which respondents felt communication and coordination between various individuals affected PO and cancer care. Three key motifs emerged, centered on the intersections of multilevel aspects of communication.

Interdepartmental Communication

Participants noted the issues that arise, even with an electronic health record, when trying to coordinate departments involved in diagnosis, treatment, and management of care. For example, several respondents identified issues with delayed next-generation sequencing (NGS) results due to gaps in communication between different medical specialties (eg, proceduralists, oncologists, and pathologists):

> *It would be nice if pathology had a way to reflexively send the specimen. I guess it's hard for them. They don't know the whole situation and you kind of need somebody to oversee it, but that again is just another frustrating thing.* [Provider 2]

Conversely, several respondents cited regular interdepartmental meetings, such as tumor boards, as helpful in reducing delays.

Interdisciplinary Communication

Interactions between providers and nurses, between providers and trainees, and between providers and administrators were all identified as key elements of PO care. Multiple providers also praised an individual coordinator responsible for PO implementation.

Patient-Provider Communication

Overall, respondents on both the patient and provider sides expressed that communication with their counterparts was an effective and important facilitator of PO. "The doctors I went to, I think, were extremely thorough in describing what was going and what the treatments were, and answering my questions" (patient 51). On the other hand, a lack of patient understanding was reported as a barrier on the part of clinicians.

XSL•FO
RenderX

Emotions and Motivation

Emotions and motivation statements reflected the interactions between how patients and clinicians highlighted the role of emotions in influencing PO practice. Unsurprisingly, most of the negative patient-reported emotions reflected fear of their cancer, its progression, and death. Providers cited concerns about the emotional and practical implications of detecting germline mutations.

Nonetheless, finding an actionable mutation was in itself a significant provider motivator:

> *I think that there's kind of a sense of happiness when we find one of those really high yield variants, when you find the EGFR or the ALK in lung cancer. You feel really happy sharing that information with patients and obviously I think they understand that when the provider is happy with something that that's probably good news for their cancer.* [Provider 5]

Thus, the balance of emotions and motivations encapsulates the inherent rewards and challenges of the PO treatment process for both clinicians and patients.

Environmental Context and Resources Shaping Opportunity

Overview

Both providers and patients talked extensively about how the care environment and its resources were critical for treatment opportunities. The overall environment and the provider's ability to navigate the resources to get patients their treatment were essential for balancing patient-level factors that influenced care choices. How they described these aspects of the environment fell into two subcategories of "system-centered" and "patient-centered" descriptions.

System: Technical

The organizational structure of the PO process functions was extensively discussed by participants. This included challenges with ordering forms, a lack of dedicated communication channels between pathology and the ordering provider, and inconsistencies in how results are eventually uploaded and displayed. For example, one respondent noted:

> *We've had to overcome some issues because there is no ordering method within the electronic medical record […] So right now we're doing this by emails. So essentially, we identify patients in the clinic that need sequencing, and I have a coordinator that is in clinic with me each time I'm in clinic and she will send specimens for sequencing. It's kind of inefficient.* [Provider 7]

Among facilitators, patient consent and electronic ordering processes were cited as technical aspects of the system that functioned well and did not delay care delivery.

System: Personnel

Respondents cited understaffed provider departments as barriers and navigators and administrative assistants as facilitators. Noting the effectiveness of a nurse navigator working in the system, a provider offered an example of facilitated communication flows:

> *The nurse navigator does usually read the report and they will scan and send me an alert, an email, and say "this is someone who's got an actionable mutation" just to make sure that I've read the report.* [Provider 12]

Understaffing reflected vacancies, hiring processes, and other factors associated with operating within a government bureaucracy.

System: Practice Resources

Clinical and institutional guidelines were noted to help providers recognize PO testing indications. These were seen as important for quickly discerning treatment plans and managing patient responses to them. As one provider stated:

> *So my overall goals do align with guidelines, such as NCCN guidelines. So they just hopefully align with standard of care.* [Provider 4]

For some providers, these guidelines served as educational tools to update their practice.

Patient Factors

Logistics, financial challenges, and how patients were educated about their treatments and options were the patient-centered factors that influenced the progression of PO care. Transportation issues were the most frequently cited barrier to care. Many veteran participants were economically vulnerable and did not always have access to reliable public transportation as an alternative. For example, one patient participant stated "sometimes I'd have an appointment that was too early and I'd request transportation and it was hard" (patient 058).

Despite economic vulnerability, the low or no-cost care provided to VA patients was viewed as a significant facilitator that helped ensure veterans were protected from the financial toxicity that can be associated with cancer care outside the VA system. Finally, patients reported a variety of strategies to educate themselves about their diseases and treatments. While some patients mentioned receiving educational materials from their doctor or other VA sources, many more reported using external resources, such as the internet or family members, to learn about their disease and treatment options. One patient stated:

> *My wife she goes online … that's where she learned about the endocarditis because she had to deal with it too. And she's had sepsis and cancer and she goes to all of the medical sites.* [Patient 056]

Educating themselves was a clear goal and motivator for patients, which highlights the value of patient engagement in the PO process. They wanted to understand the causes behind their cancer, including potential germline mutations, and to be informed about their testing results and treatment plans.

Goals: Motivation

Providers had clear goals for PO care that motivated them when working with patients throughout their cancer treatment. They formed into 2 categories focused on patient outcomes and quality of care. Both providers and patients shared a mutual outcome

XSL•FO
RenderX

goal of an increase in the length of life or a cure to their disease, as well as their desire to maintain a good quality of life. As one provider stated: "The goals of patient care would be to help people be on the best therapy with the lowest side effects that keeps them alive the longest" (Provider 010). Many provider respondents expressed the importance of the highest quality care, including safety, personalized care, and the ability to treat actionable mutations. They viewed this as inherently intertwined with a good patient outcome.

Knowledge: Capability

Overview

Patients and providers reflected on the role of knowledge in the implementation of PO. Among patients, knowledge of their disease and how it is treated was often broad, with a lack of information constituting a barrier. Conversely, among providers, a deep understanding of both the disease and its treatment facilitated greater engagement in PO.

Patient Disease

Although many patients expressed an awareness of genes and their contribution to cancer, more statements were identified in which patients specifically reported a lack of knowledge or understanding about the cause of their disease and its treatment as stated by a patient:

> *How it forms is I'm not sure. It's something to do with your cells.* [Patient 57]

Patient Treatment

Patients frequently expressed knowledge of their treatment in broad, overarching terms, with occasional confusion between modalities of treatment. A lack of familiarity with terms such as "precision medicine" and "targeted treatment" was prevalent. One patient stated:

> *I know that everyone has DNA and it's specific to each person and genes ... I've heard of gene therapy but I'm not sure if it apples to cancer.* [Patient 055]

Provider Understanding of Disease and Treatment

Providers expressed comfort with knowledge of the role of mutations in cancer pathogenesis and treatment. They also, however, pointed out the quickly changing PO landscape as a barrier and guidelines and peer-reviewed sources as facilitators. NGS reports themselves were felt to be "a little bit better in terms of outlining what's pathogenic and what's not" (provider 11).

Memory and Attention: Capability

Among both patient and provider participants, external distractors emerged as barriers to PO, with responses generally categorized as prioritization either within the clinic or outside the clinic.

Prioritization Within the Oncology Clinic

Some providers expressed difficulty with practicing PO in part because of competing oncologic concerns, such as complex treatment and comorbidity management. Furthermore, providers noted feeling overwhelmed by variable interpretations of PO testing, unclear clinical implications of certain mutations, and excessive extraneous information. Providers noted facilitators, including tumor boards and reflex testing.

Prioritization Outside the Oncology Clinic

For patients and providers, the COVID-19 pandemic and associated personal and professional distractions were noted:

> *I mean the hardest part is just always just doing too many things at once. Like sometimes Path will just call me and I'll be there in the middle of clinic. One time I was at softball practice for my 7-year-old and like they're like, "Oh yeah, we didn't get enough tissue" and my head is in a totally different place, I can't write that down.* [Provider 9]

Patients also noted the need to manage other conditions and social issues.

Skills: Capability

For the second least-frequently encountered domain, the reported "skills" involved in PO include the abilities to obtain tumor tissue, select the appropriate kit, and properly complete the ordering process:

> *They need to know the difference between the tissue sample kits and the liquid biopsy kits. And they need to make sure that the kits aren't expired.* [Provider 1]

The majority of difficulties reported were in the latter category, with convoluted and unstandardized ordering processes sometimes creating difficulties with PO.

Reinforcement: Motivation

As in the Goals domain, providers and patients noted an anticipated increase in the length or quality of life as a perceived positive consequence of PO, and providers feared missing this opportunity.

Treatment Outcomes

With regard to treatment outcomes, patients' stated goals typically pertained to survival, with one patient stating: "Well basically it's to stay alive" (Patient 51). Providers expressed similar goals as well as concerns about not meeting the goals. One provider stated:

> *I'm thinking about targeted therapy so it's a whole 15 to 17% of people you can deprive of potentially life prolonging medication. And the impact that you're doing by not testing it, can you take it? Being an oncologist, treating oncology patients, can you take it with you every night and go to sleep? I cannot, so test for it.* [Provider 3]

The negative consequences of treatment adverse effects were mentioned frequently in broad terms, but PO was seen as a way to potentially avoid these consequences.

Quality Outcomes

Respondents cited the benefits of selecting a more effective treatment as a positive outcome of PO and the detriment from the failure of PO.

XSL•FO
RenderX

Social Influences: Motivation

Sources of motivation for patients and providers came from similar and distinct sources. This relationship was always dyadic in nature.

Patient-Provider

Patients frequently mentioned their providers as their principal source of advice on treatment-related issues. For example, one respondent stated, "Yes, yes. Total trust in my doctor" (patient 073). Comfort with and deference to providers emerged as common themes, whereas instances of conflict between patients and providers were mentioned only twice.

Patient-Other

Additional relevant social influences on patients come from their community, including family and friends.

Provider-Colleague

The most important social influence between providers appeared to be their peers' consensus on the indications and use of PO:

> *I think we're constantly learning from each other. You know, we're seeing our colleagues order testing in this setting and so we're going to order.* [Provider 5]

Technical Challenges: Opportunity

How providers were able to operationalize PO treatments was the result of internal processes. These could facilitate or inhibit timely care delivery.

Ordering Process

This was the most frequently mentioned aspect of technical challenges. Providers noted frequent difficulties in completing the orders required to send testing and reported frequent breakdowns in communication with the pathology department related to initial processing. One respondent was quoted as saying:

> *Trying to get the stuff ordered is more difficult than just ordering standard tests like immunohistochemistry or flow cytometry. We have to jump through a bunch of hoops for that and also the retrieval of the information. Once it's done it requires a few phone calls, where's my stuff?* [Provider 13]

Specimen Acquisition or Processing

The expected procedural difficulties in obtaining a specimen were mentioned by several providers, and it was occasionally noted that tissue was available but at another institution. Knowing where kits were available and arranging for transport were also potential failure points, especially when dealing with peripheral blood NGS, obtained in the clinic.

Resulting Process

The principal issues with the resulting process were the long time it takes to generate a usable result and the inability to see the result once testing was completed.

Conceptual Model

To illustrate the intersections of our results, we offer a conceptual model in Figure 1. Under the auspices of our overarching theme of "The Precision Oncology Feedback Loop," the conceptual model illustrates the contribution of each of the areas of the BCW to the patient and provider interaction that ultimately influences PO implementation. Capability, opportunity, and motivation all influence and variously reinforce or negate each other, contributing to the implementation of PO.

Discussion

Principal Findings

Our study aimed to characterize the barriers and facilitators to PO. We found significant determinants of PO implementation at the patient, provider, and systems levels. In our analysis, the TDF provides a useful framework to conceptualize PO service implementation and, combined with the BCW, informs the design of interventions to improve PO implementation.

Our findings associated with system and operational issues are consistent with other studies. Prior research reported substantial structural or logistical issues in PO implementation, including coordinating or awaiting sample collection [25]. Our results add to previous work by presenting a more detailed analysis of specific factors associated with these problems, including interdepartmental communication–related barriers and the obstruction of interdepartmental communication via EHR ordering processes. Our data on the domains of Cognitive and Interpersonal Skills reiterated these challenges in communication. Furthermore, our finding about the role that staffing plays in PO care delivery, with a specific nod to the added benefit of patient navigators to help streamline test ordering, sample procurement, and results communication from pathology to providers, is novel.

From the patients' perspective, our findings are associated with the social determinants of health around patient financial, logistics, and education specific to the VA population. The existing literature reports concerns around insurance coverage, high copays, and overall costs associated with treatment, which contrasts with our findings because VA participants do not bear those financial burdens [26]. This discrepancy is likely related to the VA single-payer model, which minimizes patient costs and third-party payer interactions, thereby eliminating a barrier to PO. Notably, existing literature is concordant with our findings that logistical challenges outside of cost (eg, transportation) represent barriers [26,27].

Patient education was perceived as a facilitator in our population, but only some patients noted receiving materials from their physicians, with many others turning to external resources. Several prior studies note that education can facilitate precision medicine when available and appropriately disseminated because it promotes patient engagement with care [28]. Yet, information sources sought by patients outside of their clinical team also suggest there is a risk for patients around misinformation in understanding precision medicine [26]. Our findings are also consistent with previous work identifying

XSL•FO
RenderX

limited patient knowledge as a barrier to PO clinical trial accrual [28].

Regarding provider *Knowledge*, previous survey research reported that approximately 1 in 5 providers felt "low confidence" in their knowledge of clinical genomics and believed the tests would be used for less than 10% of patients [29]. Additional studies report that a lack of provider knowledge in genomic medicine may lead to delayed use [30]. Our findings reflect similar insecurities in providers regarding PO knowledge. Our data, however, build on previous literature by noting enthusiasm for systems-level educational interventions (ie, tumor boards noted as "immensely helpful") and increasing comfort in interpreting NGS results (ie, providers feel "better in terms of outlining what's pathogenic and what's not"). Our data suggest a role for educational initiatives at the provider and patient levels to mitigate perceived knowledge gaps. Taken together, our findings along with the existing literature suggest that a deliberately implemented educational initiative may help facilitate PO uptake while also ensuring that patients do not fall victim to misinformation.

We noted that findings regarding the domains of Memory and Attention have substantial interplay with the findings in physician *Environment*, insofar as both relate to the structural/logistical burden that physicians face when implementing PO. Our findings were concordant with previous work suggesting barriers to precision medicine in time investment by physicians reviewing records, coordinating or awaiting sample collection, determining patient eligibility, and resolving insurance-related issues [26].

With regard to the domains of Goals and Reinforcement, previous studies similarly found that patients sought the benefits of PO clinical trials over previous less-targeted approaches to cancer care [31]. Our results support similar findings not only among patients considering clinical trials but also among all patients considering PO as standard care.

Limitations

The limitations of our study include the following: (1) a small sample size comprising entirely of men, (2) VA-only study sites, (3) a risk of social desirability bias on the part of participants, and (4) bias in interpretation. Regarding sample size, our number of 33 (17 patients and 16 providers), although a relatively small cohort, was determined using theoretical saturation in which recruitment stopped when few additional new themes emerged. This approach increases efficiency while covering the majority of relevant themes. Of note, this is the largest sample of both patients and providers in any published work on the topic.

Regarding VA-only study sites, our study was conducted within the VA and represents viewpoints from a single, uniform health system. As such, external factors including insurance coverage, access to care, and access to trials are different from those in the private health care sector. Nonetheless, the VA setting does allow for geographic diversity and frequently samples patients with limited economic means, a population with well-reported underrepresentation in research [32,33].

A further consequence of VA-only sites is that our sample includes only male patients. The VA patient population is approximately 90% male, but the number of female patients is growing [34]. They may have unique PO treatment needs; thus, their perceptions may differ from those of their male counterparts. While we considered oversampling female patients, the cancers they are treated for are different from those of male patients and may change the PO oncology experience. Future research should focus specifically on female veterans to understand their needs with regard to PO.

We also chose to focus on the cancers that are most frequently treated in VA oncology clinics and to incorporate PO, that is, prostate, lung, and colorectal cancers, which yielded a predominantly male patient base. We made these decisions to focus on high-volume cancers, which can illuminate systemic challenges in a way that infrequently treated cancers cannot.

Finally, we recognize the constraints of social desirability bias along with biases in the interpretation of results. To mitigate social desirability bias, we maintained participant anonymity and confidentiality and used verbal and nonverbal cues to increase patient comfort [32]. To mitigate bias in interpretation, all interviews were coded independently by 3 reviewers, and disagreements on coding were resolved by discussion until consensus.

Conclusion

In conclusion, our interviews with patients and providers added essential data to an understudied and quickly growing area of cancer care. Patients and providers reported barriers to Knowledge, Environment, Goals, and Attention. When mapped to the TDF or BCW, these identified barriers facilitate the design of a multilevel intervention for patients, providers, and systems to educate and improve the implementation of PO. Future research should engage with broader populations of specific cancers and also examine geographic variation more closely.

Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT (Generative AI Delegation Taxonomy; 2025), the following tasks were delegated to GAI tools under full human supervision: data visualization to generate the conceptual model in Figure 1. The GAI tool used was Gemini 2.5 Nano Banana Image Generator. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Additionally, we used the Gemini image generator available within the NYU Langone Health closed AI system; thus, no data were fed into public models.

XSL•FO
RenderX

Funding

This study received funding from the Department of Veterans Affairs Health Services Research and Development Service (HSR&D 1 IK2 HX002808-01) to DB (principal investigator) from March 1, 2020, to Feburary 28, 2025, as part of the Improving the Use of Tumor Genetic Testing with Decision Science project.

Data Availability

The data are preserved in Atlas TI and are available for further review upon request.

Authors' Contributions

Conceptualization: AS, DB, DM, SS
Data curation: AB, DB, KC
Formal analysis: AB, AS, DB, KC
Funding acquisition: DB
Methodology: AS
Project administration: AV
Writing – original draft: AB, AS, CM, DB, KC
Writing – review & editing: AS, CM, DB
All authors approved of the final version.

Conflicts of Interest

None declared.

References

1. Kris MG, Johnson BE, Berry LD, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. JAMA 2014 May 21;311(19):1998-2006. [doi: 10.1001/jama.2014.3741] [Medline: 24846037]
2. NCCN guidelines. National Comprehensive Cancer Network. 2025. URL: https://www.nccn.org/guidelines/category_1 [accessed 2026-03-25]
3. Becker DJ, Lee KM, Lee SY, et al. Uptake of KRAS testing and anti-EGFR antibody use for colorectal cancer in the VA. JCO Precis Oncol 2021;5:PO.20.00359. [doi: 10.1200/PO.20.00359] [Medline: 34250412]
4. Charlton ME, Kahl AR, Greenbaum AA, et al. KRAS testing, tumor location, and survival in patients with stage IV colorectal cancer: SEER 2010-2013. J Natl Compr Canc Netw 2017 Dec;15(12):1484-1493. [doi: 10.6004/jnccn.2017.7011] [Medline: 29223986]
5. Punekar SR, Griffin MM, Masri L, et al. Socioeconomic determinants of the use of molecular testing in stage IV colorectal cancer. Am J Clin Oncol 2021 Dec 1;44(12):597-602. [doi: 10.1097/COC.0000000000000875] [Medline: 34753883]
6. Levit LA, Byatt L, Lyss AP, et al. Closing the rural cancer care gap: three institutional approaches. JCO Oncol Pract 2020 Jul;16(7):422-430. [doi: 10.1200/OP.20.00174] [Medline: 32574128]
7. Loeb S, Li R, Sanchez Nolasco T, et al. Barriers and facilitators of germline genetic evaluation for prostate cancer. Prostate 2021 Aug;81(11):754-764. [doi: 10.1002/pros.24172] [Medline: 34057231]
8. Gamble CR, Huang Y, Wright JD, Hou JY. Precision medicine testing in ovarian cancer: the growing inequity between patients with commercial vs Medicaid insurance. Gynecol Oncol 2021 Jul;162(1):18-23. [doi: 10.1016/j.ygyno.2021.04.025] [Medline: 33958212]
9. Aldrighetti CM, Niemierko A, Van Allen E, Willers H, Kamran SC. Racial and ethnic disparities among participants in precision oncology clinical studies. JAMA Netw Open 2021 Nov 1;4(11):e2133205. [doi: 10.1001/jamanetworkopen.2021.33205] [Medline: 34748007]
10. Vashistha V, Armstrong J, Winski D, et al. Barriers to prescribing targeted therapies for patients with NSCLC with highly actionable gene variants in the Veterans Affairs National Precision Oncology Program. JCO Oncol Pract 2021 Jul;17(7):e1012-e1020. [doi: 10.1200/OP.20.00703] [Medline: 33780286]
11. Wright S, Daker-White G, Newman W, Payne K. Understanding barriers to the introduction of precision medicines in non-small cell lung cancer: a qualitative interview protocol. Wellcome Open Res 2018;3:24. [doi: 10.12688/wellcomeopenres.13976.1] [Medline: 29780892]
12. Ferreira-Gonzalez A, Ko G, Fusco N, et al. Barriers and facilitators to next-generation sequencing use in United States oncology settings: a systematic review. Future Oncol 2024;20(35):2765-2777. [doi: 10.1080/14796694.2024.2390821] [Medline: 39316553]
13. Crawford JM, Bry L, Pfeifer J, et al. The business of genomic testing: a survey of early adopters. Genet Med 2014 Dec;16(12):954-961. [doi: 10.1038/gim.2014.60] [Medline: 25010053]
14. Messner DA, Koay P, Al Naber J, et al. Barriers to clinical adoption of next-generation sequencing: a policy Delphi panel's solutions. Per Med 2017;14(4):339-354. [doi: 10.2217/pme-2016-0104] [Medline: 29230253]

XSL•FO
RenderX

15. Messner DA, Al Naber J, Koay P, et al. Barriers to clinical adoption of next generation sequencing: perspectives of a policy Delphi panel. Appl Transl Genom 2016 Sep;10:19-24. [doi: 10.1016/j.atg.2016.05.004] [Medline: 27668172]
16. Wu AC, Mazor KM, Ceccarelli R, Loomer S, Lu CY. The implementation process for pharmacogenomic testing for cancer-targeted therapies. J Pers Med 2018 Oct 1;8(4):32. [doi: 10.3390/jpm8040032] [Medline: 30275409]
17. Trosman JR, Weldon CB, Kelley RK, Phillips KA. Challenges of coverage policy development for next-generation tumor sequencing panels: experts and payers weigh in. J Natl Compr Canc Netw 2015 Mar;13(3):311-318. [doi: 10.6004/jnccn.2015.0043] [Medline: 25736008]
18. Veterans Affairs: National Oncology Program. URL: https://www.cancer.va.gov/ [accessed 2026-03-25]
19. Michie S, Johnston M, Abraham C, et al. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb;14(1):26-33. [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
20. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug;42(4):1758-1772. [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]
21. Atkins L, Francis J, Islam R, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implement Sci 2017 Jun 21;12(1):77. [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
22. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
23. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
24. Cascio MA, Lee E, Vaudrin N, Freedman DA. A team-based approach to open coding: considerations for creating intercoder consensus. Field methods 2019 May;31(2):116-130. [doi: 10.1177/1525822X19838237]
25. Hamilton JG, Banerjee SC, Carlsson SV, et al. Clinician perspectives on communication and implementation challenges in precision oncology. Per Med 2021 Sep;18(6):559-572. [doi: 10.2217/pme-2021-0048] [Medline: 34674550]
26. Bright K, Mills A, Bradford JP, Stewart DJ. RAPID framework for improved access to precision oncology for lethal disease: results from a modified multi-round Delphi study. Front Health Serv 2023;3:1015621. [doi: 10.3389/frhs.2023.1015621] [Medline: 36926496]
27. Morris HN, Winslow AT, Barreiro-Rosado JA, Torian S, Charlot M. Scoping review of barriers and facilitators to recruitment of black people with Cancer in Biospecimen-Based research. JCO Precis Oncol 2024 May;8:e2300708. [doi: 10.1200/PO.23.00708] [Medline: 38748944]
28. Gray SW, Hicks-Courant K, Cronin A, Rollins BJ, Weeks JC. Physicians' attitudes about multiplex tumor genomic testing. J Clin Oncol 2014 May 1;32(13):1317-1323. [doi: 10.1200/JCO.2013.52.4298] [Medline: 24663044]
29. Cooper KE, Abdallah KE, Angove RSM, Gallagher KD, Bonham VL. Navigating access to cancer care: identifying barriers to precision cancer medicine. Ethn Dis 2022;32(1):39-48. [doi: 10.18865/ed.32.1.39] [Medline: 35106043]
30. Spector-Bagdady K, Kent M, Krenz CD, et al. Patient and provider perspectives on enrollment in precision oncology research: qualitative ethical analysis. JMIR Cancer 2022 May 3;8(2):e35033. [doi: 10.2196/35033] [Medline: 35503525]
31. Serving her country: exploring the characteristics of women veterans. US Department of Veterans Affairs. URL: https://www.data.va.gov/stories/s/wci3-yrsv [accessed 2026-03-25]
32. Ford JG, Howerton MW, Lai GY, et al. Barriers to recruiting underrepresented populations to cancer clinical trials: a systematic review. Cancer 2008 Jan 15;112(2):228-242. [doi: 10.1002/cncr.23157] [Medline: 18008363]
33. Bispo Júnior JP. Social desirability bias in qualitative health research. Rev Saude Publica 2022;56:101. [doi: 10.11606/s1518-8787.2022056004164] [Medline: 36515303]
34. Institute of Medicine (US) Committee on Cancer Research Among Minorities and the Medically Underserved. In: Haynes MA, Smedley BD, editors. The Unequal Burden of Cancer: An Assessment of NIH Research and Programs for Ethnic Minorities and the Medically Underserved: National Academies Press; 1999. [doi: 10.17226/6377] [Medline: 20845564]

Abbreviations

BCW: Behavior Change Wheel
IRB: institutional review board
NGS: next-generation sequencing
PO: precision oncology
TDF: Theoretical Domains Framework
VA: Veterans Affairs
VA-NYHHS: Veterans Affairs Medical Center – New York Harbor Healthcare System
VAMC: Veterans Affairs Medical Center

XSL•FO
RenderX

Edited by M Balcarras; submitted 29.Sep.2025; peer-reviewed by D Howell, O Olowokere; revised version received 17.Feb.2026; accepted 09.Mar.2026; published 27.Apr.2026.

Please cite as:
Becker D, Csehak K, Barbaro A, Miller C, Vo A, Roman S, Makarov D, Sherman S, Squires A
Patient and Provider Experiences With Precision Oncology: Qualitative Descriptive Study at the Department of Veterans Affairs
JMIR Cancer 2026;12:e84858
URL: https://cancer.jmir.org/2026/1/e84858
doi:10.2196/84858

© Daniel Becker, Kenneth Csehack, Alexander Barbaro, Christian Miller, Antoinette Vo, Stefanie Roman, Danil Makarov, Scott Sherman, Allison Squires. Originally published in JMIR Cancer (https://cancer.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Social Media as a Platform for Cancer Care Decision-Making Among Women: Internet Survey-Based Study on Trust, Engagement, and Preferences

Anna Rose Johnson[1], MPH, MD; Grace Anne Longfellow[2], BS; Clara N Lee[2], MPP, MD; Benjamin Ormseth[3], MD; Gary B Skolnick[1], MBA; Mary C Politi[4], PhD; Yonaira M Rivera[5], MPH, PhD; Terence Myckatyn[1], MD, FRSC(C)

[1]Division of Plastic and Reconstructive Surgery, Department of Surgery, Washington University School of Medicine, Saint Louis, MO, United States
[2]Department of Surgery, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[3]Department of Plastic Surgery, University of Texas Southwestern Medical Center, Dallas, TX, United States
[4]Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, Saint Louis, MO, United States
[5]Department of Communication, School of Communication & Information, Rutgers University, New Brunswick, NJ, United States

Corresponding Author:
Terence Myckatyn, MD, FRSC(C)
Division of Plastic and Reconstructive Surgery
Department of Surgery
Washington University School of Medicine
1020 N. Mason Rd
Ste 110
Saint Louis, MO, 63141
United States
Phone: 1 3149968800
Fax: 1 314 996 3030
Email: myckatyn@wustl.edu

Related Article:

Correction of: https://cancer.jmir.org/2025/1/e64724

(JMIR Cancer 2026;12:e87683) doi:10.2196/87683

In "Social Media as a Platform for Cancer Care Decision-Making Among Women: Internet Survey-Based Study on Trust, Engagement, and Preferences. JMIR Cancer" [1], the authors noted one omission.

A "Funding" section has been added as follows:

> *Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R01CA276408. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Johnson AR, Longfellow GA, Lee CN, Ormseth B, Skolnick GB, Politi MC, et al. Social Media as a Platform for Cancer Care Decision-Making Among Women: Internet Survey-Based Study on Trust, Engagement, and Preferences. JMIR Cancer 2025 Mar 05;11:e64724 [FREE Full text] [doi: 10.2196/64724] [Medline: 40053770]

XSL•FO
RenderX

Submitted 12.Nov.2025; this is a non–peer-reviewed article;accepted 18.Nov.2025; published 13.Feb.2026.

Please cite as:
Johnson AR, Longfellow GA, Lee CN, Ormseth B, Skolnick GB, Politi MC, Rivera YM, Myckatyn T
Correction: Social Media as a Platform for Cancer Care Decision-Making Among Women: Internet Survey-Based Study on Trust, Engagement, and Preferences
JMIR Cancer 2026;12:e87683
URL: https://cancer.jmir.org/2026/1/e87683
doi:10.2196/87683
PMID:

©Anna Rose Johnson, Grace Anne Longfellow, Clara N Lee, Benjamin Ormseth, Gary B Skolnick, Mary C Politi, Yonaira M Rivera, Terence Myckatyn. Originally published in JMIR Cancer (https://cancer.jmir.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer Using Routinely Collected Clinical Data: Multicenter Retrospective Machine Learning Model Development and External Validation Study

Gehui Xu[1*], BMed; Xiaohe Sun[1*], MM; Jiaxin Jiang[1], BMed; Jin Sun[1], PhD; Liu Li[1], PhD; Haibo Cheng[1], PhD; Yuekun Zhu[2], PhD; Haiyi Liu[3], PhD

[1]First Clinical Medical College, Nanjing University of Chinese Medicine, 138 Xianlin Rd, Qixia District, Nanjing, China

[2]Department of General Surgery, First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China

[3]Department of Colorectal Surgery, Shanxi Provincial Cancer Hospital, Taiyuan, Shanxi, China

[*]these authors contributed equally

Corresponding Author:
Haibo Cheng, PhD
First Clinical Medical College, Nanjing University of Chinese Medicine, 138 Xianlin Rd, Qixia District, Nanjing, China

Abstract

Background: Preoperative nodal assessment in colorectal cancer (CRC) remains imperfect, and uncertainty regarding pathological lymph node metastasis (LNM) may persist despite conventional imaging. Machine learning using routinely available preoperative information may provide complementary patient-level risk estimates without requiring radiomic feature extraction or molecular testing.

Objective: This study aimed to develop and externally validate machine learning models for preoperative estimation of pathologically confirmed LNM risk in patients with newly diagnosed CRC undergoing direct radical surgery without preoperative anticancer treatment, and to assess whether multivariable prediction adds information beyond clinical tumor stage (cT) alone.

Methods: This multicenter retrospective study included 2725 patients from 2 hospitals. The internal cohort (n=1074) was split into training (n=753) and internal test (n=321) sets, while 1651 patients from a second center formed the external validation cohort. Candidate predictors were routinely available preoperative clinical, laboratory, tumor-marker, imaging-based, and endoscopic biopsy variables. Predictors were selected using univariable and multivariable logistic regression (LR). Seven algorithms were evaluated: random forest (RF), support vector machine (SVM), LR, decision tree (DT), extreme gradient boosting (XGBoost), naive Bayes (NB), and light gradient boosting machine (LightGBM). Performance was assessed using discrimination, threshold-dependent classification metrics, calibration, and decision curve analysis (DCA). Shapley additive explanations (SHAP) were used for interpretation. Sensitivity analyses compared the selected predictors with all candidate predictors and cT alone.

Results: Six predictors were retained: BMI, preoperative carcinoembryonic antigen, primary tumor site, cT, histological type, and tumor differentiation. RF was selected based on internal test performance. The area under the receiver operating characteristic curve (AUROC) was 0.803 (95% CI 0.756 - 0.849) in the internal test set and 0.776 (95% CI 0.755 - 0.797) in external validation. In external validation, accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were 0.687, 0.677, 0.693, 0.569, and 0.781, respectively. The Brier score was 0.190, and the calibration slope was 1.061. cT contributed most in SHAP analysis; however, the 6-predictor RF model showed significantly higher discrimination than cT alone in both the internal test set (AUROC 0.803 vs 0.712; ΔAUROC=0.091; P<.001) and external validation cohort (AUROC 0.776 vs 0.707; ΔAUROC=0.069; P<.001). Discrimination did not differ significantly between the 6-predictor and all-predictor RF models.

Conclusions: We developed and externally validated a 6-predictor RF model for preoperative LNM risk estimation in patients with newly diagnosed CRC undergoing radical resection with regional lymph node dissection without preoperative anticancer treatment. The model showed moderate external discrimination and provided predictive information beyond cT alone. It is intended for research-stage adjunctive risk stratification and should not replace imaging, pathological evaluation, or clinical judgment. Prospective validation in broader populations and clinical settings is required before implementation.

(JMIR Cancer 2026;12:e101162) doi:10.2196/101162

XSL•FO
RenderX

KEYWORDS

colorectal cancer; lymph node metastasis; machine learning; random forest; preoperative prediction; external validation; SHAP; risk stratification; shapley additive explanations

Introduction

Colorectal cancer (CRC) remains a major global health burden. According to Global Cancer Observatory (GLOBOCAN) 2024, more than 2.0 million new CRC cases and approximately 918,000 CRC-related deaths occurred worldwide in 2024, making CRC the third most commonly diagnosed cancer and the second leading cause of cancer death globally [1]. Regional lymph node metastasis (LNM) is a key component of CRC staging and an important prognostic factor [2]. Accurate preoperative assessment of nodal involvement is therefore important for defining disease extent and stratifying patient risk before surgery.

Conventional imaging is central to preoperative CRC staging and provides essential information on local tumor extent. However, radiological assessment of nodal involvement remains challenging because benign and metastatic lymph nodes may overlap in size and morphological appearance. A recent systematic review and meta-analysis found that size-based preoperative computed tomography (CT) achieved only moderate diagnostic performance for LNM in colon cancer [3]. Similarly, an updated review of rectal magnetic resonance imaging (MRI) identified nodal staging as a persistent challenge, with size and morphological criteria subject to substantial overlap and staging pitfalls [4]. Thus, although imaging provides indispensable anatomical information, uncertainty regarding pathological nodal status may remain before surgery. Integrating routinely available clinical, laboratory, and biopsy-derived information may therefore provide a more complete patient-level estimate of LNM risk.

Machine learning has increasingly been used to combine clinical, imaging, and pathological information for LNM prediction in CRC. However, existing models differ substantially in the populations studied, predictors included, and data modalities used, ranging from conventional clinical variables to radiomics and digital pathology, which complicates comparison across studies [5]. Models using high-dimensional imaging or pathological features may capture additional predictive information but also require more complex feature extraction and interpretation. Importantly, model performance observed during development may not be maintained in independent populations, and recent CRC-specific reviews have identified limited external validation as an important source of uncertainty in reported performance [6,7]. These issues are particularly relevant to preoperative LNM prediction, for which a useful model should provide reliable risk estimates using information that can be obtained consistently before surgery.

Accordingly, we aimed to develop and externally validate machine learning models for preoperative estimation of pathologically confirmed LNM risk in patients with newly diagnosed CRC undergoing direct radical surgery without preoperative anticancer treatment. Predictors were restricted to routinely available preoperative information, including clinical characteristics, laboratory and tumor-marker measurements, imaging-based clinical assessment, and endoscopic biopsy findings. We compared multiple machine learning algorithms, evaluated the selected model in an independent external cohort, and used model interpretation to characterize predictor contributions. We also assessed whether the multivariable model provided predictive information beyond clinical T stage alone. The intended use was adjunctive patient-level LNM risk stratification within the direct-surgery pathway.

Methods

Study Design, Setting, and Participants

This was a multicenter retrospective study for the development and external validation of a clinical prediction model. After ethics approval was obtained, investigators retrospectively collected data from patients with newly diagnosed CRC who underwent radical resection at the First Affiliated Hospital of Harbin Medical University or Shanxi Province Cancer Hospital between July 2023 and December 2025.

The inclusion criteria were as follows: (1) age ≥18 years; (2) a new diagnosis of CRC confirmed by postoperative pathology; (3) direct treatment with radical resection and regional lymph node dissection; and (4) ascertainable postoperative pathological regional LNM status. Patients managed with endoscopic resection, local excision, nonoperative treatment, or preoperative anticancer therapy were not eligible for the direct-surgery cohort. The exclusion criteria were as follows: (1) preoperative radiotherapy, chemotherapy, targeted therapy, immunotherapy, or other anticancer treatment; and (2) synchronous or metachronous malignant tumors other than CRC.

The sample size was determined by all consecutive eligible cases available during the prespecified study period, and no a priori sample size estimation was performed.

The First Affiliated Hospital of Harbin Medical University constituted the internal development cohort. Stratified random sampling based on LNM status was used to split the cohort into training and internal test sets at a 7:3 ratio. The random seed was fixed at one to ensure reproducibility. Shanxi Province Cancer Hospital constituted the independent external validation cohort.

Outcome Definition and Candidate Predictors

The primary outcome was regional LNM, confirmed by the postoperative pathological report. Patients were then classified as LNM-positive or LNM-negative according to the final pathological findings.

Candidate predictors were prespecified before model construction and were all obtainable through routine preoperative clinical assessment.

The candidate predictors included the following categories: Demographic and general clinical variables were age, sex,

XSL•FO
RenderX

ethnicity, residence, smoking history, drinking history, family history, comorbidities, and BMI. Laboratory indices were neutrophil percentage, systemic immune-inflammation index (SII), and albumin-to-globulin ratio (AGR). Tumor markers were preoperative carcinoembryonic antigen (pCEA) and preoperative carbohydrate antigen 19‐9 (pCA19-9). Tumor-related variables were clinical T stage, bowel obstruction, primary tumor site, maximum tumor diameter, gross tumor type, histological type, and tumor differentiation. Comorbidities were recorded as a binary variable (yes or no). Patients with a documented history of hypertension, diabetes mellitus, or coronary heart disease were classified as having comorbidities; patients without any of these conditions were classified as having no comorbidities.

Laboratory indicators were obtained from the most recent preoperative test results, and values measured within 1 week before surgery were used whenever available. SII was calculated as platelet count×absolute neutrophil count/absolute lymphocyte count. AGR was calculated as serum albumin/globulin, where globulin=total protein−albumin. Tumor differentiation and histological type were obtained from preoperative endoscopic biopsy pathology reports, whereas maximum tumor diameter and clinical T stage were obtained from preoperative imaging and/or endoscopic evaluation.

Missing Data Handling

Before data preprocessing, the number and proportion of missing values for each candidate predictor were calculated separately in the training, internal test, and external validation sets. Patients with missing postoperative pathological LNM status were excluded before cohort formation, and the outcome was not imputed. The internal development cohort was divided into the training and internal test sets before missing-data handling, and the external validation cohort remained independent throughout. For limited missingness in continuous predictors, multiple imputation by chained equations was performed with the R package *mice* (Stef van Buuren and Karin Groothuis-Oudshoorn), generating 3 imputed data sets with a maximum of 50 iterations per chain and a random seed of 123. Missing categorical predictor values were imputed using the training-set mode. The imputation model and derived preprocessing rules were estimated from the training data and then applied without refitting to the internal test and external validation sets. Variable-level missingness before imputation is reported in Table S2 in Multimedia Appendix 1; because missingness was summarized by variable, the same patient could contribute to more than one missing-value count.

Biopsy-Resection Agreement Analysis

Among patients with both preoperative endoscopic biopsy and postoperative resection specimen pathology results, the agreement in histological type and tumor differentiation was evaluated. This analysis included only patients with nonmissing preoperative and postoperative results for the corresponding pathological indicator, and no imputed values were used.

Histological type was uniformly classified as adenocarcinoma and special histological type, with the latter including mucinous adenocarcinoma and signet-ring cell carcinoma. Agreement was evaluated using a cross-classification table, the exact agreement rate, the proportion of special histological type missed on preoperative biopsy, the positive predictive value (PPV) of a preoperative diagnosis of special histological type, and the unweighted Cohen κ coefficient with its 95% CI. A missed special histological type was defined as a preoperative biopsy diagnosis of adenocarcinoma with a postoperative resection specimen diagnosis of special histological type, and its proportion was calculated with patients confirmed as having special histological type postoperatively as the denominator. Tumor differentiation was uniformly classified as G1, G2, and G3, G4. Concordance was assessed using cross-classification tables, the exact agreement rate, the proportion with underestimated grade, the proportion with overestimated grade, and the linearly weighted Cohen κ coefficient with its 95% CI. Underestimated grade was defined as a pathologic grade on preoperative biopsy that was lower than the pathologic grade on the postoperative resection specimen, indicating that the biopsy suggested a more favorable grade; overestimated grade was defined as a pathologic grade on preoperative biopsy that was higher than the pathologic grade on the postoperative resection specimen. The 95% CIs for all percentages were calculated using the Wilson method, and the 95% CI for the Cohen κ coefficient was estimated by 20,000 multinomial bootstrap resamples. Detailed results are presented in Table S3 in Multimedia Appendix 1.

Predictor Preprocessing and Encoding

After missing-data handling, predictors were encoded according to their variable type before model development. BMI and pCEA were retained as continuous variables without standardization. Primary tumor site was one-hot encoded, with right-sided colon as the reference category and separate indicator variables for left-sided colon and rectum. Histological type was binary encoded as adenocarcinoma=0 and mucinous adenocarcinoma or signet-ring cell carcinoma=1. Clinical Tumor stage (cT) and tumor differentiation were encoded as ordinal integers ([clinical Tumor1 stage] cT1=0, [clinical Tumor2 stage] cT2=1, [clinical Tumor3 stage] cT3=2, [clinical Tumor4 stage] cT4=3; G1=0, G2=1, and G3 or G4=2, respectively).

The same preprocessing and encoding scheme was applied across all candidate machine learning models. The preprocessing rules were defined using the training data and then applied unchanged to the internal test and external validation sets. Detailed predictor preprocessing and encoding rules are provided in Table S5 in Multimedia Appendix 1.

Predictor Selection

Predictor selection was performed exclusively in the training set to avoid information leakage. Prespecified candidate predictors were first evaluated using univariable binary logistic regression (LR). Variables associated with LNM at $P<.05$ were subsequently entered into a multivariable binary LR model. Variables remaining statistically significant at $P<.05$ in the multivariable analysis were retained as the final predictor set for subsequent machine learning model development. Associations were reported as odds ratios (ORs) with 95% CIs. The internal test and external validation sets were not used for predictor selection.

XSL•FO
RenderX

Model Development and Hyperparameter Tuning

Using the predictors retained through the prespecified selection procedure, 7 machine learning models were developed: random forest (RF), support vector machine (SVM), LR, decision tree (DT), extreme gradient boosting (XGBoost), naive Bayes (NB), and light gradient-boosting machine (LightGBM). All model development and hyperparameter tuning were performed exclusively using the training set. In the training set, LNM was present in 278 of 753 (36.9%) patients. RF therefore used class_weight=balanced to account for the moderate outcome imbalance. Final hyperparameter settings are provided in Table S5 in Multimedia Appendix 1.

For RF, SVM, DT, XGBoost, and LightGBM, hyperparameters were optimized using grid search with 3-fold cross-validation within the training set, with the area under the receiver operating characteristic curve (AUROC) as the optimization metric. LR and NB were fitted without hyperparameter tuning. The final hyperparameter settings for all candidate models are provided in Table S5 in Multimedia Appendix 1.

Model Performance Evaluation and Validation

Model performance was evaluated in the training, internal test, and independent external validation sets. Discrimination was assessed using receiver operating characteristic (ROC) curves and the AUROC with 95% CIs. Classification performance was evaluated using accuracy, sensitivity, specificity, F_1 score, PPV, and negative predictive value (NPV).

For each model, an evaluation operating point was determined by maximizing the Youden index in the training set and was then applied unchanged to the internal test and external validation sets to calculate threshold-dependent classification metrics. No operating-point reoptimization was performed using the internal test or external validation data. These training-derived operating points were used for model evaluation and were not considered validated clinical decision cutoffs.

Calibration was assessed using calibration curves, the Brier score, calibration intercept, and calibration slope. Clinical utility was explored using decision curve analysis (DCA) by estimating net benefit across a prespecified threshold-probability display interval of 0.05 - 0.80. The final model was selected based on prespecified overall performance in the internal test set, considering discrimination, classification performance, calibration, clinical utility, and the balance of sensitivity, specificity, and F_1 score. After model selection, the selected model was evaluated in the independent external validation cohort without further fitting, hyperparameter tuning, or threshold optimization. The DCA interval was descriptive and was not treated as a validated clinical decision threshold.

After selection of the final model, its AUROC was compared pairwise with those of the other candidate models in the internal test and external validation sets using the DeLong test. To account for multiple comparisons, *P* values were adjusted separately within each data set using the Holm method.

Model Interpretability

Shapley additive explanations (SHAP) analysis was performed for the model selected according to the prespecified evaluation framework to characterize the contribution of individual predictors to model output. Global feature importance was quantified using the mean absolute SHAP value for each encoded feature. A SHAP summary plot was used to visualize both the magnitude and direction of predictor contributions across patients.

Patient-level interpretation was performed using an individual SHAP explanation to illustrate how multiple predictors jointly contributed to a single prediction. SHAP dependence plots were additionally generated to examine the relationship between predictor values and their contributions to model output. For primary tumor site, the 2 dummy variables representing left-sided colon and rectum were interpreted separately, with right-sided colon serving as the reference category; the dummy-variable values were not interpreted as an ordinal risk gradient.

Sensitivity Analyses

Sensitivity analyses were performed to examine the robustness of the predictor-selection strategy and the incremental predictive value of the selected predictor set beyond clinical T stage alone. Using the algorithm selected as the final model, 3 models were constructed: a model based on the predictor set retained through the prespecified selection procedure, a model incorporating all prespecified candidate predictors, and a model using clinical T stage alone.

All models were developed using the training set and evaluated in the internal test and external validation sets. Their performance was compared in terms of discrimination, threshold-dependent classification performance, calibration, and clinical utility. Pairwise AUROC comparisons between the 6-predictor model and the 2 comparator models were performed using the DeLong test. Detailed sensitivity-analysis results are presented in Table S10 and Figure S4 in Multimedia Appendix 1.

Statistical Analysis and Software

Two-sided tests were used for all analyses. Statistical significance was defined at α=.05. We assessed normality of continuous variables with the Shapiro-Wilk test. Normally distributed data are presented as mean (SD). These data were compared with the independent-samples *t* test. Skewed data are presented as median (IQR). They were analyzed with the Mann-Whitney *U* test. We summarized categorical variables as frequencies and percentages. Group differences were examined with the chi-square test or Fisher exact test, as appropriate. When Fisher exact test was used, chi-square statistics were not reported.

Associations between candidate predictors and LNM were examined using univariable and multivariable binary LR models, with effect estimates reported as ORs and 95% CIs.

Statistical analyses and missing-data handling were performed using R version 4.5.1 (R Core Team). Machine learning model development, performance evaluation, and model interpretation were performed using Python (Python Software Foundation) version 3.12.3 with scikit-learn (scikit-learn developers) version 1.4.2, XGboost (Distributed Machine Learning Community)

XSL•FO
RenderX

version 2.1.1, lightGBM (Microsoft Corporation) version 4.5.0, and SHAP (Scott Lundberg) version 0.45.1.

Ethical Considerations

This study was approved by the Ethics Committee of the First Affiliated Hospital of Harbin Medical University (approval No. 202452) and the Ethics Committee of Cancer Hospital, Chinese Academy of Medical Sciences, Shanxi Hospital (Shanxi Province Cancer Hospital; approval No. KY2024163). At the First Affiliated Hospital of Harbin Medical University, the requirement for informed consent was waived because of the retrospective study design. All patient-level data used for the present analysis were deidentified before analysis.

Results

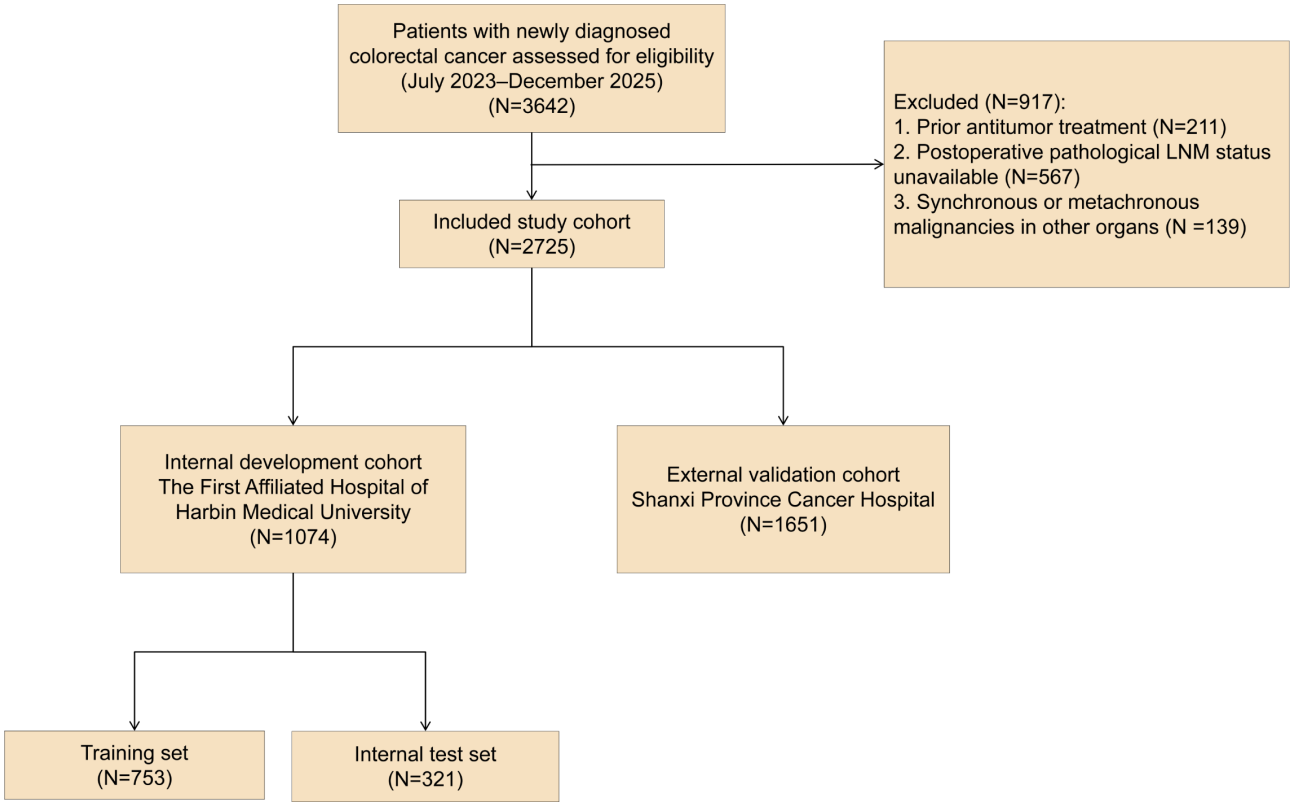
Study Population and Data Characteristics

Patient Inclusion and Cohort Allocation

During the study period, 3642 patients with newly diagnosed CRC were screened for eligibility. Of these, 211 patients were excluded because they had received preoperative anticancer treatment, 567 because postoperative pathological LNM status could not be ascertained, and 139 because of synchronous or metachronous malignancies other than CRC. The final study cohort therefore comprised 2725 patients from 2 independent medical centers (Figure 1).

The internal development cohort included 1074 patients from the First Affiliated Hospital of Harbin Medical University. Stratified random splitting at a 7:3 ratio yielded 753 patients for model training and 321 for internal testing. The independent external validation cohort comprised 1651 patients from Shanxi Province Cancer Hospital.

Figure 1. Study flowchart for patient screening, cohort allocation, model development, and validation. LNM: lymph node metastasis.



Baseline Characteristics and Missing Data

The development and external validation cohorts differed significantly in pCA19-9 (P=.004), SII (P=.03), residence (P<.001), family history (P<.001), smoking status (P=.03), and the number of lymph nodes examined (P=.01). No significant differences were observed in the other baseline characteristics, including LNM status (Table 1). No significant differences were observed between the training and internal test sets (Table S1 in Multimedia Appendix 1). Missingness was low overall, with missing proportions below 5% for all prespecified candidate predictors across the 3 data sets. The maximum missingness was 35/753 (4.65%) in the training set, 14/321 (4.36%) in the internal test set, and 78/1651 (4.72%) in the external validation set (Table S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Baseline characteristics of the development and external validation cohorts.

Characteristic[a]	Development cohort (n=1074)	External validation cohort (n=1651)	Statistics[b]	*P* value
Age, median (IQR) years	66.00 (60.00 - 73.00)	67.00 (58.00 - 73.00)	0.100[c]	.92
Sex, n (%)			2.880 (1)[d]	.09
Male	679 (63.2)	989 (59.9)		
Female	395 (36.8)	662 (40.1)		
Ethnicity, n (%)			3.650 (1)[d]	.06
Han	1059 (98.6)	1641 (99.4)		
Other	15 (1.4)	10 (0.6)		
Residence, n (%)			108.589 (1)[d]	<.001
Urban	768 (71.5)	848 (51.4)		
Rural	306 (28.5)	803 (48.6)		
BMI, kg/m², median (IQR)	23.88 (21.26 - 26.22)	23.71 (21.57 - 26.25)	-0.216[c]	.83
Smoking status, n (%)			6.943 (2)[d]	.03
Never smoking	874 (81.4)	1335 (80.9)		
Current smoking	187 (17.4)	272 (16.5)		
Former smoking	13 (1.2)	44 (2.7)		
Drinking status, n (%)			4.415[d] (2)	.11
Never drinking	926 (86.2)	1392 (84.3)		
Current drinking	136 (12.7)	248 (15.0)		
Former drinking	12 (1.1)	11 (0.7)		
Family history, n (%)			115.977[d] (1)	<.001
Yes	6 (0.6)	191 (11.6)		
No	1068 (99.4)	1460 (88.4)		
Comorbidities, n (%)			2.196[d] (1)	.14
Yes	942 (87.7)	1414 (85.6)		
No	132 (12.3)	237 (14.4)		
Bowel obstruction, n (%)			3.417[d] (1)	.07
Yes	215 (20.0)	283 (17.1)		
No	859 (80.0)	1368 (82.9)		
pCEA[e], median (IQR) ng/mL	4.21 (2.35 - 10.31)	4.47 (2.04 - 9.59)	0.938[c]	.35
pCA19-9[f], median (IQR) U/mL	9.96 (5.14 - 19.32)	11.59 (6.24 - 21.62)	-2.909[c]	.004
Neutrophil percentage, %	63.60 (56.02 - 70.30)	62.40 (55.50 - 69.85)	1.594[c]	.11
SII[g], median (IQR)	598.63 (400.50 - 993.44)	565.20 (392.04 - 897.17)	2.163[c]	.03
AGR[h], median (IQR)	1.46 (1.27 - 1.66)	1.45 (1.29 - 1.62)	0.857[c]	.39
Primary tumor site, n (%)			4.280 (2)[d]	.12
Rectum	454 (42.3)	670 (40.6)		
Left-sided colon	332 (30.9)	478 (29.0)		
Right-sided colon	288 (26.8)	503 (30.5)		

XSL•FO
RenderX

Characteristic[a]	Development cohort (n=1074)	External validation cohort (n=1651)	Statistics[b]	*P* value
Maximum tumor diameter, median (IQR) cm	3.80 (3.00 - 4.50)	3.50 (3.00 - 5.00)	-1.895[c]	.06
cT[i], n (%)			2.490 [d](3)	.48
cT1	65 (6.1)	82 (5.0)		
cT2	223 (20.8)	330 (20.0)		
cT3	713 (66.4)	1136 (68.8)		
cT4	73 (6.8)	103 (6.2)		
Gross tumor type, n (%)			3.191 (1)[d]	.07
Ulcerative or infiltrative	725 (67.5)	1169 (70.8)		
Protruding	349 (32.5)	482 (29.2)		
Histological type, n (%)			3.557[d] (1)	.06
Adenocarcinoma	898 (83.6)	1425 (86.3)		
Mucinous adenocarcinoma or signet-ring cell carcinoma	176 (16.4)	226 (13.7)		
Tumor differentiation, n (%)			4.291[d] (2)	.12
G1	51 (4.7)	58 (3.5)		
G2	830 (77.3)	1259 (76.3)		
G3 or G4	193 (18.0)	334 (20.2)		
Lymph node metastasis, n (%)			0.082 (1)[d]	.77
Yes	396 (36.9)	619 (37.5)		
No	678 (63.1)	1032 (62.5)		
Number of lymph nodes examined, median (IQR)	14.00 (12.00 - 17.00)	14.00 (12.00 - 17.00)	2.527[c]	.01

[a]Values are median (IQR) or n (%) unless otherwise indicated. The Mann-Whitney U test was used for continuous variables, and the chi-square test or Fisher exact test was used for categorical variables, as appropriate.

[b]Values are median (IQR) or n (%) unless otherwise indicated. The Mann-Whitney *U* test was used for continuous variables, and the chi-square test was used for categorical variables. Z denotes the standardized test statistic from the Mann-Whitney *U* test. Z values are reported for continuous variables, whereas chi-square values are reported for categorical variables, with degrees of freedom given in parentheses.

[c]Z value.

[d]χ^2 (df).

[e]pCEA: preoperative carcinoembryonic antigen.

[f]pCA19-9: preoperative carbohydrate antigen 19 - 9.

[g]SII: systemic immune-inflammation index.

[h] AGR: albumin-to-globulin ratio.

[i]cT: clinical Tumor stage.

Biopsy-Resection Agreement

Agreement between preoperative biopsy and postoperative resection pathology was moderate for both histological type and tumor differentiation. For histological type, the exact agreement rate was 88.5% (2392/2703), with an unweighted Cohen κ of 0.616 (95% CI 0.577 - 0.653); 42.5% (248/583) of postoperative special histological types had been classified as adenocarcinoma on preoperative biopsy. For tumor differentiation, the exact agreement rate was 77% (2074/2692), with a linearly weighted Cohen κ of 0.521 (95% CI 0.488 - 0.552); undergrading and overgrading occurred in 13.9% (374/2692) and 9.1% (244/3692) of cases, respectively (Table S3 in Multimedia Appendix 1).

Predictor Selection

Univariable LR Analysis

Univariable analysis identified BMI, pCEA, bowel obstruction, maximum tumor diameter, gross tumor type, primary tumor site, clinical T stage, histological type, and tumor differentiation as significantly associated with LNM (Table S4 in Multimedia Appendix 1). Higher BMI and pCEA levels, bowel obstruction, right-sided colon or rectal location, mucinous adenocarcinoma or signet-ring cell carcinoma, and poorer tumor differentiation

XSL•FO
RenderX

were associated with higher odds of LNM, whereas protruding gross tumor type was associated with lower odds. Compared with cT1, cT3 and cT4 were associated with substantially higher odds of LNM (OR 8.238 and 11.520, respectively; both *P*<.001). cT2 showed a lower odds estimate than cT1 in the univariable analysis (OR 0.186, 95% CI 0.043 - 0.810; *P*=.03). The remaining candidate predictors were not significantly associated with LNM.

Multivariable LR Analysis

Multivariable LR identified BMI, pCEA, primary tumor site, cT, histological type, and tumor differentiation as factors independently associated with LNM (Table 2). Higher BMI (OR 1.154, 95% CI 1.095 - 1.216; *P*<.001) and pCEA (OR 1.005, 95% CI 1.001 - 1.010; *P*=.03) were associated with higher odds of LNM. Compared with left-sided colon tumors, right-sided colon tumors (OR 1.756, 95% CI 1.101 - 2.801; *P*=.02) and rectal tumors (OR 1.838, 95% CI 1.196 - 2.824; *P*=.005) also showed higher odds of LNM.

Clinical T stage showed the strongest associations: compared with cT1, cT3 and cT4 were associated with markedly higher odds of LNM (OR 11.798 and 17.743, respectively; both *P*<.001), whereas cT2 was not statistically significant (*P*=.10). Mucinous adenocarcinoma or signet-ring cell carcinoma was associated with higher odds than adenocarcinoma (OR 1.880, 95% CI 1.172 - 3.017; *P*=.009). Compared with G1 differentiation, G2 (OR 5.122, 95% CI 1.082 - 24.237; *P*=.04) and G3 or G4 (OR 13.593, 95% CI 2.755 - 67.070; *P*=.001) were associated with higher odds of LNM. Gross tumor type, bowel obstruction, and maximum tumor diameter were no longer statistically significant after multivariable adjustment.

XSL•FO
RenderX

Table . Multivariable logistic regression analysis of factors associated with lymph node metastasis in the training set. Reference categories are shown directly in the table.

Variable	β coefficient	SE	Adjusted OR[a] (95% CI)	*P* values
BMI, kg/m²	.143	0.027	1.154 (1.095 - 1.216)	<.001
pCEA[b], ng/mL	.005	0.002	1.005 (1.001 - 1.010)	.03
Gross tumor type				
Ulcerative or infiltrative (Reference)				
Protruding	−.134	0.216	0.875 (0.573 - 1.336)	.57
Bowel obstruction				
No (Reference)				
Yes	−.003	0.221	0.997 (0.647 - 1.538)	.99
Maximum tumor diameter, cm	−.112	0.059	0.894 (0.797 - 1.003)	.06
Primary tumor site				
Left-sided colon (Reference)				
Right-sided colon	.563	0.238	1.756 (1.101 - 2.801)	.02
Rectum	.609	0.219	1.838 (1.196 - 2.824)	.005
Clinical T stage				
cT1 (Reference)				
cT2	−1.335	0.811	0.263 (0.054 - 1.291)	.100
cT3	2.468	0.581	11.798 (3.775 - 36.866)	<.001
cT4	2.876	0.650	17.743 (4.963 - 63.425)	<.001
Histological type				
Adenocarcinoma (Reference)				
Mucinous adenocarcinoma or signet-ring cell carcinoma	.631	0.241	1.880 (1.172 - 3.017)	.009
Tumor differentiation				
G1 (Reference)				
G2	1.634	0.793	5.122 (1.082 - 24.237)	.04
G3 or G4	2.610	0.814	13.593 (2.755 - 67.070)	.001

[a]OR: odds ratio.

[b]pCEA: preoperative carcinoembryonic antigen.

Model Performance and Model Selection

Discrimination and Classification Performance

The discrimination and classification performance of the 7 candidate models is summarized in Table 3 and Figure 2. In the internal test set, AUROCs ranged from 0.708 to 0.809. The RF model achieved an AUROC of 0.803 (95% CI 0.756 - 0.849), with an accuracy of 0.720, sensitivity of 0.763, specificity of 0.695, and F_1 score of 0.667. In the external validation set, RF showed the highest AUROC among the candidate models (0.776, 95% CI 0.755 - 0.797), with an accuracy of 0.687, sensitivity of 0.677, specificity of 0.693, and F_1 score of 0.618.

Pairwise DeLong comparisons showed that, in the internal test set, the RF AUROC differed significantly only from that of the DT after Holm adjustment. In the external validation set, RF showed significantly higher AUROCs than SVM, LR, DT, and NB, whereas the differences versus XGBoost and LightGBM were not statistically significant after adjustment (Table S6 in Multimedia Appendix 1). Using the RF training-derived Youden operating point of 0.545, the PPV and NPV were 0.592 and 0.834 in the internal test set and 0.569 and 0.781 in the external validation set, respectively. The corresponding false-positive and false-negative counts were 62 and 28 in the internal test set and 317 and 200 in the external validation set. Threshold-dependent classification metrics are summarized in Table 3, while PPV, NPV, and full confusion-matrix counts are provided in Table S7 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Predictive performance of the candidate machine learning models in the training, internal test, and external validation sets.

Model	Data set[a]	AUROC[b] (95% CI)	Accuracy (95% CI)[c]	Sensitivity (95% CI)[c]	Specificity (95% CI)[c]	F_1 score (95% CI)[c]
RF[d]	Training	0.862 (0.837 - 0.887)	0.768 (0.735 - 0.799)	0.809 (0.760 - 0.853)	0.743 (0.703 - 0.785)	0.720 (0.680 - 0.762)
RF	Internal test	0.803 (0.756 - 0.849)	0.720 (0.670 - 0.769)	0.763 (0.677 - 0.834)	0.695 (0.630 - 0.759)	0.667 (0.593 - 0.731)
RF	External validation	0.776 (0.755 - 0.797)	0.687 (0.664 - 0.709)	0.677 (0.644 - 0.716)	0.693 (0.665 - 0.719)	0.618 (0.588 - 0.649)
SVM[e]	Training	0.798 (0.767 - 0.832)	0.709 (0.679 - 0.741)	0.827 (0.779 - 0.876)	0.640 (0.597 - 0.684)	0.677 (0.635 - 0.717)
SVM	Internal test	0.790 (0.744 - 0.836)	0.704 (0.654 - 0.756)	0.847 (0.776 - 0.914)	0.621 (0.559 - 0.693)	0.678 (0.612 - 0.735)
SVM	External validation	0.751 (0.725 - 0.775)	0.652 (0.629 - 0.676)	0.732 (0.696 - 0.768)	0.605 (0.576 - 0.634)	0.612 (0.584 - 0.641)
LR[f]	Training	0.754 (0.717 - 0.789)	0.664 (0.633 - 0.699)	0.827 (0.783 - 0.873)	0.568 (0.526 - 0.611)	0.645 (0.606 - 0.683)
LR	Internal test	0.778 (0.724 - 0.828)	0.713 (0.660 - 0.766)	0.856 (0.794 - 0.918)	0.631 (0.566 - 0.702)	0.687 (0.625 - 0.748)
LR	External validation	0.671 (0.644 - 0.696)	0.593 (0.571 - 0.616)	0.703 (0.670 - 0.737)	0.527 (0.495 - 0.557)	0.564 (0.538 - 0.597)
DT[g]	Training	0.942 (0.927 - 0.956)	0.849 (0.821 - 0.874)	0.863 (0.821 - 0.903)	0.840 (0.808 - 0.876)	0.808 (0.771 - 0.843)
DT	Internal test	0.708 (0.649 - 0.766)	0.651 (0.598 - 0.704)	0.644 (0.552 - 0.738)	0.655 (0.593 - 0.727)	0.576 (0.493 - 0.652)
DT	External validation	0.675 (0.646 - 0.699)	0.637 (0.612 - 0.661)	0.572 (0.529 - 0.610)	0.675 (0.648 - 0.706)	0.541 (0.505 - 0.573)
XGBoost[h]	Training	0.871 (0.846 - 0.896)	0.786 (0.758 - 0.817)	0.809 (0.766 - 0.855)	0.773 (0.736 - 0.810)	0.736 (0.696 - 0.779)
XGBoost	Internal test	0.790 (0.741 - 0.833)	0.723 (0.670 - 0.769)	0.729 (0.644 - 0.802)	0.719 (0.655 - 0.773)	0.659 (0.590 - 0.719)
XGBoost	External validation	0.768 (0.747 - 0.789)	0.681 (0.660 - 0.704)	0.653 (0.617 - 0.688)	0.698 (0.670 - 0.726)	0.605 (0.575 - 0.636)
NB[i]	Training	0.798 (0.767 - 0.829)	0.707 (0.672 - 0.740)	0.856 (0.813 - 0.896)	0.619 (0.575 - 0.665)	0.683 (0.640 - 0.718)
NB	Internal test	0.809 (0.766 - 0.854)	0.732 (0.682 - 0.777)	0.890 (0.828 - 0.943)	0.640 (0.571 - 0.706)	0.709 (0.654 - 0.767)
NB	External validation	0.734 (0.708 - 0.758)	0.643 (0.621 - 0.667)	0.743 (0.710 - 0.774)	0.583 (0.552 - 0.613)	0.610 (0.580 - 0.637)
LightGBM[j]	Training	0.872 (0.849 - 0.895)	0.780 (0.752 - 0.807)	0.831 (0.788 - 0.871)	0.749 (0.711 - 0.788)	0.736 (0.698 - 0.772)
LightGBM	Internal test	0.791 (0.752 - 0.835)	0.717 (0.673 - 0.765)	0.746 (0.664 - 0.822)	0.700 (0.639 - 0.761)	0.659 (0.597 - 0.726)
LightGBM	External validation	0.769 (0.751 - 0.792)	0.680 (0.660 - 0.701)	0.670 (0.640 - 0.710)	0.685 (0.658 - 0.714)	0.611 (0.583 - 0.643)

[a]Data set sizes were 753 for training, 321 for internal testing, and 1651 for external validation.

[b]AUROC: area under the receiver operating characteristic curve.

[c]Classification metrics were calculated using model-specific training-derived operating points selected by maximizing the Youden index in the training set: RF, 0.545; SVM, 0.316; LR, 0.343; DT, 0.429; XGBoost, 0.439; NB, 0.208; and LightGBM, 0.413. These operating points were applied unchanged to the internal test and external validation sets and were used for performance evaluation rather than as validated clinical decision cutoffs. Values are presented to 3 decimal places; full-precision values were used in the analyses. PPV, NPV, and full confusion-matrix counts are reported in Table S7 in Multimedia Appendix 1.

[d]RF: random forest.

[e]SVM: support vector machine.

XSL•FO
RenderX

[f]LR: logistic regression.

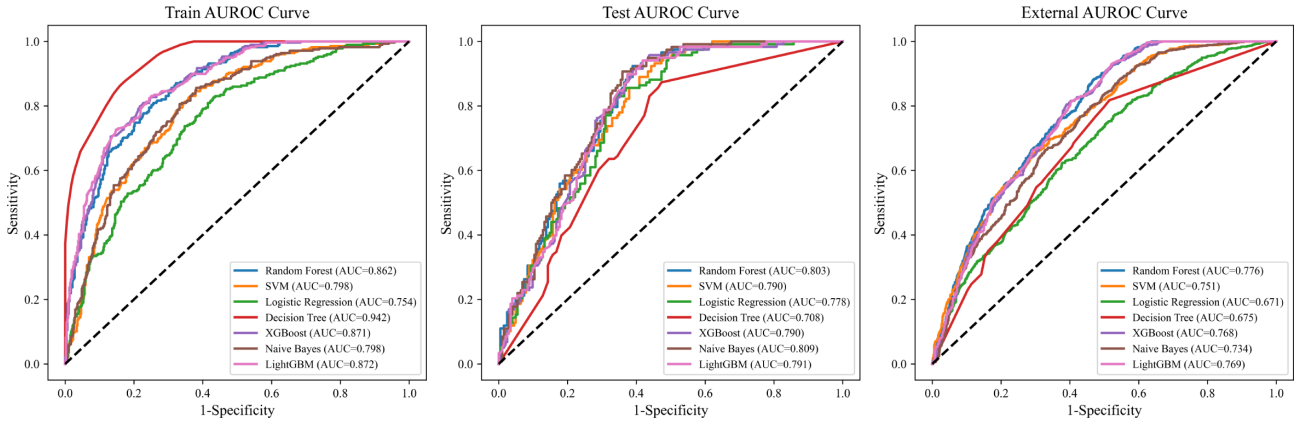
[g]DT: decision tree.

[h]XGBoost: extreme gradient boosting.

[i]NB: Naive Bayes.

[j]LightGBM: light gradient boosting machine.

Figure 2. Area under receiver operating characteristic curves of the 7 candidate machine learning models in the training, internal test, and external validation sets. AUROC: area under receiver operating characteristic; LightGBM: light gradient boosting machine; SVM: support vector machine; XGBoost: extreme gradient boosting.

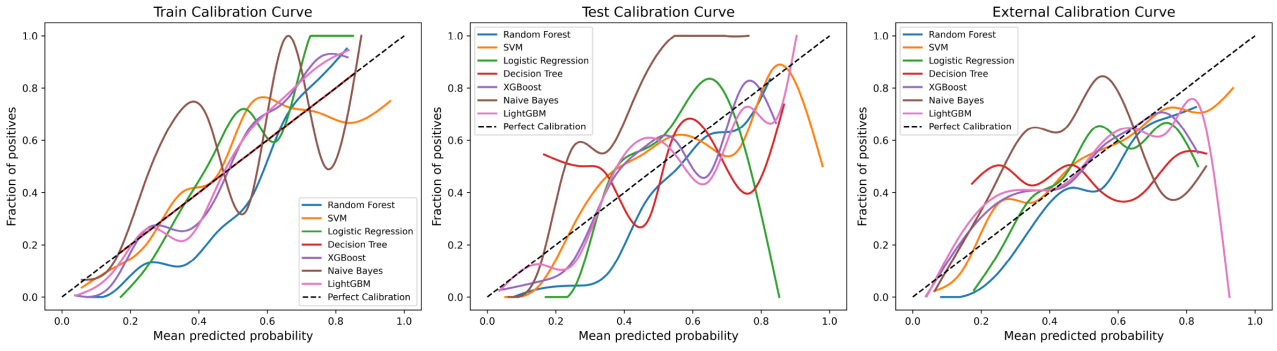


Calibration and Clinical Utility

Calibration curves are shown in Figure 3, and quantitative calibration statistics, including the Brier score, calibration intercept, calibration slope, and observed-to-expected ratio, are reported in Multimedia Appendix 1, Table S8 in Multimedia Appendix 1. Corresponding calibration results from the sensitivity analyses are provided in Table S10 in Multimedia Appendix 1. For the RF model, the Brier score was 0.179 in the internal test set and 0.190 in the external validation set. The calibration intercepts were −0.176 and −0.120, the calibration slopes were 1.161 and 1.061, and the observed-to-expected ratios were 0.806 and 0.803, respectively. These estimates indicated reasonably stable calibration across the 2 validation sets, although the CIs should be considered when interpreting precision.

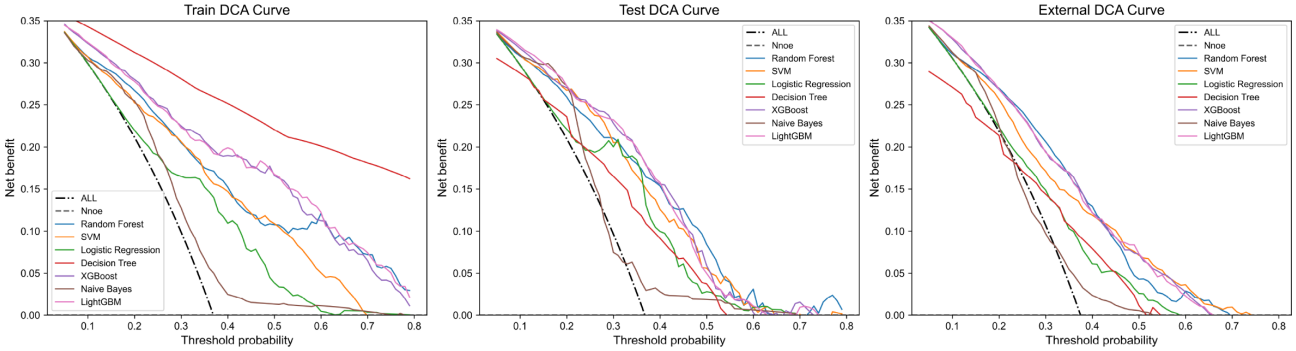
DCA is shown in Figure 4, and the corresponding model-specific net-benefit threshold-probability ranges are summarized in Table S9 in Multimedia Appendix 1. The lower bound of 0.05 was the prespecified lower limit of the display interval; these ranges were exploratory and should not be interpreted as validated clinical decision thresholds.

Figure 3. Calibration curves of the 7 candidate machine learning models in the training, internal test, and external validation sets. AUROC: area under receiver operating characteristic; LightGBM: light gradient boosting machine; SVM: support vector machine; XGBoost: extreme gradient boosting.



XSL•FO
RenderX

Figure 4. Decision curve analysis of the 7 candidate machine learning models in the training, internal test, and external validation sets. AUROC: area under the receiver operating curve; DCA: decision curve analysis; LightGBM: light gradient boosting machine; XGBoost: extreme gradient boosting.



Final Model Selection

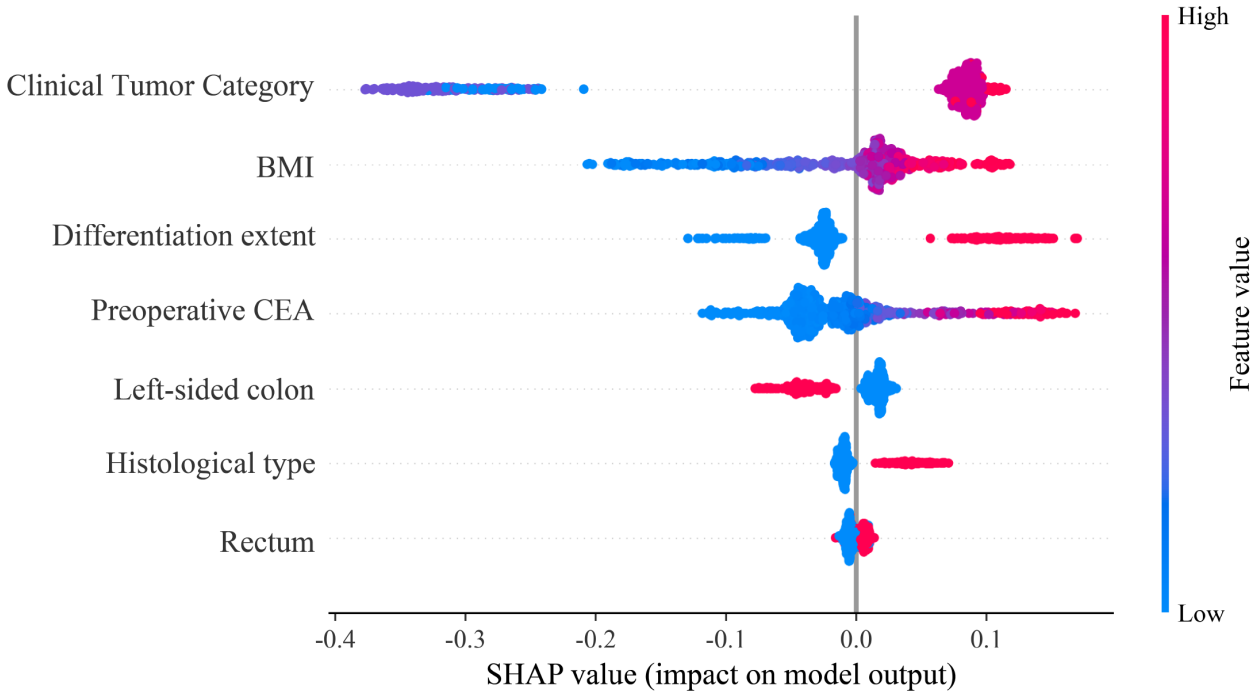
Considering discrimination, classification performance, calibration, and clinical utility in the internal test set, RF was selected as the final model before external validation. RF achieved the highest AUROC in the external validation set (0.776), while maintaining relatively balanced sensitivity (0.677) and specificity (0.693). Its calibration slope was close to 1 in both the internal test (1.161) and external validation (1.061) sets. Although the external-validation AUROC differences between RF and XGBoost or LightGBM were not statistically significant after Holm adjustment, RF showed a favorable overall balance across the prespecified internal-test criteria and was therefore selected for subsequent external validation and interpretability analyses.

Model Interpretation

Global Feature Importance

SHAP analysis showed that clinical T stage had the greatest overall contribution to the RF model, followed by BMI, tumor differentiation, pCEA, the left-sided colon indicator, histological type, and the rectum indicator (Figure S1 in Multimedia Appendix 1). The SHAP summary plot further showed that higher clinical T stage, BMI, poorer tumor differentiation, higher pCEA, and special histological type generally shifted model predictions toward LNM positivity (Figure 5). In contrast, the left-sided colon indicator was predominantly associated with negative SHAP values relative to the right-sided colon reference category.

Figure 5. Shapely additive explanations summary plot of the encoded features in the final random forest model. CEA: carcinoembryonic antigen; SHAP: Shapely additive explanations.



Individual Prediction and Feature Dependence

For the representative patient shown in Figure S2 in Multimedia Appendix 1, the RF model estimated an LNM probability of 0.804, compared with a baseline probability of 0.504. G3 or G4 differentiation, cT3 stage, elevated pCEA (22.9 ng/mL), and higher BMI (25.9 kg/m²) contributed positively to the prediction, with tumor differentiation showing the largest individual contribution.

SHAP dependence plots further demonstrated distinct patterns across predictors (Figure S3 in Multimedia Appendix 1). Higher

XSL•FO
RenderX

BMI and pCEA were generally associated with increasingly positive SHAP values. cT1 and cT2 contributed predominantly toward LNM-negative predictions, whereas cT3 and cT4 contributed toward LNM-positive predictions. G3 or G4 differentiation and mucinous adenocarcinoma or signet-ring cell carcinoma also showed predominantly positive contributions. For tumor location, the left-sided colon indicator was associated with lower SHAP values when present, whereas the rectum indicator showed comparatively smaller and more variable contributions.

Sensitivity Analyses

Sensitivity analyses showed that the RF model based on the 6 selected predictors performed similarly to the model incorporating all prespecified candidate predictors. The AUROC differences were not statistically significant in either the internal test set (0.803 vs 0.833; *P*=.15) or the external validation set (0.776 vs 0.787; *P*=.22). The all-candidate-predictor model showed slightly higher accuracy and lower Brier scores, but the overall performance profiles of the 2 models were comparable.

In contrast, the 6-predictor RF model showed significantly better discrimination than the model using clinical T stage alone, with AUROC differences of 0.091 in the internal test set and 0.069 in the external validation set (both *P*<.001). Calibration and decision curve analyses were consistent with improved probabilistic accuracy and broader exploratory net-benefit ranges for the multivariable model than for clinical T stage alone (Table S10 and Figure S4 in Multimedia Appendix 1).

Discussion

Principal Findings

In this multicenter retrospective study, we developed and externally validated a machine learning model for preoperative estimation of LNM risk in patients with CRC using routinely available clinical information. Six preoperative predictors—BMI, preoperative carcinoembryonic antigen (CEA), primary tumor site, clinical T stage, histological type, and tumor differentiation—were retained through the prespecified predictor-selection procedure. Among the 7 candidate algorithms, RF was selected as the final model based on its overall performance in the internal test set, with particular consideration of discrimination and the balance of threshold-dependent classification performance. The model achieved an AUROC of 0.803 in the internal test set and maintained an AUROC of 0.776 in the independent external validation cohort. In external validation, sensitivity and specificity were 0.677 and 0.693, respectively, while the calibration slope was 1.061 and the Brier score was 0.190, indicating that the model retained moderate discrimination with relatively stable classification and calibration performance when transported to an independent center.

An important finding was that the predictive information captured by the final model extended beyond clinical T stage alone. Although clinical T stage was the most influential predictor in the SHAP analysis, the clinical T stage–only RF model showed substantially lower discrimination than the final 6-predictor model in both the internal test set (AUROC 0.712 vs 0.803) and the external validation set (AUROC 0.707 vs 0.776). The corresponding AUROC differences of 0.091 and 0.069 were statistically significant in both datasets (both *P*<.001). These findings indicate that clinical T stage provides an important foundation for LNM risk estimation, but that BMI, preoperative CEA, tumor location, histological type, and tumor differentiation contribute additional complementary information for distinguishing individual risk.

Conversely, expanding the RF model from the selected 6 predictors to all prespecified candidate predictors did not result in a statistically significant improvement in discrimination. The all-predictor model yielded AUROCs of 0.833 and 0.787 in the internal test and external validation sets, respectively, compared with 0.803 and 0.776 for the 6-predictor model; the corresponding differences were not statistically significant (*P*=.15 and *P*=.22, respectively). Although the expanded model showed small numerical improvements in some performance measures, the absence of a clear gain in discrimination suggests that substantially increasing the number of predictors added limited incremental predictive information. The selected 6-predictor model therefore represents a relatively parsimonious approach that preserves most of the predictive information available from the broader set of routinely collected variables.

The SHAP analysis showed that the final RF model used information from several predictor domains. cT contributed most strongly, followed by BMI, tumor differentiation, pCEA, primary tumor site, and histological type. Higher cT, poorer tumor differentiation, higher BMI, and elevated pCEA generally shifted predictions toward a higher probability of LNM. These patterns describe model associations and should not be interpreted as causal effects or as evidence that the model explicitly identifies causal interactions.

Interpretation of Key Predictors

This finding is consistent with previous studies identifying deeper tumor invasion as a strong correlate of LNM in CRC [2,8,9]. The strong contribution of cT in the present model is therefore consistent with its established relationship to local tumor invasion and regional nodal spread.

Tumor differentiation was also an important predictor in the model. In our study, differentiation was assessed on preoperative endoscopic biopsy. Poor differentiation has been associated with an increased risk of LNM, particularly in early CRC [10]. However, limited tissue sampling may prevent preoperative biopsy from fully representing the tumor's overall differentiation. Barresi et al [11] reported concordant grading in approximately 75% of CRCs using a grading system based on poorly differentiated clusters, and the biopsy-derived grade was associated with nodal involvement and pTNM stage. Although their grading system differed from that used in our study, these findings support the clinical relevance of adverse histological features identified on preoperative biopsy. In our biopsy-resection analysis, tumor differentiation showed 77.0% (2074/2692) exact agreement, with a linearly weighted Cohen κ of 0.521, and 13.9% (374/2692) of tumors were undergraded on preoperative biopsy. Biopsy-reported differentiation therefore provides useful preoperative information, although it may not fully capture the highest-grade component of the tumor. Poor

XSL•FO
RenderX

differentiation on biopsy may serve as a meaningful high-risk feature, whereas a more favorable grade should be interpreted with awareness of sampling limitations.

Histological type showed a similar pattern. Mucinous adenocarcinoma and signet-ring cell carcinoma have distinct clinicopathological characteristics and are generally associated with more aggressive disease than conventional adenocarcinoma [12]. In our cohort, histological type showed 88.5% (2392/2703) exact agreement between preoperative biopsy and resection pathology, with a Cohen κ of 0.616. However, 42.5% (248/583) of tumors ultimately classified as mucinous adenocarcinoma or signet-ring cell carcinoma on the resection specimen had not been identified as a special histological type on preoperative biopsy. Previous studies have also reported that mucinous colorectal adenocarcinoma may be missed on initial endoscopic biopsy, reflecting the limitations of tissue sampling [13]. Importantly, despite incomplete biopsy-resection agreement, biopsy-derived histological type remained independently associated with LNM and contributed to the RF predictions, supporting its value as a routinely available preoperative predictor.

pCEA was independently associated with LNM, consistent with previous studies identifying elevated CEA as a risk factor for nodal involvement in CRC [14]. Our findings support pCEA as a routinely available preoperative marker for LNM risk prediction.

We found that right-sided colon and rectal tumors were associated with higher odds of LNM than left-sided colon tumors. CRCs arising at different anatomical sites show distinct clinicopathological and biological characteristics [15,16], which may partly explain the observed differences in nodal involvement. Our findings suggest that primary tumor site provides additional information for preoperative LNM risk prediction.

BMI was independently associated with LNM and ranked highly in the SHAP analysis. Obesity-related phenotypes have been linked to CRC development and progression [17,18], although BMI is only a simple proxy for overall adiposity. The observed association should not be interpreted as causal; BMI may instead function as a readily available patient-level marker correlated with other biological or clinical factors.

Taken together, these routinely available preoperative features capture complementary information on tumor extent, morphology, circulating markers, anatomical site, and patient characteristics, providing a broader basis for individualized LNM risk estimation.

Comparison With Previous Studies

Conventional imaging remains the foundation of preoperative staging in CRC, with CT commonly used for colon cancer and pelvic MRI for rectal cancer. Imaging provides important information on local tumor extent, as also reflected by the strong contribution of cT in our model. However, direct assessment of nodal metastasis still relies largely on lymph node size and morphological characteristics, which have limited accuracy for distinguishing metastatic from benign nodes. A meta-analysis of MRI-based nodal staging in rectal cancer reported only moderate diagnostic performance across different size and morphological criteria [19], while a recent systematic review similarly found limited performance for preoperative N staging in colon cancer using MRI or 18F-fluorodeoxyglucose positron emission tomography/computed tomography (18F-FDG PET/CT) [20]. These limitations provide a rationale for complementing conventional imaging with additional preoperative clinical and pathological information when estimating patient-level LNM risk.

Previous CRC metastasis prediction models have commonly included cancer stage, tumor differentiation, histological type, CEA, and demographic factors [21]. However, external validation has been limited: only 4 of 12 models identified in an umbrella review had undergone external validation, and calibration was incompletely reported [21]. More recent meta-analytic evidence also found higher reported AUROCs in studies without external validation [5]. Our model used routinely available preoperative variables and was evaluated in a large independent external cohort.

Quantitative imaging approaches, including radiomics and deep learning, aim to extract information beyond conventional visual assessment. A recent systematic review and meta-analysis reported promising performance of radiomics for preoperative LNM prediction in CRC, although substantial heterogeneity across studies remains [22]. Multimodal models further combine imaging-derived features with clinical variables to capture complementary information. However, the added value of these features may not always persist in external validation. In a multicenter rectal cancer study, Li et al [23] found that a clinical-radiomics model outperformed the clinical model in the training and internal validation cohorts, but this advantage was not maintained in the external cohort.

Taken together, previous studies illustrate a trade-off between the amount of predictive information incorporated and the complexity of data acquisition and model implementation. Differences in study populations, predictor sets, imaging modalities, and validation strategies make direct comparison of reported AUROCs across studies difficult [5,21,22]. In this context, our study focused on routinely available preoperative information, including clinical staging, laboratory markers, and biopsy-derived pathological features, without requiring radiomic feature extraction or additional molecular testing. The selected model was evaluated in a large independent external cohort, and sensitivity analyses showed that the 6-predictor model provided additional discriminatory information beyond cT alone, whereas expanding the model to all candidate predictors did not significantly improve discrimination. These findings support a relatively parsimonious model for preoperative LNM risk estimation, but they do not establish superiority over every existing clinical model.

Clinical Implications and Intended Use

Preoperative nodal assessment remains imperfect in CRC, leaving uncertainty even after conventional imaging and clinical staging. Our model addresses this uncertainty by integrating routinely available clinical, laboratory, imaging-derived, and biopsy-based information into an individualized estimate of pathological LNM risk. Individualized risk estimation is an

XSL•FO
RenderX

important application of clinical prediction models because it translates multiple patient characteristics into a probability that can be interpreted at the patient level [24]. In our study, the 6-predictor model also provided additional discriminatory information beyond cT alone, suggesting that routinely collected features can complement conventional assessment of local tumor extent.

A practical advantage of the model is that all predictors are already available during routine preoperative assessment and do not require dedicated radiomic feature extraction, molecular assays, or additional specialized examinations. This may simplify future evaluation in routine-care settings, although the completeness and reliability of these inputs, as well as transportability across centers, must be confirmed before implementation [25,26].

The model was developed in patients with newly diagnosed CRC who proceeded directly to radical surgery without preoperative anticancer treatment, and its intended use is therefore preoperative LNM risk stratification in a similar direct-surgery population. Its performance has not been validated in patients being considered for endoscopic resection, local excision, neoadjuvant therapy, or nonoperative management. The predicted probability is not intended to function as an isolated treatment criterion. A possible future workflow would be to calculate the probability after routine preoperative data collection and initial imaging review, then use discordance between the model estimate and the clinical assessment to prompt verification of the input data, focused radiology or pathology review, or multidisciplinary discussion. This workflow is a proposed implementation scenario for prospective evaluation, not a validated decision rule.

Strengths and Limitations

This study has several strengths. It included a large multicenter cohort with independent external validation, and all predictors were routinely available before surgery without requiring radiomics, molecular testing, or additional specialized examinations. Model evaluation included discrimination, calibration, DCA, threshold-dependent performance, and SHAP interpretation. Biopsy-resection agreement and sensitivity analyses against cT alone and the full predictor set provided additional evidence regarding predictor reliability and model robustness.

Several limitations should be acknowledged. First, the retrospective design and inclusion of only 2 centers in China may limit generalizability. Second, eligibility was restricted to patients with newly diagnosed CRC who underwent direct radical resection with regional lymph node dissection without preoperative anticancer treatment; therefore, the findings should not be extrapolated to patients undergoing neoadjuvant therapy, endoscopic resection, local excision, or nonoperative management. Patients with unavailable postoperative pathological LNM status were excluded, which may have introduced selection bias. Third, clinical regional lymph node category (cN category) was not included in the prespecified predictor set; therefore, the model was not directly benchmarked against clinical regional lymph node category-based assessment. The model also depends partly on imaging-derived cT, and its performance may vary with imaging quality, reader expertise, staging protocols, laboratory measurement, pathology processing, lymph-node yield, and biopsy practices across centers. Biopsy-derived differentiation and histological type showed imperfect agreement with resection pathology and remain subject to sampling limitations. Comorbidities were available only as a binary clinical-record variable without harmonized disease-specific definitions across centers. Early-stage tumors, particularly cT1-cT2 disease, were relatively underrepresented, and subgroup or fairness analyses were not performed. Finally, the training-derived evaluation operating point, DCA ranges, and threshold-dependent classifications were exploratory; no decision-impact or patient-outcome benefit was established. These factors define the current validation boundaries of the model: it has been evaluated in a direct-surgery population from 2 centers in China, with limited representation of early-stage disease, and its performance has not been established in other treatment pathways or broader clinical settings. Prospective multicenter validation across diverse populations and settings is required.

Conclusions

We developed and externally validated a 6-predictor RF model for preoperative LNM risk estimation in patients with newly diagnosed CRC who proceed directly to radical resection with regional lymph node dissection without preoperative anticancer treatment. The model used routinely available clinical, laboratory, imaging-derived, and biopsy-based information and showed moderate discrimination in an independent external cohort, with predictive information beyond cT alone. Its intended role is adjunctive, research-stage risk stratification rather than replacement of imaging, pathology, or clinical judgment. Further prospective validation in broader populations, particularly patients with early-stage disease and those managed through different treatment pathways, is required before clinical implementation.

Acknowledgments

The authors thank all members of the study team involved in data collection, data management, and clinical coordination. No GenAI tools were used in the preparation, writing, or editing of this manuscript.

Funding

This study was supported by the National Science and Technology Innovation 2030 Major Program (No. 2025ZD0544900); the Innovation Team and Talents Cultivation Program of the National Administration of Traditional Chinese Medicine (No. ZYYCXTD-C-202208); and NATCM's Project of High-level Construction of Key TCM Disciplines (Document No. GZYYRJH

XSL•FO
RenderX

[2023] 85). The funders had no role in study design, data collection, data analysis, data interpretation, manuscript preparation, or the decision to submit the manuscript for publication.

Data Availability

The fitted final RF model and a prediction script implementing the 6-predictor model are provided in Multimedia Appendix 2. The implementation is intended for patients with newly diagnosed CRC who proceed directly to radical resection with regional lymph node dissection without preoperative anticancer treatment and has not been validated for patients managed with endoscopic resection, local excision, neoadjuvant therapy, or nonoperative treatment. Complete values for all 6 predictors are required because the released implementation does not include the original training-time imputation object and does not perform missing-value imputation. Predictor definitions, units, accepted categories, encoding rules, and software requirements are provided in the accompanying README. The script returns the predicted probability of LNM and, for reproducibility of the reported threshold-dependent analyses, a binary classification based on the training-derived Youden operating point of 0.545. This operating point was used for performance evaluation and is not a validated clinical decision cutoff. Additional analysis code used for model development and evaluation is available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: GX, XS, HC.
Data curation: GX, XS, JJ, JS, LL, YZ, HL.
Formal analysis: GX, XS.
Funding acquisition: HC.
Literature review: JJ.
Machine learning model development: GX, XS.
Methodology: GX, XS, JJ, HC.
Supervision: HC.
Validation: JS, LL, YZ, HL.
Visualization: GX, XS.
Writing – original draft: GX, XS.
Writing – review & editing: JJ, JS, LL, YZ, HL, HC.
GX and XS contributed equally to this work and share first authorship.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables and figures for model development, validation, interpretation, and sensitivity analyses.
[DOCX File, 3245 KB - cancer_v12i1e101162_app1.docx]

Multimedia Appendix 2
Fitted random forest model, prediction script, README, and software requirements.
[PDF File, 17 KB - cancer_v12i1e101162_app2.pdf]

References

1. Sung H, Filho AM, Laversanne M, et al. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries. CA Cancer J Clin 2026;76(4):e70090. [doi: 10.3322/caac.70090] [Medline: 42417444]
2. Zhu X, Lin SQ, Xie J, et al. Biomarkers of lymph node metastasis in colorectal cancer: update. Front Oncol 2024;14:1409627. [doi: 10.3389/fonc.2024.1409627] [Medline: 39328205]
3. Takayama Y, Shimizu T, Watanabe J, Ogawa S, Kanemitsu Y. Diagnostic accuracy of size-based preoperative CT assessment for predicting lymph node metastasis in colon cancer: a systematic review and meta-analysis. Ann Gastroenterol Surg 2026 Sep;10(5):1450-1462. [doi: 10.1002/ags3.70218] [Medline: 42495671]
4. Salmerón-Ruiz A, Luengo Gómez D, Medina Benítez A, Láinez Ramos-Bossini AJ. Primary staging of rectal cancer on MRI: an updated pictorial review with focus on common pitfalls and current controversies. Eur J Radiol 2024 Jun;175:111417. [doi: 10.1016/j.ejrad.2024.111417] [Medline: 38484688]
5. Abbaspour E, Mansoori B, Karimzadhagh S, et al. Machine learning and deep learning models for preoperative detection of lymph node metastasis in colorectal cancer: a systematic review and meta-analysis. Abdom Radiol (NY) 2025 May;50(5):1927-1941. [doi: 10.1007/s00261-024-04668-z] [Medline: 39522103]

XSL•FO
RenderX

6. Keel B, Quyn A, Jayne D, Relton SD. State-of-the-art performance of deep learning methods for pre-operative radiologic staging of colorectal cancer lymph node metastasis: a scoping review. BMJ Open 2024 Dec 2;14(12):e086896. [doi: 10.1136/bmjopen-2024-086896] [Medline: 39622569]
7. Bedrikovetski S, Dudi-Venkata NN, Kroon HM, et al. Artificial intelligence for pre-operative lymph node staging in colorectal cancer: a systematic review and meta-analysis. BMC Cancer 2021 Sep 26;21(1):1058. [doi: 10.1186/s12885-021-08773-w] [Medline: 34565338]
8. Niu X, Cao J. Predicting lymph node metastasis in colorectal cancer patients: development and validation of a column chart model. Updates Surg 2024 Aug;76(4):1301-1310. [doi: 10.1007/s13304-024-01884-6] [Medline: 38954377]
9. Xia C, Liu F, Xia C, et al. Machine learning to develop and validate a model for predicting the risk of lymph node metastasis in colorectal cancer patients. Front Oncol 2026;16:1860520. [doi: 10.3389/fonc.2026.1860520] [Medline: 42553131]
10. Dykstra MA, Gimon TI, Ronksley PE, Buie WD, MacLean AR. Classic and novel histopathologic risk factors for lymph node metastasis in T1 colorectal cancer: a systematic review and meta-analysis. Dis Colon Rectum 2021 Sep 1;64(9):1139-1150. [doi: 10.1097/DCR.0000000000002164] [Medline: 34397562]
11. Barresi V, Bonetti LR, Ieni A, Branca G, Baron L, Tuccari G. Histologic grading based on counting poorly differentiated clusters in preoperative biopsy predicts nodal involvement and pTNM stage in colorectal cancer patients. Hum Pathol 2014 Feb;45(2):268-275. [doi: 10.1016/j.humpath.2013.07.046] [Medline: 24289972]
12. Fadel MG, Malietzis G, Constantinides V, Pellino G, Tekkis P, Kontovounisios C. Clinicopathological factors and survival outcomes of signet-ring cell and mucinous carcinoma versus adenocarcinoma of the colon and rectum: a systematic review and meta-analysis. Discov Onc 2021 Dec;12(1). [doi: 10.1007/s12672-021-00398-6]
13. Xiao S, Huang J, Zhang Y, et al. Endoscopy biopsy is not efficiency enough for diagnosis of mucinous colorectal adenocarcinoma. Discov Oncol 2021 Oct 25;12(1):44. [doi: 10.1007/s12672-021-00443-4] [Medline: 35201497]
14. Xiong X, Wang C, Cao J, Gao Z, Ye Y. Lymph node metastasis in T1-2 colorectal cancer: a population-based study. Int J Colorectal Dis 2023 Apr 13;38(1):94. [doi: 10.1007/s00384-023-04386-w] [Medline: 37055602]
15. Benlice C, Elhan AH, Seker ME, Gorgun E, Kuzu MA. Oncologic outcomes and trends in each colon cancer location and stages over the last two decades: insights from the SEER registry. Tech Coloproctol 2024 Nov 2;28(1):147. [doi: 10.1007/s10151-024-03020-3] [Medline: 39487239]
16. Ciepiela I, Szczepaniak M, Ciepiela P, et al. Tumor location matters, next generation sequencing mutation profiling of left-sided, rectal, and right-sided colorectal tumors in 552 patients. Sci Rep 2024 Feb 26;14(1):4619. [doi: 10.1038/s41598-024-55139-w] [Medline: 38409377]
17. Gonzalez-Gutierrez L, Motiño O, Barriuso D, et al. Obesity-associated colorectal cancer. Int J Mol Sci 2024 Aug 14;25(16):8836. [doi: 10.3390/ijms25168836] [Medline: 39201522]
18. Mandic M, Li H, Safizadeh F, Niedermaier T, Hoffmeister M, Brenner H. Is the association of overweight and obesity with colorectal cancer underestimated? An umbrella review of systematic reviews and meta-analyses. Eur J Epidemiol 2023 Feb;38(2):135-144. [doi: 10.1007/s10654-022-00954-6] [Medline: 36680645]
19. Zhuang Z, Zhang Y, Wei M, Yang X, Wang Z. Magnetic resonance imaging evaluation of the accuracy of various lymph node staging criteria in rectal cancer: a systematic review and meta-analysis. Front Oncol 2021;11:709070. [doi: 10.3389/fonc.2021.709070] [Medline: 34327144]
20. Sikkenk DJ, Henskens IJ, van de Laar B, et al. Diagnostic performance of MRI and FDG PET/CT for preoperative locoregional staging of colon cancer: systematic review and meta-analysis. AJR Am J Roentgenol 2024 Nov;223(5):e2431440. [doi: 10.2214/AJR.24.31440] [Medline: 39230407]
21. Xu W, He Y, Wang Y, et al. Risk factors and risk prediction models for colorectal cancer metastasis and recurrence: an umbrella review of systematic reviews and meta-analyses of observational studies. BMC Med 2020 Jun 26;18(1):172. [doi: 10.1186/s12916-020-01618-6] [Medline: 32586325]
22. Abbaspour E, Karimzadhagh S, Monsef A, Joukar F, Mansour-Ghanaei F, Hassanipour S. Application of radiomics for preoperative prediction of lymph node metastasis in colorectal cancer: a systematic review and meta-analysis. Int J Surg 2024 Jun 1;110(6):3795-3813. [doi: 10.1097/JS9.0000000000001239] [Medline: 38935817]
23. Li H, Chen XL, Liu H, Lu T, Li ZL. MRI-based multiregional radiomics for predicting lymph nodes status and prognosis in patients with resectable rectal cancer. Front Oncol 2022;12:1087882. [doi: 10.3389/fonc.2022.1087882] [Medline: 36686763]
24. Beaulieu-Jones BK, Yuan W, Brat GA, et al. Machine learning for patient risk stratification: standing on, or looking over, the shoulders of clinicians? NPJ Digit Med 2021 Mar 30;4(1):62. [doi: 10.1038/s41746-021-00426-3] [Medline: 33785839]
25. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ 2024 Jan 8;384:e074819. [doi: 10.1136/bmj-2023-074819] [Medline: 38191193]
26. Santos CS, Amorim-Lopes M. Externally validated and clinically useful machine learning algorithms to support patient-related decision-making in oncology: a scoping review. BMC Med Res Methodol 2025 Feb 21;25(1):45. [doi: 10.1186/s12874-025-02463-y] [Medline: 39984835]

XSL•FO
RenderX

Abbreviations

18F-FDG PET/CT: 18F-fluorodeoxyglucose positron emission tomography/computed tomography
AGR: albumin-to-globulin ratio
AUROC: area under the receiver operating characteristic curve
CEA: carcinoembryonic antigen
CRC: colorectal cancer
CT: computed tomography
cT: clinical tumor stage
cT1: clinical Tumor1 stage
cT2: clinical Tumor2 stage
cT3: clinical Tumor3 stage
cT4: clinical Tumor4 stage
DCA: decision curve analysis
DT: decision tree
GLOBOCAN: Global Cancer Observatory
LightGBM: light gradient-boosting machine
LNM: lymph node metastasis
LR: logistic regression
MRI: magnetic resonance imaging
NB: naive Bayes
NPV: negative predictive value
OR: odds ratio
pCA19-9: preoperative carbohydrate antigen 19‐9
pCEA: preoperative carcinoembryonic antigen
PPV: positive predictive value
RF: random forest
ROC: receiver operating characteristic
SHAP: Shapley additive explanations
SII: systemic immune-inflammation index
SVM: support vector machine
XGBoost: extreme gradient boosting

Edited by M Balcarras; submitted 12.May.2026; peer-reviewed by MH Tsai, S Das; revised version received 01.Sep.2026; accepted 02.Sep.2026; published 07.Oct.2026.

Please cite as:
Xu G, Sun X, Jiang J, Sun J, Li L, Cheng H, Zhu Y, Liu H
Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer Using Routinely Collected Clinical Data: Multicenter Retrospective Machine Learning Model Development and External Validation Study
JMIR Cancer 2026;12:e101162
URL: https://cancer.jmir.org/2026/1/e101162
doi:10.2196/101162

© Gehui Xu, Xiaohe Sun, Jiaxin Jiang, Jin Sun, Liu Li, Haibo Cheng, Yuekun Zhu, Haiyi Liu. Originally published in JMIR Cancer (https://cancer.jmir.org), 7.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Psychosocial Distress and Radiation Treatment Interruptions in Patients With Cancer Using Machine Learning and Regression: Retrospective Cohort Study

Soheil Hashtarkhani[1,2], PhD; Christopher L Brett[3], MD; Fekede Asefa Kumsa[4], PhD; Rezaur Rashid[1], PhD; Parnian Kheirkhah Rahimabad[4], MD, PhD; Saba Kheirinejad[1], PhD; Robert L Davis[1], MD, MPH; David L Schwartz[4,5*], MD; Arash Shaban-Nejad[1*], MPH, PhD

[1]Center for Biomedical Informatics, College of Medicine, University of Tennessee Health Science Center, 50 N Dunlap Street, Memphis, TN, United States

[2]Health Equity Program, Rhodes College, Memphis, TN, United States

[3]University of Tennessee Graduate School of Medicine, Knoxville, TN, United States

[4]Department of Radiation Oncology, University of Tennessee Health Science Center, College of Medicine, Memphis, TN, United States

[5]Department of Preventive Medicine, University of Tennessee Health Science Center, College of Medicine, Memphis, TN, United States

[*]these authors contributed equally

Corresponding Author:
Arash Shaban-Nejad, MPH, PhD
Center for Biomedical Informatics, College of Medicine, University of Tennessee Health Science Center, 50 N Dunlap Street, Memphis, TN, United States

Abstract

Background: Uninterrupted delivery of radiation therapy is critical to cancer treatment quality. Unplanned radiation treatment interruptions compromise tumor control and survival. Although previous studies have explored clinical and sociodemographic factors influencing treatment adherence, the role of psychosocial distress and related symptoms remains unclear.

Objective: This study aims to identify patient-reported psychosocial predictors of radiation treatment interruptions from patient-reported distress screening data integrated with clinical and socioeconomic correlates via traditional and machine learning–enhanced modeling.

Methods: This institutional cohort study included 1953 patients who underwent external beam radiation therapy between August 2020 and May 2024. Patient demographics, clinical characteristics, treatment details, and baseline distress levels were extracted from a hospital-level electronic health record data repository. Distress was assessed using the Distress Thermometer developed by the National Comprehensive Cancer Network. These data were linked with census tract–level socioeconomic indicators. Predictors of treatment interruption (measured as missed treatment days) were evaluated using negative binomial regression and an extreme gradient boosting (XGBoost) model with a Poisson objective function.

Results: The study cohort missed a mean of 1.19 (SD 2.18) treatment days (range: 0-15). Higher distress scores, the presence of learning or communication barriers, and genitourinary malignancy were significantly associated with increased interruption risk in the negative binomial regression model. The XGBoost model confirmed these features as top contributors to prediction performance, while yielding an incrementally lower root mean squared error of 2.19 (vs 2.21 for the regression model).

Conclusions: Patient-reported distress and health education barriers were modestly associated with treatment interruptions, underscoring the importance of incorporating psychosocial assessments into radiation oncology workflows. Both parametric and machine learning models provided stable and closely aligned predictions of missed radiation treatment days. Key predictors identified as statistically significant in the regression model were also among the top contributors in the XGBoost model, reinforcing the robustness of these findings and supporting integration of both approaches into future intervention strategies.

(JMIR Cancer 2026;12:e92899) doi:10.2196/92899

KEYWORDS

radiotherapy; cancer; treatment interruption; risk factors; psychosocial distress; machine learning; extreme gradient boosting; XGBoost; negative binomial regression

XSL•FO
RenderX

Introduction

Radiation therapy is a fundamental component of cancer management, with more than half of all patients requiring radiotherapy during their treatment course [1]. The effectiveness of radiation therapy depends on consistent treatment delivery, as unplanned interruptions have been associated with reduced tumor control, disease recurrence, and decreased overall survival [2,3]. Despite ongoing technological and procedural advances, radiation treatment interruptions (RTIs) remain a major challenge in oncology care. These interruptions arise from multiple factors, including technical and operational issues [4,5], as well as patient-specific factors such as poor performance status, treatment-related toxicities, and unplanned hospitalizations [6]. Social determinants of health further compound these challenges, as individuals with limited financial means, lower education, or unstable support systems are more likely to experience barriers to treatment completion [7].

While previous studies have identified several demographic, clinical, and socioeconomic factors associated with RTIs [8-10], less is known about how psychological and symptom-related domains interact with these traditional predictors to influence treatment adherence. Cancer-related distress, encompassing emotional, physical, and practical challenges, is a prevalent yet often underrecognized factor that may directly contribute to treatment noncompliance [11,12]. Patients experiencing high levels of distress or other psychosocial difficulties may have challenges maintaining daily treatment schedules due to fatigue, anxiety, or logistical stressors, leading to preventable treatment delays. We integrated regression-based analyses with machine learning methods to examine the influence of psychological and distress-related symptoms, along with other key variables, on RTIs.

The present study aims to comprehensively evaluate the predictors of RTIs within a Central Appalachian micropolitan and rural catchment region. By integrating patient-reported distress screening data with clinical and socioeconomic information, this analysis sought to capture the multidimensional influences on treatment adherence. The overarching goal was to identify key risk factors that may inform targeted interventions, improve continuity of care, and ultimately enhance the quality and equity of radiation oncology delivery in underserved and geographically diverse communities.

Methods

Dataset

Demographic, clinical, and treatment-related data, including age, sex, race, cancer site, insurance type, and marital status, were obtained from the centralized oncology data repository at the University of Tennessee Medical Center in Knoxville, Tennessee. The analytic cohort included patients who received external beam radiation therapy between August 2020 and May 2024. This cohort draws from the same institutional data source as our group's prior analysis of demographic and clinical predictors of treatment interruption [9]; the present study extends that work by incorporating patient-reported psychosocial distress screening data and census tract–level socioeconomic indicators, which were not examined previously.

Patients undergoing treatment were asked to complete a standardized distress screening tool at baseline before starting radiation therapy. This instrument canvassed emotional, physical, practical, and nutritional concerns, along with a self-reported distress score ranging from 0 ("no distress") to 10 ("extreme distress"). Paper-based forms were reviewed by nursing staff with the patient, with the distress score and pertinent findings entered into an electronic form within the medical record, enabling integration of psychological and symptom-related measures with clinical and demographic records.

To complement patient-level data, area-level social determinants of health were appended using data from the US Census Bureau's American Community Survey 2019 - 2023 [13], aggregated at the census tract level. Socioeconomic indicators included measures such as median household income and educational attainment in patients' home census tracts, enabling contextual assessment of neighborhood characteristics.

Preprocessing and Cohort Selection

A total of 3134 consecutive patients underwent radiation therapy during the study period. Several data-quality and eligibility criteria were applied to define the analytic cohort. First, 269 patients were excluded for missing more than fifteen treatment days at the data extraction stage. Interruptions of this length were considered unlikely to represent a single episodic treatment interruption and more often reflect distinct clinical scenarios, such as sequential courses of radiotherapy for different disease sites; consequently, detailed records for these patients were not retained for further characterization. Patients receiving treatment courses lasting fewer than 5 days (n=599) or with a total delivered dose below 2000 cGy (n=84) were also removed, as these treatment patterns generally correspond to very short palliative or stereotactic regimens in which the clinical significance of interruptions differs from that of conventional fractionation. Additional exclusions were made for missing demographic or administrative variables (52 cases). These steps resulted in 2130 patients with complete information. Among these individuals, 177 patients did not complete a baseline distress screening form and could not be included in analyses incorporating psychological variables, resulting in a final analytic sample of 1953.

Categorical variables were harmonized and standardized, with missing categories explicitly coded to maintain transparency in downstream modeling. All continuous variables and clinical indicators were cleaned using context-specific rules to preserve interpretability. Socioeconomic and demographic measures were grouped into broader categories to reduce data sparsity.

Multicollinearity among predictors was evaluated using the variance inflation factor (VIF) and generalized VIF (GVIF) for categorical predictors with multiple levels. All predictors exhibited acceptable collinearity levels (GVIF-adjusted statistics <10, most <2.5), indicating no significant multicollinearity concerns. Therefore, all candidate variables were retained for inclusion in subsequent modeling, including both negative

XSL•FO
RenderX

binomial regression and extreme gradient boosting (XGBoost) analyses.

Study Outcome Measures

The primary outcome of this study was the number of missed radiation treatment days. Missed treatment days were calculated based on the anticipated versus actual completion dates of therapy, adjusted for weekends and federal holidays. This count variable, which fell within a 0 to 15 day range, represents the total number of treatment days not completed as scheduled.

Automated Variable Categorization Using Large Language Models

Several extracted variables were available only in unstructured text format and required additional processing before they could be used for analysis. For example, *International Classification of Diseases* (*ICD*) diagnosis codes and insurance providers needed to be mapped into standardized categorical groups. In addition, approximately 7% of the distress screening forms included free-text entries by nursing staff rather than the designated radio buttons for practical, emotional, physical, and nutritional concerns.

To address this, we incorporated GPT-4o APIs into the preprocessing pipeline. GPT-4o was selected based on our group's prior comparative evaluation of 5 language models (GPT-3.5, GPT-4o, Llama 3.2, Gemini 1.5, and BioBERT) for clinical text categorization tasks in an oncology setting, in which GPT-4o and BioBERT outperformed the other models [14]. The present study required categorization of a structured variable (*International Classification of Diseases*, 10th Revision [*ICD*-10], with over 50 unique values), a semistructured variable (insurance, with over 700 unique values), and an unstructured variable (free-text distress screening entries).

The categorization framework and mapping schema were developed in consultation with 2 radiation oncology specialists. A few-shot prompting strategy was used, in which the model was provided with a small set of representative examples to guide consistent categorization for each variable type. LLM-generated categorizations were subsequently reviewed by our 2 oncology coauthors, allowing us to both correct any misclassifications and assess the feasibility of using LLM-based categorization independently, without manual review, in future prediction pipelines.

Model Generation

Missed radiation treatment day values were heavily right-skewed, with a large proportion of zero counts and a long upper tail. This distribution confirmed that the data were discrete and overdispersed, warranting the use of count-based regression rather than continuous linear models.

The dataset was divided into training (391/1953, 80%) and testing (391/1953, 20%) subsets, with categorical variables one-hot encoded for the machine learning algorithm. To assess the stability of performance estimates given this single split, we additionally conducted 100 repeated stratified resamples preserving the distribution of missed treatment days across train and test sets, and calculated mean, SD, and 95% CIs (2.5th-97.5th percentile) for root mean square error (RMSE), mean absolute error (MAE), and Poisson deviance across resamples.

A negative binomial regression model was fitted to the training data to account for overdispersion, as the initial Poisson model exhibited a dispersion statistic of 3.69. The negative binomial model estimated incidence rate ratios (IRRs) for each covariate, providing interpretable measures of relative change in missed treatment days associated with each factor. To ensure model parsimony while preserving interpretability, the full negative binomial model including all candidate covariates was first examined. Backward elimination was then applied as a sensitivity approach to remove nonsignificant predictors ($P \geq .05$); this process did not materially alter effect estimates or the substantive conclusions of the model. Given the large proportion of zero outcomes, we additionally evaluated a zero-inflated negative binomial model as part of model selection; however, the zero-inflation component was not statistically significant, predictive performance on the hold-out test set did not improve, and a Vuong test favored the standard negative binomial formulation. Therefore, the conventional negative binomial model was retained for all primary analyses.

In parallel, an XGBoost model was trained on the same training data using a Poisson objective function to handle count-based outcomes. Hyperparameters, including tree depth, learning rate, subsampling rate, and column sampling, were tuned through internal validation and early stopping to prevent overfitting. Both models were subsequently applied to the test dataset to generate predicted counts and assess generalization performance.

Model Comparison

Predictive accuracy was assessed using multiple complementary metrics—RMSE, MAE, and Poisson deviance—to capture different aspects of model fit and prediction error. All performance metrics were calculated on a naive 20% hold-out test set, which was separated from the data prior to model training or hyperparameter tuning to ensure an unbiased evaluation of out-of-sample performance. In addition to evaluating predictive performance, a direct comparison of these two modeling approaches enabled integration of statistical inference and machine learning–based feature importance.

Ethical Considerations

This study was reviewed and approved by the Institutional Review Board of the University of Tennessee Health Science Center for analysis without personal identifiers (protocol: 16 - 04888-XP).

Results

Descriptive Analysis

Table 1 presents the demographic, socioeconomic, and clinical characteristics of the study cohort. The total number of RTIs ranged from 0 to 15 days, with a mean of 1.19 (SD 2.18) days (median 0, IQR 0 - 1). Study patients had a mean age of 64.5 (SD 12.1) years and were predominantly White (1837/1953, 94.1%), with smaller proportions identifying as Black/African American (88/1953, 4.5%) or other/minority (28/1953, 1.4%). The cohort included 1017 (52.1%) females and most participants

XSL•FO
RenderX

were married (1130/1953, 57.9%) and received health insurance coverage through Medicare (647/1953, 33.1%) or commercial payers (776/1953, 39.7%). Approximately one-third of patients resided in low-income (596/1953, 30.5%) or low-education (589/1953, 30.2%) areas.

XSL•FO
RenderX

Table . Characteristics of the study cohort (N=1953).

Variable	Description	Range/coding	Distribution
Outcome			
Missing treatment days (outcome)	Number of missed radiation treatment sessions compared to prescribed schedule	0 - 15 days	• Mean 1.19 (SD 2.18) • Median 0 (IQR 0 - 1)
Distress and symptom measures			
Distress score	Self-reported oncology distress thermometer score assessing psychological distress during treatment	0 - 10	• Mean 3.1 (SD 2.9) • Median 3 (IQR 0-5; range 0 - 10)
Emotional symptom count	Number of emotional or psychological concerns selected (eg, depression, sadness, fears, and nervousness)	0="no concern" and 1 - 6="number of emotional issues"	• 0: 1769 (90.6%) • 1: 66 (3.4%) • 2: 52 (2.7%) • 3: 19 (1.0%) • 4: 15 (0.8%) • 5: 16 (0.8%) • 6: 16 (0.8%)
Practical symptom count	Number of practical or logistical concerns endorsed (eg, financial, transportation, and insurance issues)	0="none" and 1 - 5="number of problems"	• 0: 1855 (95.0%) • 1: 54 (2.8%) • 2: 31 (1.6%) • 3: 5 (0.3%) • 4: 3 (0.2%) • 5: 5 (0.3%)
Physical symptom count	Number of physical concerns (eg, fatigue, sleep disturbance, and mobility problems)	0 - 13	• 0: 1738 (89.0%) • 1: 85 (4.4%) • 2: 29 (1.5%) • 3: 21 (1.1%) • 4: 23 (1.2%) • ≥5: 57 (2.9%)
Nutrition symptom count	Number of nutrition-related problems (eg, nausea, appetite loss, and diarrhea)	0 - 5	• 0: 1847 (94.6%) • 1: 64 (3.3%) • 2: 25 (1.3%) • 3: 13 (0.7%) • ≥4: 5 (0.3%)
Primary pain intensity	Self-reported primary pain score at assessment	0 - 10 scale	• 0: 1435 (73.5%) • 1-3: 195 (10.0%) • 4-6: 172 (8.8%) • 7-10: 151 (7.7%)
Fall risk and clinical measures			
Fall risk (medication)	Risk level based on use of medications contributing to falls	No risk, polypharmacy, and high-risk medications	• No risk: 964 (49.4%) • Polypharmacy: 805 (41.2%) • High-risk medications: 184 (9.4%)
Fall risk (elimination)	Risk based on urinary or bowel elimination issues	Independent and at risk	• Independent: 1919 (98.3%) • At risk: 34 (1.7%)
Fall history	Any history of prior falls	Yes/no	• Yes: 80 (4.1%) • No: 1873 (95.9%)
Falls assessment score	Composite nursing assessment fall-risk score	0 - 9	• Median 1 (IQR 0 - 2) • 0: 595 (30.5%) • 1: 630 (32.3%) • 2: 444 (22.7%) • ≥3: 284 (14.5%)

XSL•FO
RenderX

Variable	Description	Range/coding	Distribution
Falls risk band	Categorized fall risk level derived from score	Low/moderate/high	• Low (62.8%) • Moderate (27.1%) • High (10.1%)
ECOG[a] performance status	Clinical performance status scale (0 - 4)	0="Fully active"; 4="Completely disabled"	• 0: 1505 (77.1%) • 1: 371 (19.0%) • 2: 51 (2.6%) • 3: 26 (1.3%)
Has barrier to learning	Presence of nurse-identified communication or learning barriers during care	0="No"; 1="Yes"	• No: 1662 (85.1%) • Yes: 291 (14.9%)
Influenza vaccine	Record of influenza vaccination	Yes/no	• Yes: 769 (39.4%) • No: 1184 (60.6%)
Fears or anxieties	Presence of fears or anxiety-related symptoms	Yes/no	• Yes: 259 (13.3%) • No: 1694 (86.7%)
Living will/advance care plan	Documented advance care planning	Yes/no	• Yes: 499 (25.6%) • No: 1454 (74.4%)
Demographic and socioeconomic variables			

XSL•FO
RenderX

Variable	Description	Range/coding	Distribution
Age	Age in years at the start of radiation therapy	Continuous variable	• Mean 64.5 (SD 12.1; range 21 - 95)
Race	Self-identified race	White, Black/African American, and other	• White: 1837 (94.1%) • Black: 88 (4.5%) • Other: 28 (1.4%)
Sex	Self-identified sex	Female and male	• Female: 1017 (52.1%) • Male: 936 (47.9%)
Marital status	Marital status at diagnosis	Married, divorced, single, widowed, and other	• Married: 1130 (57.9%) • Single: 312 (16.0%) • Divorced: 176 (9.0%) • Widowed: 159 (8.1%) • Other: 176 (9.0%)
ICD-10[b] category[c] (cancer site)	Primary cancer site category	14 *ICD*-10–based groups	• Metastasis: 395 (20.2%) • Breast: 315 (16.1%) • Prostate: 221 (11.3%) • Head/neck: 133 (6.8%)
Insurance category	Primary insurance type at treatment	Medicare, Medicaid, commercial, and so on	• Medicare: 647 (33.1%) • Commercial: 776 (39.7%) • Medicaid: 112 (5.7%) • Dual: 61 (3.1%) • Other: 168 (8.6%) • Missing category
Median household income	Area-level median income group	Low, moderate, and high	• Low: 596 (30.5%) • Moderate: 649 (33.2%) • High: 707 (36.2%)
Percent population ≥ bachelor's degree	Educational attainment level of census area	Low, moderate, and high	• Low: 589 (30.2%) • Moderate: 651 (33.3%) • High: 713 (36.5%)
Smoking status	Smoking history	Current, former, never, and unknown	• Current: 183 (9.4%) • Former: 355 (18.2%) • Never: 430 (22.0%) • Unknown: 985 (50.4%)
Alcohol use	Self-reported alcohol use	Current, past, denies, and unknown	• Denies: 628 (32.2%) • Current: 463 (23.7%) • Past: 110 (5.6%) • Unknown: 752 (38.5%)
Distance to radiation facility	Straight-line distance to treating radiation center (miles)	Continuous variable	• Median 13.5 (IQR 7.8 - 25.1; range: 0.9 - 255) miles

[a]ECOG: Eastern cooperative oncology group.

[b]*ICD*-10: *International Classification of Diseases*, 10th Revision.

[c]The 4 most common of 14 groups are shown.

The mean distress score was 3.1 (SD 2.9), with a range of 0 - 10 indicating generally moderate distress with considerable variability across the cohort. Emotional, physical, practical, and nutritional symptom variables represent the number of symptom categories patients endorsed within each domain. For emotional symptoms, 90.6% (n=1769) reported no emotional concerns, 3.4% (n=66) endorsed one emotional symptom, and fewer than 3% (n≤52) of patients in each remaining category endorsed two or more emotional symptoms. The majority of patients demonstrated good functional status, with 96% (1876/1953) scoring 0 - 1 on the Eastern Cooperative Oncology Group (ECOG) performance scale and exhibited low to moderate fall risk. The median travel distance from home address to the radiation treatment facility was 13.5 (IQR 7.8 - 25.1) miles.

A total of 177 patients were excluded from the analytic cohort due to missing distress score values. To assess whether their exclusion introduced systematic bias, we compared key demographic characteristics with those of the included cohort. The excluded patients were similar in age (median 64, IQR 56.8-73.7 y vs median 64.5, IQR 57.6-72.9 y), sex distribution

XSL•FO
RenderX

(female: 102/177, 57.7% vs 1017/1953, 52.1%), and race (White: 167/177, 94.3% vs 1838/1953, 94.1%). Marital status and insurance patterns also aligned closely with the main study population, indicating no meaningful demographic differences. Excluded patients also had a similar rate of missed treatment days (mean 1.09, SD 1.96; 67/177, 38.1% with any interruption) compared with included patients (mean 1.19, SD 2.18; 842/1953, 43.1% with any interruption), with no significant difference between groups by Mann-Whitney *U* test (*P*=.41), indicating no meaningful demographic or outcome-related differences between groups.

Negative Binomial Regression Modeling

Negative binomial regression identified multiple variables significantly associated with treatment interruptions (Table 2). Among psychosocial and clinical variables, patients' barriers to health-related communication or learning during care were associated with a 43% higher rate of missed sessions (IRR 1.43, 95% CI 1.07 - 1.92; *P*=.02). Each one-point increase in total distress score corresponded to a 5% higher rate of missed treatments (IRR 1.05, 95% CI 1.02 - 1.09; *P*=.01).

Table . Negative binomial regression model results.

Variable	IRR[a] (95% CI)	*P* value
ICD-10[b]: genitourinary	2.57 (1.33 - 4.97)	.006
Distress score (per unit)	1.05 (1.02 - 1.09)	.01
Marital status: married	0.67 (0.49 - 0.92)	.02
Learning barrier (yes vs no)	1.43 (1.07 - 1.92)	.02
Physical symptom count	0.82 (0.72 - 0.96)	.02
Nutrition symptom count	1.43 (0.99 - 2.11)	.051
ICD-10: prostate	1.52 (0.99 - 2.32)	.053
Marital status: widowed	0.65 (0.42 - 1.02)	.06
Race: White vs others	0.69 (0.46 - 1.02)	.06
Insurance: other insurance	0.37 (0.11 - 1.06)	.07
Distance to facility (per mile)	1.00 (0.99 - 1.00)	.08
Marital status: single	0.76 (0.53 - 1.09)	.13
Age (per y)	0.99 (0.98 - 1.00)	.14
ICD-10: head and neck	1.42 (0.87 - 2.31)	.16
Fall risk (elimination; independent vs at risk)	1.82 (0.78 - 4.18)	.17
Living will/advance care (yes vs no)	1.16 (0.93 - 1.43)	.18
ICD-10: skin	1.40 (0.81 - 2.43)	.22
ECOG[c] performance	1.10 (0.91 - 1.33)	.30
Primary pain intensity	1.02 (0.98 - 1.06)	.31
Influenza vaccine (yes vs no)	0.91 (0.75 - 1.11)	.35
ICD-10: other	1.31 (0.74 - 2.34)	.35
Falls assessment score	1.07 (0.91 - 1.27)	.37
Emotional symptom count	1.05 (0.92 - 1.27)	.40
Fall history (yes vs no)	0.75 (0.35 - 1.60)	.45
Practical symptom count	1.05 (0.78 - 1.48)	.66
Fears or anxieties (yes vs no)	0.96 (0.73 - 1.28)	.75
Fall risk (medication; polypharmacy vs no risk)	1.02 (0.73 - 1.42)	.86
Insurance: Medicaid	1.02 (0.31 - 2.92)	.97

[a]IRR: incident rate ratio.

[b]*ICD*-10: *International Classification of Diseases*, 10th Revision.

[c]ECOG: Eastern Cooperative Oncology Group.

Higher nutrition symptom counts were associated with greater interruption rates, approaching statistical significance (IRR

XSL•FO
RenderX

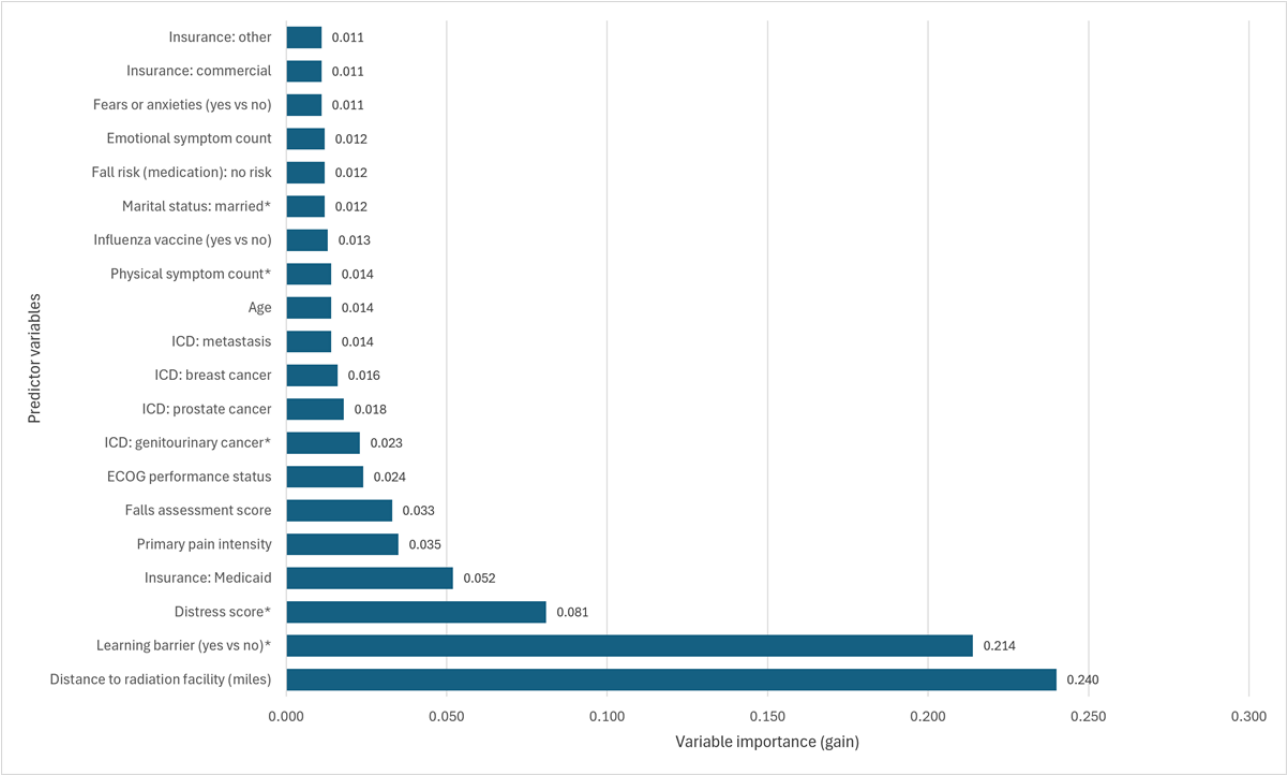
1.43, 95% CI 0.99 - 2.11; *P*=.051). In contrast, greater physical symptom burden was associated with lower interruption rates (IRR 0.82, 95% CI 0.72 - 0.96; *P*=.02).

Marital status and primary cancer site were also significant predictors. Married patients exhibited a lower rate of missed treatments (IRR 0.67, 95% CI: 0.49 - 0.92; *P*=.02) compared with unmarried patients. Patients diagnosed with genitourinary cancers had a significantly higher interruption rate (IRR 2.57, 95% CI 1.33 - 4.97; *P*=.006), with similar trends observed for head and neck, prostate, and skin cancers.

XGBoost Modeling

The XGBoost model achieved optimal performance using a maximum tree depth of 4, a learning rate of 0.05, and a subsampling rate of 0.9. Variable importance was assessed using the Gain metric, representing each predictor's relative contribution to reducing model error (Figure 1).

Figure 1. Variable importance in the extreme gradient boosting (XGBoost) model based on the gain metric. Higher values indicate greater relative contribution to prediction accuracy. Variables marked with an asterisk (*) were also statistically significant (*P*<.05) in the negative binomial regression model. ECOG: Eastern Cooperative Oncology Group; *ICD*: *International Classification of Diseases*.



Variables identified with the highest gain were patients' travel distance to the radiation facility, the presence of learning or communication barriers during treatment, and the total distress score, followed by Medicaid insurance and primary pain intensity. Moderate contributions were observed for fall risk assessment score, ECOG performance status, and genitourinary cancer site.

The Gain metric can favor high-cardinality continuous predictors, such as travel distance, and may produce less stable rankings among correlated features [15]. We therefore interpret the relative ranking of predictors in Figure 1 with this consideration in mind, particularly for travel distance, which may be ranked more highly than its true relative contribution due to this metric's known bias toward continuous variables with many unique values.

Predictive Model Comparison

Model performance metrics for the negative binomial and XGBoost models are summarized in Table 3. Both models provided stable and closely aligned predictions of missed radiation treatment days. Each achieved similar predictive accuracy, with XGBoost demonstrating incrementally lower root mean squared error (RMSE=2.19) and mean absolute error (MAE=1.36) vs the negative binomial model (RMSE=2.21; MAE=1.42). Poisson deviance values were also comparable between models, indicating consistent goodness-of-fit across approaches.

XSL•FO
RenderX

Table . Model performance comparison between negative binomial and extreme gradient boosting (XGBoost) models.

Model	RMSE[a], mean (SD; 95% CI)	MAE[b], mean (SD; 95% CI)	Poisson deviance, mean (SD; 95% CI)
XGBoost (Poisson)	2.19 (0.14; 1.95 - 2.49)	1.36 (0.05; 1.28 - 1.47)	973 (62; 869 - 1109)
Negative binomial	2.21 (0.13, 1.99 - 2.50)	1.42 (0.05; 1.33 - 1.53)	979 (66; 858 - 1125)

[a]RMSE: root mean squared error.

[b]MAE: mean absolute error.

To place these values in clinical context, we compared both models against a naive baseline predicting the sample mean for all patients, evaluated using the same resampling procedure. The naive baseline achieved a mean RMSE of 2.223 (SD 0.133; 95% CI 1.99 - 2.49), only modestly higher than the negative binomial and XGBoost models. This indicates that, while both models offer a measurable improvement in absolute prediction accuracy over a covariate-free estimate, the improvement is limited given the low-count, zero-inflated nature of the outcome.

Discussion

In this study of patients receiving radiation therapy, we examined how psychological and distress-related symptoms, along with other key clinical and social risk factors, were associated with RTI. Among all studied variables, total distress score emerged as a key determinant, with higher baseline distress levels independently linked to an increased likelihood of missed treatment days across both the negative binomial and XGBoost models. This finding underscores the critical influence of psychological well-being on treatment adherence. Other variables with demonstrated significance in both models, learning or communication barriers during treatment and genitourinary cancer site, highlight complex, multilayered mechanisms leading to RTIs. Strong concordance between traditional regression and machine learning results supports the robustness of identified associations and reinforces the need to incorporate patient-reported data to inform personalized psychosocial and supportive interventions to mitigate preventable interruptions across a multiweek course of daily radiation delivery.

Our study expands upon prior literature demonstrating the influence of emotional distress on radiation therapy adherence. Clover et al [16] found that radiation therapists rated 11% of sessions as disrupted by anxiety at the pretreatment simulation visit, rising to 24% at the first treatment session, with disruptions primarily linked to procedural fears (eg, mask-related anxiety) rather than overall distress. Sreeraman et al [17] found that, among 280 patients with head and neck cancer, poor Karnofsky Performance Status (≤70) and pre-existing psychiatric illness were each significantly associated with a greater likelihood of missing more than 5 treatment days, and Chen et al [18] demonstrated that pretreatment depression independently predicted both prolonged radiation breaks and reduced survival.

Our findings build on this mixed evidence by demonstrating a consistent and independent association between patient-reported distress and RTIs across 2 distinct analytic frameworks. Unlike prior work that relied on diagnostic psychiatric categories or clinician-reported psychological factors [17,19,20], the use of the National Comprehensive Cancer Network (NCCN) Distress Thermometer allowed us to capture real-time, patient-centered indicators of emotional strain. This may explain the stronger and more stable associations observed. Moreover, recent analyses of NCCN Distress Thermometer scores indicate that higher distress correlates with greater acute treatment-related toxicity during radiotherapy, lending additional biological plausibility to distress as a mediator of treatment course deviations [21].

Beyond psychological distress, several additional predictors emerged as important determinants of radiation treatment adherence. Patients with learning or communication barriers consistently showed higher rates of missed sessions, underscoring the role of health literacy and communication effectiveness in sustaining treatment continuity. This aligns with evidence that health system and informational needs are the second most frequently reported unmet supportive care needs among people affected by cancer, behind only psychological and emotional needs, with patients commonly reporting difficulty understanding their diagnosis, treatment, and self-management, as well as challenges navigating and coordinating care [22]. Cancer site was also associated with treatment interruption, with patients with genitourinary and prostate cancers experiencing higher rates of missed treatments. Although prostate cancer is anatomically classified within the genitourinary system, we analyzed it separately because of its high prevalence in our cohort and its distinct treatment patterns. Prior studies have reported higher rates of treatment interruption among patients with head and neck cancer [16,23]; however, head and neck cancer was not among the highest-ranked predictors of treatment interruption in our models. While indicators of functional status such as ECOG performance and fall-risk assessments were not statistically significant, they ranked moderately in the XGBoost model, suggesting that physical limitations may still contribute indirectly to treatment nonadherence. Importantly, evidence from a multiple-baseline intervention study demonstrates that structured pretreatment education delivered by radiation therapists can reduce psychological distress and improve treatment preparedness, supporting the potential of communication-focused interventions to mitigate avoidable treatment disruptions [24]. Taken together, these findings emphasize the value of targeted educational and supportive strategies, particularly for patients with communication challenges or undergoing complex treatment regimens.

Notably, higher physical symptom burden was associated with fewer, rather than more, treatment interruptions in our cohort. This finding contrasts with prior evidence that greater symptom burden is generally associated with worse treatment adherence;

XSL•FO
RenderX

for example, each additional moderate-severity symptom during adjuvant endocrine therapy in patients with breast cancer has been associated with 4 to 9 additional missed treatment days [25]. One possible explanation for the divergent direction observed in our study is that patients experiencing greater physical symptoms during radiotherapy may receive closer clinical attention, as prior work has shown that higher symptom severity scores are associated with a greater likelihood of being offered a clinical intervention [26]. Whether this increased clinical contact translates into improved treatment continuity remains untested and should be examined in future prospective work. Alternatively, this association may reflect a form of "healthy adherer" behavior, in which patients more engaged with their overall care are both more likely to accurately report symptoms and more likely to complete treatment as scheduled. Further prospective research incorporating longitudinal symptom tracking is needed to clarify this relationship.

Interestingly, some variables demonstrated high predictive importance in the XGBoost model despite lacking statistical significance in the negative binomial regression. Travel distance to treatment facilities was among the top contributors in the machine learning model, suggesting that geographic accessibility meaningfully influences treatment adherence, even when this effect is not captured well by traditional linear modeling. This could be due to complex, nonlinear relationships, such as threshold effects where adherence declines sharply beyond certain travel distances, or interactions with socioeconomic and logistical factors such as transportation availability or employment constraints. Similarly, insurance coverage and marital status, though not statistically significant, contributed to model accuracy, implying subtle cumulative effects of socioeconomic stability and social support on treatment access. These results highlight the complementary value of machine learning approaches in uncovering interaction and nonlinear effects which conventional regression models may underestimate.

In this study, large language models were incorporated into data preprocessing to improve the usability of unstructured clinical text and enable consistent classification of fields that were otherwise incomplete, inconsistently formatted, or missing structured responses. Model performance varied by variable type. For structured *International Classification of Diseases*, 10th Revision (*ICD*-10) diagnosis codes, GPT-4o achieved an accuracy of 90.8% (95% CI 87.7 - 93.9) while for less structured entries such as insurance categorization, accuracy was 81.6% (95% CI 76.7 - 86.5). Performance declined further for the 168 distress screening cases in which a nurse had recorded free-text descriptions rather than completing the structured checklist fields; owing to the greater linguistic complexity and variability of this free text, categorization accuracy for these cases was only 48.1% (95% CI 43.0 - 53.2). All large language model–generated categorizations were subsequently reviewed and manually refined by our oncology team, and no record was retained in an incorrect category.

Distress was assessed at baseline only, before treatment began, which creates a temporal mismatch with our outcome, as missed treatment days can accumulate over a 4-week to 7-week course of daily fractionated radiation. Baseline distress represents the most actionable time point for early psychosocial intervention and is the only assessment consistently collected within our institution's standard workflow; however, it may not reflect a patient's psychological state at the time interruptions occur. Among the 1391 (71.2%) patients with more than one distress screening during treatment (median 2 assessments per patient), 43.7% (n=608) showed a decreased score and 30.3% (n=421) showed an increased score between their first and last measurement, confirming that distress is not static for many patients. Because the number of distress forms did not consistently align with the timing of treatment sessions or interruptions, we were unable to examine whether the distress-RTI association differs between early and later interruptions. Future work incorporating repeated, treatment-aligned distress assessments is needed to clarify this temporal relationship.

Some limitations should be considered when interpreting our findings. First, our retrospective analysis was limited to a single-institutional experience, constraining the generalizability of results to other health systems and patient populations. Also, reliance on electronic health record data introduces the possibility of incomplete or inconsistent documentation, particularly for psychosocial and symptom-related variables. Moreover, area-level socioeconomic indicators derived from census data may not accurately represent individual circumstances. Our outcome measure also does not distinguish patient-initiated absences from provider- or operationally-driven interruptions, as reliably coded reason data were not consistently available, although such interruptions are uncommon in our practice. Finally, while the XGBoost model improved prediction and identified nonlinear effects, it does not infer causality, and variable importance should be interpreted as associative rather than deterministic.

This study integrated clinical, socioeconomic, and patient-reported data to examine how psychological and distress-related symptoms, along with other key clinical and social determinants of health, affect unplanned radiation treatment interruption rates. Both the negative binomial and XGBoost models consistently identified associations of patient-reported distress, learning or communication barriers, and specific cancer types with interrupted treatment. These findings highlight the potential influence of psychological and functional factors on cancer treatment access and a critical need to collect patient-reported hardships to inform social support interventions.

By combining traditional regression and machine learning approaches, this study provides an early but potentially comprehensive understanding of the multifactorial determinants of radiation treatment access. Future work will need to validate reproducible predictive factors across larger, multi-institutional cohorts. Ultimately, such data may inform personalized intervention strategies, such as enhanced psychosocial screening, patient education, and community-centered assistance to support high-quality patient access and treatment efficacy for all patients with cancer requiring radiotherapy.

XSL•FO
RenderX

Acknowledgments

The authors thank the University of Tennessee IT team for their support with data access and infrastructure. The authors used ChatGPT-4o (OpenAI) solely to check grammar and improve language clarity during manuscript preparation; the tool was not used to generate scientific content, analyze data, or draw conclusions, and all edits were reviewed and verified by the authors for accuracy. This use is distinct from GPT-4o's role as an analytical preprocessing tool within the study's data pipeline, described in detail in the "Automated Variable Categorization Using Large Language Models" section.

Funding

This research was supported by a grant from the Tennessee Department of Health.

Data Availability

The data that support the findings of this study are available upon reasonable request from the corresponding author. The dataset is not publicly available due to institutional privacy policies.

Authors' Contributions

SH conducted data integration and statistical and machine learning analyses and drafted the manuscript. FAK and RR assisted with data processing and analyses. PKR and SK contributed to data curation and manuscript review. CLB and DLS, as radiation oncologists, provided clinical expertise and validated the clinical relevance of the study design and findings. RLD contributed to study design and methodological oversight. AS-N and DLS conceptualized and supervised the study, reviewed, and edited the manuscript, interpreted the results, and acquired funding. The two senior authors (AS-N and DLS) contributed equally to the work. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Baskar R, Lee KA, Yeo R, Yeoh KW. Cancer and radiation therapy: current advances and future directions. Int J Med Sci 2012;9(3):193-199. [doi: 10.7150/ijms.3635] [Medline: 22408567]
2. Chow R, Hasan S, Choi JI, et al. Effect of treatment interruptions on overall survival in patients with triple-negative breast cancer. J Natl Cancer Inst 2023 Sep 7;115(9):1029-1035. [doi: 10.1093/jnci/djad127] [Medline: 37399094]
3. Hecht M, Meier F, Zimmer L, et al. Clinical outcome of concomitant vs interrupted BRAF inhibitor therapy during radiotherapy in melanoma patients. Br J Cancer 2018 Mar 20;118(6):785-792. [doi: 10.1038/bjc.2017.489] [Medline: 29438368]
4. Zink JA, Rashid R, Chinthala L, et al. X-CART: a multimodal real-world data pipeline for explainable AI-driven patient navigation in radiation oncology. ESMO Real World Data Digit Oncol 2026 Sep;13:100771. [doi: 10.1016/j.esmorw.2026.100771] [Medline: 42751204]
5. Hubler A, Wakefield DV, Makepeace L, et al. Independent predictors for hospitalization-associated radiation therapy interruptions. Adv Radiat Oncol 2022;7(6):101041. [doi: 10.1016/j.adro.2022.101041] [Medline: 36158745]
6. Yang XL, Zhou GQ, Lin L, et al. Prognostic value of radiation interruption in different periods for nasopharyngeal carcinoma patients in the intensity-modulated radiation therapy era. Cancer Med 2021 Jan;10(1):143-155. [doi: 10.1002/cam4.3580] [Medline: 33107201]
7. Akinlade BI, Folasire AM, Elumelu-Kupoluyi TN, Adenipekun AA, Iyobosa UB, Campbell OB. Radiation therapy interruption in a poor resource setting: causes and management. Afr J Med Med Sci 2014 Dec;43(4):333-337. [Medline: 26234121]
8. Sawaf T, Virgen CG, Renslo B, et al. Association of social-ecological factors with delay in time to initiation of postoperative radiation therapy: a prospective cohort study. JAMA Otolaryngol Head Neck Surg 2023 Jun 1;149(6):477-484. [doi: 10.1001/jamaoto.2023.0308] [Medline: 37079327]
9. Kumsa FA, Brett CL, Hashtarkhani S, et al. Leveraging centralized health system data management and large language model-based data preprocessing to identify predictors for radiation therapy interruption. JCO Clin Cancer Inform 2025 Oct;9(9):e2500218. [doi: 10.1200/CCI-25-00218] [Medline: 41150999]
10. Wakefield DV, Carnell M, Dove APH, et al. Location as destiny: Identifying geospatial disparities in radiation treatment interruption by neighborhood, race, and insurance. Int J Radiat Oncol Biol Phys 2020 Jul 15;107(4):815-826. [doi: 10.1016/j.ijrobp.2020.03.016] [Medline: 32234552]
11. Syrjala KL, Sutton SK, Jim HSL, et al. Cancer and treatment distress psychometric evaluation over time: a BMT CTN 0902 secondary analysis. Cancer 2017 Apr 15;123(8):1416-1423. [doi: 10.1002/cncr.30454] [Medline: 27893933]

XSL•FO
RenderX

12. Sharma A, Loades ME, Baker L, Jordan A, James V. Parental experiences of adolescent cancer-related distress: a qualitative study. Eur J Cancer Care (Engl) 2021 Jul;30(4):e13417. [doi: 10.1111/ecc.13417] [Medline: 33512062]
13. American Community Survey 5-year data (2009-2024). US Census Bureau. 2024. URL: https://www.census.gov/data/developers/data-sets/acs-5year.html [accessed 2026-09-30]
14. Hashtarkhani S, Rashid R, Brett CL, et al. Cancer diagnosis categorization in electronic health records using large language models and BioBERT: model performance evaluation study. JMIR Cancer 2025 Oct 2;11:e72005. [doi: 10.2196/72005] [Medline: 41037674]
15. Nembrini S, König IR, Wright MN. The revival of the Gini importance? Bioinformatics 2018 Nov 1;34(21):3711-3718. [doi: 10.1093/bioinformatics/bty373] [Medline: 29757357]
16. Clover K, Oultram S, Adams C, Cross L, Findlay N, Ponman L. Disruption to radiation therapy sessions due to anxiety among patients receiving radiation therapy to the head and neck area can be predicted using patient self-report measures. Psychooncology 2011 Dec;20(12):1334-1341. [doi: 10.1002/pon.1854] [Medline: 20878722]
17. Sreeraman R, Vijayakumar S, Chen AM. Correlation of radiation treatment interruptions with psychiatric disease and performance status in head and neck cancer patients. Support Care Cancer 2013 Dec;21(12):3301-3306. [doi: 10.1007/s00520-013-1907-3] [Medline: 23892904]
18. Chen AM, Hsu S, Felix C, Garst J, Yoshizaki T. Effect of psychosocial distress on outcome for head and neck cancer patients undergoing radiation. Laryngoscope 2018 Mar;128(3):641-645. [doi: 10.1002/lary.26751] [Medline: 28714543]
19. Holmes EG, Holmes JA, Park EM. Psychiatric care of the radiation oncology patient. Psychosomatics 2017;58(5):457-465. [doi: 10.1016/j.psym.2017.04.004] [Medline: 28599843]
20. Jenkins C, Carmody TJ, Rush AJ. Depression in radiation oncology patients: a preliminary evaluation. J Affect Disord 1998 Jul;50(1):17-21. [doi: 10.1016/s0165-0327(98)00039-1] [Medline: 9716274]
21. Zhao M, Benson R, Friesner I, et al. Association between NCCN distress thermometer screening data and natural language processing-identified toxicities during radiotherapy. Int J Radiat Oncol Biol Phys 2024 Oct;120(2):S173-S174. [doi: 10.1016/j.ijrobp.2024.07.2210]
22. Riedl D, Gastl R, Gamper E, et al. Cancer patients' wish for psychological support during outpatient radiation therapy: findings from a psychooncological monitoring program in clinical routine. Strahlenther Onkol 2018 Jul;194(7):655-663. [doi: 10.1007/s00066-018-1288-0] [Medline: 29532098]
23. Russo G, Haddad R, Posner M, Machtay M. Radiation treatment breaks and ulcerative mucositis in head and neck cancer. Oncologist 2008 Aug;13(8):886-898. [doi: 10.1634/theoncologist.2008-0024] [Medline: 18701763]
24. Halkett G, O'Connor M, Jefford M, et al. RT Prepare: a radiation therapist-delivered intervention reduces psychological distress in women with breast cancer referred for radiotherapy. Br J Cancer 2018 Jun;118(12):1549-1558. [doi: 10.1038/s41416-018-0112-z] [Medline: 29855611]
25. Hu X, Walker MS, Stepanski E, et al. Racial differences in patient-reported symptoms and adherence to adjuvant endocrine therapy among women with early-stage, hormone receptor-positive breast cancer. JAMA Netw Open 2022 Aug 1;5(8):e2225485. [doi: 10.1001/jamanetworkopen.2022.25485] [Medline: 35947386]
26. Rau A, Yannitsos D, Grendarova P, Qi S, Watson L, Barbera L. High symptom burden in patients receiving radiotherapy and factors associated with being offered an intervention. Curr Oncol 2024 Feb 27;31(3):1253-1265. [doi: 10.3390/curroncol31030094] [Medline: 38534927]

Abbreviations

ECOG: Eastern Cooperative Oncology Group
GVIF: generalized variance inflation factor
***ICD*:** *International Classification of Diseases*
***ICD*-10:** *International Classification of Diseases*, 10th Revision
IRR: incidence rate ratio
MAE: mean absolute error
NCCN: National Comprehensive Cancer Network
RMSE: root mean square error
RTI: radiation treatment interruption
VIF: variance inflation factor
XGBoost: extreme gradient boosting

XSL•FO
RenderX

Edited by M Balcarras; submitted 05.Feb.2026; peer-reviewed by E Miranda, M Jani, N Rebecca; revised version received 17.Sep.2026; accepted 18.Sep.2026; published 19.Oct.2026.

Please cite as:
Hashtarkhani S, Brett CL, Kumsa FA, Rashid R, Kheirkhah Rahimabad P, Kheirinejad S, Davis RL, Schwartz DL, Shaban-Nejad A Psychosocial Distress and Radiation Treatment Interruptions in Patients With Cancer Using Machine Learning and Regression: Retrospective Cohort Study
JMIR Cancer 2026;12:e92899
URL: https://cancer.jmir.org/2026/1/e92899
doi:10.2196/92899

© Soheil Hashtarkhani, Christopher L Brett, Fekede Asefa Kumsa, Rezaur Rashid, Parnian Kheirkhah Rahimabad, Saba Kheirinejad, Robert L Davis, David L Schwartz, Arash Shaban-Nejad. Originally published in JMIR Cancer (https://cancer.jmir.org), 19.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Machine Learning–Based Survival Prediction Models for Young Patients With Gastric Cancer: Model Development and Validation Study

Ha Ye Jin Kang[1], PhD; Wooyeong Jang[2], MS; Minsam Ko[3], PhD; Kwang Sun Ryu[4], PhD

[1]Department of Computer Science, Semyung University, Jecheon-si, Chungcheongbuk-do, Republic of Korea

[2]Department of Biomedical Engineering, School of Medicine, Chungbuk National University, Seowon-gu, Cheongju-si, Chungcheongbuk-do, Republic of Korea

[3]Department of Human-Computer Interaction, Hanyang University, Sangnok-gu, Ansan-si, Gyeonggi-do, Republic of Korea

[4]National Cancer Data Center, National Cancer Center, 323 Ilsan-ro, Ilsandong-gu, Goyang-si, Gyeonggi-do, Republic of Korea

Corresponding Author:
Kwang Sun Ryu, PhD
National Cancer Data Center, National Cancer Center, 323 Ilsan-ro, Ilsandong-gu, Goyang-si, Gyeonggi-do, Republic of Korea

Abstract

Background: Despite a global decline in the incidence of gastric cancer (GC), the number of cases diagnosed among younger individuals continues to increase. Several studies have been conducted to develop predictive models of mortality in patients with GC.

Objective: We developed 3- and 5-year survival prediction models for young patients with GC based on machine learning–based survival modeling approaches.

Methods: Data from 813 young patients (≤50 years) diagnosed with GC between 2013 and 2015 were retrieved from the Gastric Cancer Public Staging Database. Among these 813 patients, data from 569 (70%) and 244 (30%) were allocated to the model training and testing datasets, respectively. Random survival forest, gradient boosting survival analysis, extra survival tree, and the Cox proportional hazards model were applied to predict survival outcomes at the 3- and 5-year time horizons. Model performance was assessed and quantified using the concordance index (C-index) metric. For the machine learning prediction models, the mean C-index values and corresponding 95% CIs were estimated across 100 repeated training iterations.

Results: For the random survival forest model, the C-index for predicting 3-year mortality was 95.89% (95% CI 95.80% - 95.97%), whereas the C-index for predicting 5-year mortality was 91.82% (95% CI 91.68% - 91.96%). In the gradient boosting survival analysis model, the C-index for predicting 3-year mortality was 95.32% (95% CI 95.31% - 95.33%), and the C-index for predicting 5-year mortality was 89.98% (95% CI 89.95% - 90.01%). For the extra survival tree model, the C-index for predicting 3-year mortality was 95.53% (95% CI 95.46% - 95.60%), whereas the C-index for predicting 5-year mortality was 94.60% (95% CI 94.50% - 94.70%). In addition, the Cox proportional hazards model showed a C-index of 94.15% for predicting 3-year mortality and 82.26% for predicting 5-year mortality. Tumor stage and tumor size were the primary predictive variables used to train the models for mortality prediction at different time points. Other variables exhibited varying levels of predictive contribution across different time points.

Conclusions: These findings may facilitate the identification of high-risk young patients with GC who may benefit from more aggressive treatment strategies by enabling the prediction of mortality risk at different time points.

(JMIR Cancer 2026;12:e86418) doi:10.2196/86418

KEYWORDS

survival prediction model; young patients with gastric cancer; machine learning; gastric cancer; mortality risk; predictive modeling

Introduction

Despite a consistent global decline in the incidence of gastric cancer (GC) resulting from improved understanding of the key etiological factors influencing its development, more than 1 million new cases are still diagnosed annually worldwide, and the disease remains associated with a high mortality rate [1]. Recent studies have shown that the prevalence of GC among younger individuals varies substantially across geographic regions and is increasing worldwide. This trend is particularly evident in Asian countries such as South Korea, Japan, and China [2-4] where GC occurs at a comparatively higher rate among younger populations [2]. Although the precise mechanisms underlying GC development in younger individuals remain unclear, both disease progression and prognosis are

XSL•FO
RenderX

influenced by a complex interplay of factors, including pathological characteristics, molecular alterations, genomic features, dietary patterns, and lifestyle-related exposures. Consequently, several studies have been conducted to evaluate prognostic outcomes in younger patients with GC [5,6]. Previous studies predicting survival outcomes in young patients with GC have primarily relied on traditional statistical nomogram models. These models represent the relationships between clinical covariates and patient outcomes in an interpretable format and have demonstrated relatively stable predictive performance [7].

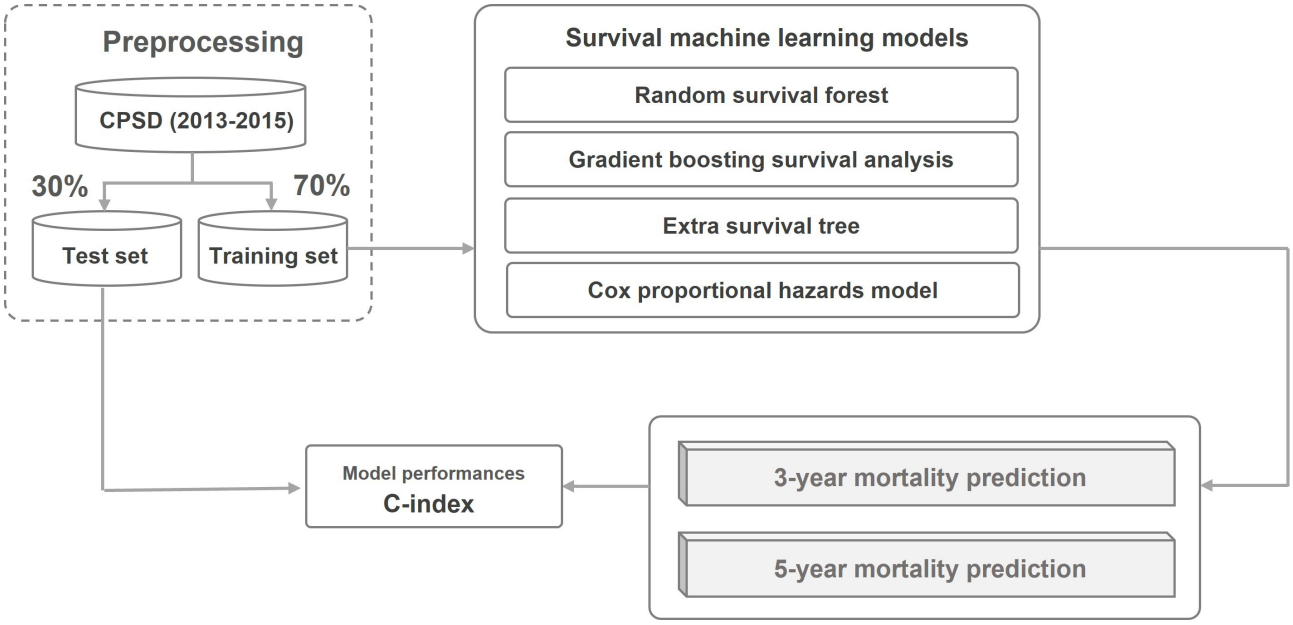
Recently, machine learning (ML)–based approaches have been explored as complementary tools to traditional prediction models [8,9], demonstrating particular strengths in capturing complex nonlinear relationships among predictor variables and analyzing large-scale, high-dimensional clinical datasets [10,11]. Several studies have demonstrated that ML-based survival models can provide meaningful prognostic insights for predicting time-to-event outcomes using structured clinical data collected in real-world health care settings [12-16]. However, in young patients with GC, baseline prognostic factors influencing survival change over time [17]. In addition, existing mortality prediction models typically provide predictions for a single time point, which may limit their ability to capture time-varying changes in prognostic factors that may occur in younger patients with GC [15]. Therefore, to account for the complex nature of GC prognosis, this study developed 3- and 5-year mortality prediction ML-based models for young patients with GC and constructed 3- and 5-year mortality prediction models to capture potential time-varying risk factors. This ML-based survival prediction study aimed to identify how key prognostic factors associated with mortality vary across different prediction horizons.

Methods

Study Design

We developed a predictive model to estimate all-cause mortality among younger patients with GC using deidentified data obtained from the Cancer Public Staging Database (CPSD) provided by the National Cancer Data Center (NCDC; Figure 1). To achieve this objective, we first extracted data for younger patients (≤50 years) diagnosed with GC between 2013 and 2015 from the CPSD and subsequently performed data preprocessing on the dataset. The data were randomly divided into training and testing subsets comprising 70% and 30% of the data, respectively. Using the training dataset, we developed 3- and 5-year mortality prediction models using random survival forest (RSF), gradient boosting survival analysis (GBSA), extra survival tree (EST), and Cox proportional hazards (PH).

Figure 1. Research framework. CPSD: Cancer Public Staging Database.



Ethical Considerations

This study was approved by the institutional review board of the National Cancer Center of South Korea (NCC2023-0260). The requirement for informed consent was waived because HYJK, WJ, and KSR accessed only anonymized data for the purposes of analysis. The pseudonymized dataset was analyzed within a secure research environment provided by the NCDC, ensuring that only aggregated analytical results were exported.

Data Source

This study used the CPSD, which is provided by the NCDC. This database was constructed by linking data from the NCDC's Cancer Public Library Database with collaborative staging data for cancer cases established in the Korea Central Cancer Registry (KCCR) [18]. The collaborative staging system was developed using probability-based sampling methods and mandatory retrospective medical record surveys of cancer cases registered in the KCCR. These surveys provide information on cancer staging, treatment modalities, and patient prognoses [19]. The Cancer Public Library Database was developed by integrating linked data from 4 major population-based public databases: the Korea National Cancer Incidence Database maintained by the KCCR, cause of death records from Statistics Korea, the National Health Information Database of the National

XSL•FO
RenderX

Health Insurance Service, and the National Health Insurance Research Database maintained by the Health Insurance Review and Assessment Service [20]. The use of the CPSD dataset complied with the data access and use policies established by the NCDC, and all data were handled in accordance with relevant guidelines and regulations to ensure confidentiality and data security. As the dataset contained no personally identifiable information, this study was exempt from institutional review board review.

Definitions

Comorbid conditions, including myocardial infarction, stroke, heart failure, diabetes mellitus, hypertension, dyslipidemia, chronic obstructive pulmonary disease, peripheral vascular disease, liver disease, atrial fibrillation, and chronic kidney disease, were identified and defined according to the diagnostic codes from the *International Classification of Diseases, 10th Revision*. Additionally, primary study end points were defined to represent short- and medium-term mortality, specifically all-cause mortality occurring within 3 and 5 years after cancer diagnosis, respectively. Information on all-cause mortality was obtained from official cause of death records provided by South Korea's national statistical authority. Deaths occurring within each follow-up period were treated as cumulative mortality events; therefore, events occurring within 3 years were also included in the cumulative event count for the 5-year follow-up period.

Experimental Dataset

The dataset was randomly divided into a training set (n=569) and a test set (n=244) using a fixed random seed (42) to ensure the reproducibility of the dataset partitioning. The primary outcome was all-cause mortality measured at 3 and 5 years after cancer diagnosis, and separate survival prediction models were developed for each prediction horizon. To minimize potential bias caused by outcome imbalance, the dataset was randomly partitioned using stratified sampling, ensuring comparable proportions of mortality events between the training and testing datasets. Model performance was evaluated on the corresponding test dataset using the Harrell C-index, a widely used discrimination metric in survival analysis. For the ML-based survival models (RSF, GBSA, and EST), the training and evaluation procedures were repeated across 100 iterations, with the random seed varied in each iteration while maintaining a fixed data split to quantify stochastic variability and assess the robustness of model performance. At each iteration, the models were refitted using the same training dataset and subsequently evaluated on the identical test dataset. The resulting training and testing datasets were fixed after partitioning and were consistently used in all subsequent model development and evaluation experiments.

Models

The RSF is an ensemble-based ML method that models complex nonlinear relationships between predictor variables and estimates variable importance to reduce generalization error. The RSF model can also estimate the cumulative hazard function for individual observations even when the PH assumption is not satisfied [21,22]. A gradient boosting machine framework [23] was adapted for survival analysis, enabling the model to effectively handle light-censored survival data. Gradient boosting machine is an ensemble learning technique in which successive decision trees are sequentially constructed, with each tree iteratively correcting the prediction errors of the preceding model. GBSA builds on this boosting framework for survival modeling, allowing the algorithm to handle censored observations and capture complex interactions between predictor variables and survival time [24,25]. The EST model, an extension of the randomized trees algorithm introduced by Geurts et al [26] in 2006, introduces additional randomness by selecting split points randomly during tree construction. The EST algorithm can censor survival data and does not rely on the PH assumption [27]. The Cox PH model is a widely used semiparametric survival analysis method that examines the relationship between survival time and multiple predictor variables. This model does not require explicit specification of the baseline hazard function [28,29].

Computational Environment Settings

Survival analysis is a statistical methodology used to analyze time-to-event outcomes such as death or disease progression [30], and ML approaches are increasingly being adopted in health care research to address the analytical complexities associated with survival modeling [31]. Nonparametric tree-based ensemble methods are generally considered well suited for analyzing censored time-to-event outcomes and performing dynamic survival predictions [32]. The predictor variables were further ranked and stratified according to their relative feature importance within the survival prediction models. In this study, we developed a survival prediction model based on GBSA, RSF, EST, and the Cox PH. During model development, feature importance scores were computed based on the mean decrease in impurity using the scikit-learn library (version 1.0.2; Google Summer of Code project) [33]. The models were implemented using the scikit-survival library (version 0.17.2) [34] within a Python (version 3.7.5; Python Software Foundation) computational environment, with TensorFlow (version 1.15.5; Google Brain Team) used as the underlying ML framework.

All survival models were trained using predefined and fixed hyperparameter configurations. For the RSF, GBSA, and EST models, the number of estimators was fixed at 200, whereas all remaining hyperparameters were retained at their default library settings. For the Cox PH model, elastic net regularization was applied with an L1 ratio of 1×10^{-6} and predefined alpha values to enhance numerical stability during model optimization. Continuous variables were not normalized for tree-based survival models because tree-based algorithms are inherently scale invariant and do not require feature normalization. Missing values were handled through data imputation using the median value for continuous variables and the most frequent category for categorical variables followed by one-hot encoding of categorical features. Importantly, all preprocessing procedures—including imputation and feature encoding—were fit exclusively on the training dataset and subsequently applied to the test dataset to prevent information leakage during model evaluation. For the Cox PH model, the predictor variables were

XSL•FO
RenderX

standardized prior to model training to improve numerical stability and optimization convergence.

Model Performance Evaluation

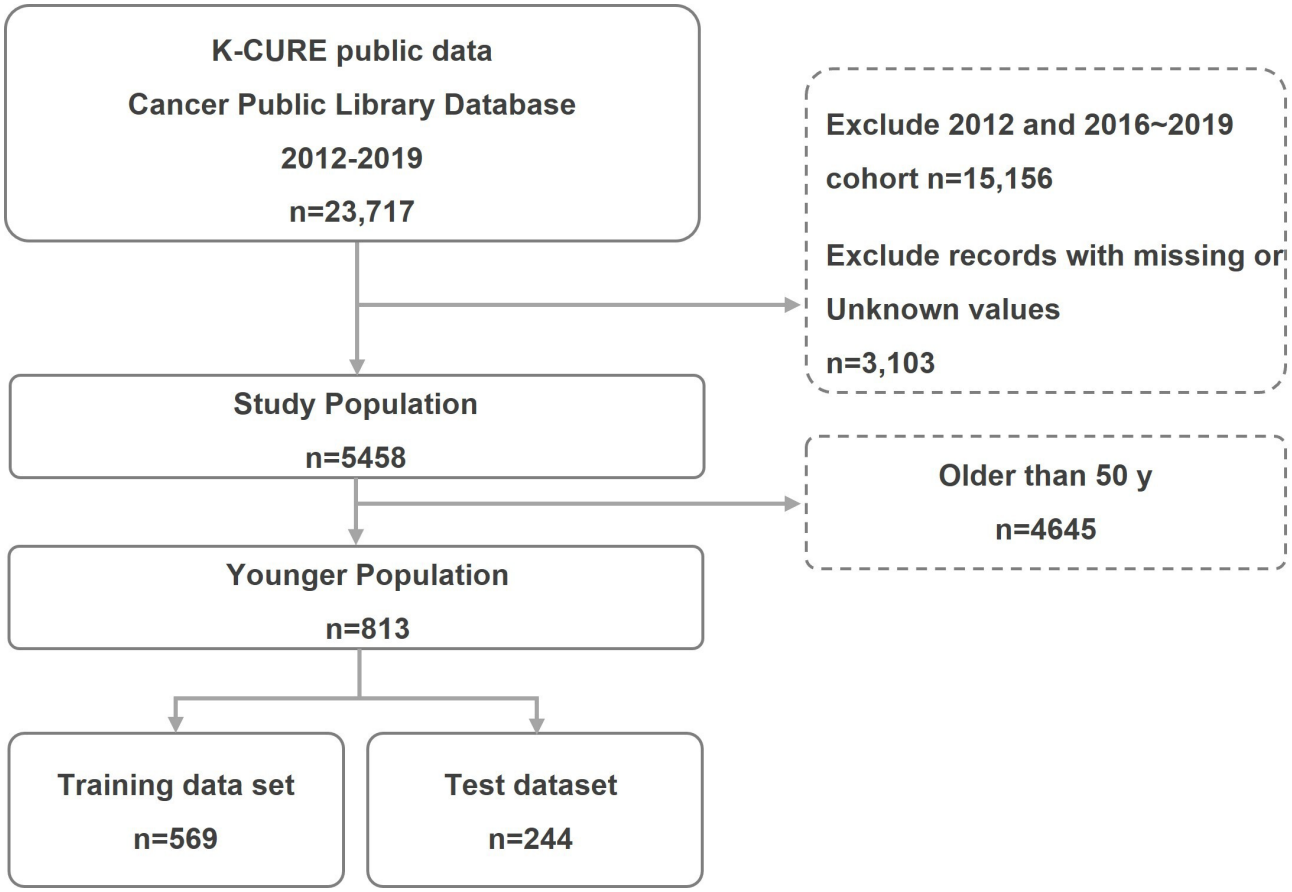
The C-index [35] is the most widely used evaluation metric for assessing the performance of survival prediction models [36]. This C-index is computed by comparing the predictive risk scores and observed survival times for pairs of randomly selected patients. To determine the number of concordant and discordant patient pairs, this pairwise comparison process was repeated across all possible patient pairs within the study cohort. The predictive model assigns a numerical risk score to each individual, representing the predicted risk of mortality, whereas the C-index evaluates the model's ability to correctly rank the relative risk between pairs of individuals [37,38]. Therefore, the C-index was used to evaluate the predictive performance and discrimination ability of the survival prediction models.

Results

Study Population

The CPSD contains data for 23,717 individuals diagnosed with GC at the primary anatomical sites C160 to C166, C168, and C169 between 2012 and 2019. From this dataset, we excluded patients diagnosed with GC between 2016 and 2019 for 5-year mortality prediction analysis, as well as the 2012 cohort, for which prediagnostic screening information was unavailable. The dataset subsequently underwent data cleaning procedures, during which 3103 records containing missing or unknown values were removed, including tumor size (T-size; n=1140, 36.7%), height (n=1849, 59.6%), urine protein (n=44, 1.4%), gamma-glutamyl transpeptidase (n=4, 0.1%), low-density lipoprotein cholesterol (n=17, 0.5%), and estimated glomerular filtration rate (eGFR; n=49, 1.6%). Patients aged 50 years or younger at the time of diagnosis were classified as the "younger" patient group, as illustrated in Figure 2.

Figure 2. Generation process of the experimental datasets. K-CURE: Korea Clinical Data Utilization Network for Research Excellence.



Baseline Characteristics

The study population showed differences between younger and older patients. Among younger patients, there was a higher proportion of women, whereas older patients exhibited a higher burden of comorbidities, including cardiovascular and metabolic diseases. Differences were also observed in tumor-related characteristics, including cancer stage according to the seventh edition of the American Joint Committee on Cancer (AJCC) cancer staging manual, with younger patients tending to present with more advanced stages. In addition, older patients showed higher all-cause mortality at both 3 and 5 years. Detailed baseline characteristics are provided in Multimedia Appendix 1.

Model Performance

The RSF, GBSA, EST, and Cox PH models were trained and evaluated to compare their predictive performance for survival outcomes at 2 time horizons (3 and 5 years). Table 1 presents the mean C-index values with corresponding 95% CIs, reflecting the predictive accuracy of the models in estimating mortality risk among young patients with GC at different prediction time

XSL•FO
RenderX

points. For the ML prediction models, we reported the mean C-index and the corresponding empirical 95% CIs across 100 repeated training iterations, whereas the Cox PH model was evaluated using a single model run; therefore, CIs for the C-index were not estimated or reported. The C-index values obtained from each of the 100 training iterations for the prediction models are provided in Multimedia Appendix 2. For 3-year survival prediction, the RSF, GBSA, EST, and Cox PH models achieved C-index values of 95.89% (95% CI 95.80% - 95.97%), 95.32% (95% CI 95.31% - 95.33%), 95.53% (95% CI 95.46% - 95.60%), and 94.15%, respectively. For 5-year survival prediction, the corresponding C-index values were 91.82% (95% CI 91.68% - 91.96%), 89.98% (95% CI 89.95% - 90.01%), 94.60% (95% CI 94.50% - 94.70%), and 82.26%, as summarized in Table 1.

Table . Concordance index (C-index) with 95% CIs for survival prediction models and the Cox proportional hazards (PH) model.

Model	3 year, C-index (%; 95% CI)	5 year, C-index (%; 95% CI)
RSF[a]	95.89 (95.80-95.97)	91.82 (91.68-91.96)
GBSA[b]	95.32 (95.31-95.33)	89.98 (89.95-90.01)
EST[c]	95.53 (95.46-95.60)	94.60 (94.50-94.70)
Cox PH	94.15[d]	82.26

[a]RSF: random survival forest.

[b]GBSA: gradient boosting survival analysis.

[c]EST: extra survival tree.

[d]The C-index values reported for the machine learning models represent the mean values and corresponding 95% CIs obtained from 100 repeated model training iterations. The Cox PH model was evaluated using a single model run; therefore, CIs for the C-index were not estimated or reported.

For 3-year survival prediction, the RSF model identified several important predictors, including the AJCC stage, T-size, age, serum glutamic oxaloacetic transaminase (aspartate aminotransferase), eGFR, systolic blood pressure, serum glutamic pyruvic transaminase (aspartate aminotransferase), hemoglobin level, and tumor topography codes (C162 and C160). For 5-year survival prediction, the RSF model identified AJCC stage, T-size, deep vein thrombosis, age, tumor topography code (C163), gamma-glutamyl transpeptidase, tumor grade, diastolic blood pressure, hemoglobin level, and liver disease as key prognostic predictors. For 3-year survival prediction, the GBSA model identified several important predictors, including AJCC stage, hemoglobin level, tumor topography code (C169), T-size, fasting blood glucose, total cholesterol, weekly alcohol consumption, walking activity level, systolic blood pressure, and serum glutamic oxaloacetic transaminase. For 5-year survival prediction, the GBSA model identified the following important predictors: AJCC stage, T-size, tumor topography code (C169), weekly alcohol consumption, fasting blood glucose, hemoglobin level, diastolic blood pressure, serum glutamic pyruvic transaminase, tumor grade, and total cholesterol. For 3-year survival prediction, the EST model identified several important predictors, including AJCC stage, T-size, diabetes mellitus, tumor topography codes (C169 and C162), proteinuria, waist circumference, body weight, daily alcohol consumption, and serum glutamic oxaloacetic transaminase levels. For 5-year survival prediction, the most influential predictors identified by the EST model included AJCC stage, T-size, tumor topography code (C163), deep vein thrombosis, liver disease, diabetes mellitus, tumor grade, BMI, waist circumference, and height. For 3-year mortality prediction, the Cox PH model identified AJCC stage and T-size as the most influential predictors, followed by serum glutamic oxaloacetic transaminase, eGFR, triglycerides, age, high-density lipoprotein cholesterol, hemoglobin level, total cholesterol, and gamma-glutamyl transpeptidase. For 5-year mortality prediction, AJCC stage and T-size remained the dominant predictors, together with serum glutamic oxaloacetic transaminase, eGFR, age, high-density lipoprotein cholesterol, total cholesterol, hemoglobin level, triglycerides, and weekly alcohol consumption.

Feature importance analysis showed that AJCC stage and T-size were consistently the most influential predictors across all models and prediction time points, whereas other variables demonstrated model- and time-dependent variability. Detailed feature importance results are provided in Multimedia Appendix 3, the corresponding hazard ratio analyses are presented in Multimedia Appendix 4, and Multimedia Appendix 5 provides the mapping between heat map feature labels and the corresponding clinical variables reported in the manuscript tables.

Discussion

Principal Findings

In this study, we developed ML-based survival prediction models for young patients with GC to estimate 3- and 5-year mortality. Among the evaluated models, RSF and EST demonstrated superior predictive performance compared to the Cox PH model. Across all models and prediction horizons, AJCC stage and T-size consistently emerged as the most important predictors of mortality risk.

GC has traditionally been considered a disease predominantly affecting older adults; however, recent epidemiological studies have reported an increasing incidence among younger individuals [39-41]. These findings highlight the need for tailored prognostic models for younger patient populations. Our analysis results indicated that the older patient group exhibited a poorer prognosis than the younger patient group. Furthermore, the stage-specific survival curves demonstrated adverse poorer

XSL•FO
RenderX

prognoses in the older patient cohort. However, for stage 3 and 4 disease, no statistically meaningful differences in survival prognosis were observed between the younger and older groups, as shown in Multimedia Appendix 6. Owing to age-related physiological and biological differences, treatment responses and survival outcomes may vary, which should be considered when developing survival prediction models [42,43]. Previous studies have reported promising results, demonstrating the potential of predictive modeling to generate meaningful insights into survival outcomes [5,6,44-48]. However, uncertainty remains regarding the consistency of model performance across different predictive time horizons in young patients with GC. The superior performance of ML-based models compared with the Cox PH model may be attributed to their ability to capture complex nonlinear relationships and interactions among clinical variables, which are not adequately addressed by traditional statistical approaches. These findings suggest that ML-based models may provide more accurate risk stratification, potentially supporting personalized treatment strategies and clinical decision-making for young patients with GC. In particular, the use of longitudinal clinical and comorbidity data in this study may have contributed to improved model performance, which is in line with previous studies exploring the use of ML approaches for survival prediction. Furthermore, although many predictive models have been developed using the Surveillance, Epidemiology, and End Results database, which contains comprehensive cancer registry information, there remains limited availability of longitudinal data for tracking comorbidities and patient health characteristics, such as disease diagnosis codes (*International Classification of Diseases, 10th Revision*), medication records, and routine health examination data. Such information is not consistently available in the Healthcare Examiners and Evaluators database or the National Health Service database, which may influence the predictive performance and reliability of ML models. To address these limitations, we classified patients with GC as younger (≤50 years) based on criteria used in recent studies [49] and evaluated their 3- and 5-year survival predictions using multiple survival prediction models. Using feature importance analysis, we identified clinically relevant predictors, among which 2 variables consistently influenced survival prognosis across all models and time horizons: AJCC stage and T-size. These variables demonstrated significant associations with all-cause mortality at both the 3- and 5-year follow-up intervals, as shown in Multimedia Appendix 7. The tumor, node, metastasis (TNM) staging system represents a comprehensive framework for evaluating the anatomical extent of cancer progression and continues to serve as the fundamental basis for prognostic stratification in oncology. Although T-size contributes to the T component of the TNM classification, the continuous measurement of T-size may provide additional prognostic granularity beyond categorical stage classifications. The TNM staging system aggregates information from the tumor, nodal involvement, and metastasis categories into discrete stage groups. However, a continuous T-size variable may capture heterogeneity within the same stage group, which may not be fully represented by categorical staging alone. This may prove why both tumor stage and T-size were retained as significant predictors in the ML survival prediction models. These findings suggest that, within established staging frameworks, incorporating quantitative tumor characteristics such as continuous T-size measurements can provide additional prognostic information, particularly for long-term survival prediction. However, as illustrated by the heat map of key predictors for the 3- and 5-year models, the relative importance of variables used to predict mortality varied across prediction time points in Multimedia Appendix 3. The proposed survival prediction models enabled the identification of high-risk subgroups among young patients with GC by accounting for temporal variations in prognostic risk factors. These findings may support the development of more targeted and potentially aggressive treatment strategies for young patients with GC. To evaluate the performance of the proposed model, a comparative analysis was conducted between the traditional TNM staging system and the ML survival model for 1-, 3-, and 5-year mortality prediction, with the Harrell C-index used to quantify model discrimination performance. The findings demonstrated that ML-based approaches exhibited superior performance for 3- and 5-year mortality, whereas at the 1-year prediction horizon, the TNM staging classification showed slightly higher discriminatory power than the ML survival model, as shown in Multimedia Appendix 8. Consequently, the ML-based 1-year mortality prediction model was excluded from subsequent analyses. These findings suggest that short-term mortality outcomes are predominantly influenced by the initial stage of disease, a factor that is effectively captured by the conventional TNM staging classification system. Conversely, the additional predictive benefit of ML-based survival models became more pronounced in long-term predictions (3 and 5 years), where multiple complex clinical and biological factors may contribute to heterogeneity in patient outcomes. The validation of predictive model performance was strongly influenced by the data-splitting strategy used for model training and evaluation. This study compared the results obtained using a random data-splitting strategy (training: 70%; testing: 30%) applied to the entire dataset from 2013 to 2015 with those obtained using a year-based split (training: 2013 - 2014; testing: 2015), as shown in Multimedia Appendix 9. The analysis demonstrated that the random data-splitting strategy produced a higher predictive performance compared with the year-based splitting approach. This result was interpreted as reflecting the inability of the year-based split to adequately represent the characteristics of the patient cohort collected in a specific year (2015) within the training dataset, which led to larger differences in key variables—particularly AJCC stage—between the training and testing datasets. In summary, although the dataset in this study was longitudinal and incorporated patient-specific follow-up durations, the year-based segmentation approach functioned more as a method for separating patient cohorts collected at different time points than as a validation framework representing a true temporal prediction scenario. This heterogeneity in key predictor variables—particularly AJCC stage—between the training and test datasets was considered to limit the generalization of latent patterns learned by the models during training, thereby reducing the overall predictive performance.

This study has several limitations. It was conducted using data derived from a single ethnic population, which may have introduced potential population-specific bias into the developed

XSL•FO
RenderX

prediction models. In addition, external validation using independent datasets was not performed, which may limit the robustness, generalizability, and clinical applicability of these predictive models. Although this study developed and validated ML models for predicting 3- and 5-year mortality in young patients with GC, several important methodological considerations, including model interpretability and robustness to complex or incomplete data, were not fully addressed.

In summary, this study developed and validated ML models for predicting 3- and 5-year mortality in young patients with GC, demonstrating their potential utility for risk stratification and clinical decision-making. Validation using larger and more diverse patient populations is necessary to ensure robustness, external validity, and generalizability. The integration of explainable artificial intelligence methodologies may enhance the interpretability and transparency of ML models, thereby making predictive outputs more clinically actionable for health care practitioners [50]. Recent ML research has addressed several methodological challenges, including modeling temporal dependencies, discriminative feature selection, and improving robustness to noisy or incomplete input data. Prior studies in nonclinical ML domains have proposed advanced methodologies, including attention-based temporal models, structure-aware feature selection frameworks, and confidence-driven learning strategies [51-53]. Although these approaches do not directly fall within the scope of this study, they provide conceptually important insights for survival prediction problems, particularly those requiring careful handling of longitudinal data patterns, high-dimensional clinical variables, and uncertainty in medical datasets. These methodological directions may be adapted and extended to medical prognostic prediction models, particularly in survival analysis and clinical outcome prediction research. In addition to these methodological considerations, evaluation of the clinical utility of the key prognostic variables identified in this study, including cancer stage and T-size, is required through prospective or independent validation studies, whereas the use of internationally standardized and multiethnic datasets may further improve the reliability, external validity, and global generalizability of the predictive models.

Conclusions

The number of young patients diagnosed with GC continues to increase worldwide, and nomogram-based prognostic models have been proposed to identify patients at high risk of adverse outcomes. However, relatively few studies have focused on predicting survival outcomes in young patients with GC using ML approaches. Therefore, we developed ML-based prediction models for estimating 3- and 5-year mortality in young patients with GC. These models may assist clinicians in identifying young patients with GC who may require more aggressive or personalized treatment strategies. However, because this analysis was conducted using data derived from the Korean patient population, future studies should externally validate these prediction models in diverse international populations.

Acknowledgments

The authors attest that there was no use of generative artificial intelligence technology in the generation of text, figures, or other informational content of this manuscript.

Funding

This study was supported by a grant (2310440-4) from the National Cancer Center of South Korea.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy protection requirements and national data governance restrictions for cancer registry–linked health data. However, researchers may request access to the original dataset through the Korea Clinical Data Utilization Network for Research Excellence portal [54], subject to registration and approval through the required data application and review process. Applicants must submit a data access application form, including a detailed research proposal describing the intended use of the dataset, which is reviewed and approved by both the Korea Clinical Data Utilization Network for Research Excellence and the National Cancer Data Center. This data access service is currently restricted to researchers based in South Korea. All source code associated with this study is publicly available through the GitHub repository [55].

Authors' Contributions

Conceptualization: HYJK, MK, KSR
Data curation: HYJK, WJ, KSR
Investigation: HYJK
Methodology: HYJK, MK, KSR
Validation: HYJK, WJ, MK, KSR
Writing—original draft: HYJK, KSR
All authors contributed to drafting and revising the manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Baseline characteristics of the older and younger patient groups.
[DOCX File, 31 KB - cancer_v12i1e86418_app1.docx]

Multimedia Appendix 2
C-index values of the survival prediction models.
[DOCX File, 38 KB - cancer_v12i1e86418_app2.docx]

Multimedia Appendix 3
Relative feature importance of predictors in the survival prediction model.
[DOCX File, 103 KB - cancer_v12i1e86418_app3.docx]

Multimedia Appendix 4
Univariate analysis results for all causes of death.
[DOCX File, 34 KB - cancer_v12i1e86418_app4.docx]

Multimedia Appendix 5
Mapping between heat map feature labels and clinical variables.
[DOCX File, 25 KB - cancer_v12i1e86418_app5.docx]

Multimedia Appendix 6
Survival curve for younger and older patients.
[DOCX File, 226 KB - cancer_v12i1e86418_app6.docx]

Multimedia Appendix 7
Survival curve for stage and tumor size.
[DOCX File, 175 KB - cancer_v12i1e86418_app7.docx]

Multimedia Appendix 8
Comparison of predictive performance between American Joint Committee on Cancer stage and survival machine learning models.
[DOCX File, 17 KB - cancer_v12i1e86418_app8.docx]

Multimedia Appendix 9
Comparison of year-based data splitting and the 7:3 random split.
[DOCX File, 34 KB - cancer_v12i1e86418_app9.docx]

References

1. Wong MCS, Huang J, Chan PSF, et al. Global incidence and mortality of gastric cancer, 1980-2018. JAMA Netw Open 2021 Jul 1;4(7):e2118457. [doi: 10.1001/jamanetworkopen.2021.18457] [Medline: 34309666]
2. Wu SL, Zhang Y, Fu Y, Li J, Wang JS. Gastric cancer incidence, mortality and burden in adolescents and young adults: a time-trend analysis and comparison among China, South Korea, Japan and the USA. BMJ Open 2022 Jul 21;12(7):e061038. [doi: 10.1136/bmjopen-2022-061038] [Medline: 35863834]
3. Gao K, Wu J. National trend of gastric cancer mortality in China (2003–2015): a population‐based study. Cancer Commun 2019 Dec;39(1):1-5. [doi: 10.1186/s40880-019-0372-x]
4. Yang Q, Xu D, Yang Y, Lu S, Wang D, Wang L. Global, regional, and national burden of gastric cancer in adolescents and young adults, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Am J Gastroenterol 2024 Mar 1;119(3):454-467. [doi: 10.14309/ajg.0000000000002551] [Medline: 37800697]
5. Wu C, Wang N, Zhou H, Wang T, Zhao D. Development and validation of a nomogram to individually predict survival of young patients with nonmetastatic gastric cancer: a retrospective cohort study. Saudi J Gastroenterol 2019;25(4):236-244. [doi: 10.4103/sjg.SJG_378_18] [Medline: 30719999]
6. Chen YR, Tian ZY, Wang MQ, Sun ML, Wu JZ, Wang XY. Development and validation of prognostic nomograms based on lymph node ratio for young patients with gastric cancer: a SEER-based study. Technol Cancer Res Treat 2023;22:15330338231157923. [doi: 10.1177/15330338231157923] [Medline: 37248609]

XSL•FO
RenderX

7. Lee W, Lam SK, Zhang Y, Yang R, Cai J. Review of methodological workflow, interpretation and limitations of nomogram application in cancer study. Radiat Med Prot 2022 Dec;3(4):200-207. [doi: 10.1016/j.radmp.2022.08.004]
8. Tran TT, Lee J, Gunathilake M, et al. A comparison of machine learning models and Cox proportional hazards models regarding their ability to predict the risk of gastrointestinal cancer based on metabolic syndrome and its components. Front Oncol 2023;13:1049787. [doi: 10.3389/fonc.2023.1049787] [Medline: 36937438]
9. Kurt Omurlu I, Ture M, Tokatli F. The comparisons of random survival forests and Cox regression analysis with simulation and an application related to breast cancer. Expert Syst Appl 2009 May;36(4):8582-8588. [doi: 10.1016/j.eswa.2008.10.023]
10. Díez-Sanmartín C, Sarasa Cabezuelo A. Application of artificial intelligence techniques to predict survival in kidney transplantation: a review. J Clin Med 2020 Feb 19;9(2):572. [doi: 10.3390/jcm9020572] [Medline: 32093027]
11. Wang J, Chen N, Guo J, Xu X, Liu L, Yi Z. SurvNet: a novel deep neural network for lung cancer survival analysis with missing values. Front Oncol 2021;10:588990. [doi: 10.3389/fonc.2020.588990]
12. Mainali G. Artificial intelligence in medical science: perspective from a medical student. JNMA J Nepal Med Assoc 2020 Sep 27;58(229):709-711. [doi: 10.31729/jnma.5257] [Medline: 33068098]
13. Seifert R, Weber M, Kocakavuk E, Rischpler C, Kersting D. Artificial intelligence and machine learning in nuclear medicine: future perspectives. Semin Nucl Med 2021 Mar;51(2):170-177. [doi: 10.1053/j.semnuclmed.2020.08.003] [Medline: 33509373]
14. Lu T, Fang Y, Liu H, et al. Comparison of machine learning and logic regression algorithms for predicting lymph node metastasis in patients with gastric cancer: a two-center study. Technol Cancer Res Treat 2024;23:15330338231222331. [doi: 10.1177/15330338231222331] [Medline: 38190617]
15. Alabi RO, Mäkitie AA, Pirinen M, Elmusrati M, Leivo I, Almangush A. Comparison of nomogram with machine learning techniques for prediction of overall survival in patients with tongue cancer. Int J Med Inform 2021 Jan;145:104313. [doi: 10.1016/j.ijmedinf.2020.104313] [Medline: 33142259]
16. Huang Y, Li J, Li M, Aparasu RR. Application of machine learning in predicting survival outcomes involving real-world data: a scoping review. BMC Med Res Methodol 2023 Nov 13;23(1):268. [doi: 10.1186/s12874-023-02078-1] [Medline: 37957593]
17. Chen QY, Zhong Q, Wang W, et al. Prognosis of young survivors of gastric cancer in China and the U.S.: determining long-term outcomes based on conditional survival. Oncologist 2019 Jun;24(6):e260-e274. [doi: 10.1634/theoncologist.2018-0220] [Medline: 30470692]
18. National Cancer Data Center. URL: https://www.cancerdata.re.kr/en/index [accessed 2024-05-21]
19. Korea Central Cancer Registry. KCCR survey. Korea Central Cancer Registry. URL: https://kccrsurvey.cancer.go.kr/index.do [accessed 2024-05-21]
20. Choi DW, Guk MY, Kim HR, et al. Data resource profile: the Cancer Public Library Database in South Korea. Cancer Res Treat 2024 Oct;56(4):1014-1026. [doi: 10.4143/crt.2024.207] [Medline: 38697846]
21. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. Ann Appl Stat 2008;2(3):841-860. [doi: 10.1214/08-AOAS169]
22. Mogensen UB, Ishwaran H, Gerds TA. Evaluating random forests for survival analysis using prediction error curves. J Stat Softw 2012 Sep;50(11):1-23. [doi: 10.18637/jss.v050.i11] [Medline: 25317082]
23. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Statist 2001;29(5):1189-1232. [doi: 10.1214/aos/1013203451]
24. Chen Y, Jia Z, Mercola D, Xie X. A gradient boosting algorithm for survival analysis via direct optimization of concordance index. Comput Math Methods Med 2013;2013:873595. [doi: 10.1155/2013/873595] [Medline: 24348746]
25. Bai M, Zheng Y, Shen Y. Gradient boosting survival tree with applications in credit scoring. J Oper Res Soc 2022;73(1):39-55. [doi: 10.1080/01605682.2021.1919035]
26. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn 2006 Apr;63:3-42. [doi: 10.1007/s10994-006-6226-1]
27. Zaenal MS, Fitrianto A, Wijayanto H. Comparison of extremely randomized survival trees and random survival forests: a simulation study. Sci J Inform 2024;11(3):635-644. [doi: 10.15294/sji.v11i3.8464]
28. Cox DR. Regression models and life-tables. J R Stat Soc Ser B Methodol 1972 Jan;34(2):187-202. [doi: 10.1111/j.2517-6161.1972.tb00899.x]
29. Cygu S, Seow H, Dushoff J, Bolker BM. Comparing machine learning approaches to incorporate time-varying covariates in predicting cancer survival time. Sci Rep 2023 Jan 25;13(1):1370. [doi: 10.1038/s41598-023-28393-7] [Medline: 36697455]
30. Schober P, Vetter TR. Survival analysis and interpretation of time-to-event data: the tortoise and the hare. Anesth Analg 2018 Sep;127(3):792-798. [doi: 10.1213/ANE.0000000000003653] [Medline: 30015653]
31. Deng Y, Qin HY, Zhou YY, et al. Artificial intelligence applications in pathological diagnosis of gastric cancer. Heliyon 2022 Dec;8(12):e12431. [doi: 10.1016/j.heliyon.2022.e12431] [Medline: 36619448]
32. Pickett KL, Suresh K, Campbell KR, Davis S, Juarez-Colunga E. Random survival forests for dynamic predictions of a time-to-event outcome using a longitudinal biomarker. BMC Med Res Methodol 2021 Oct 17;21(1):216. [doi: 10.1186/s12874-021-01375-x] [Medline: 34657597]

XSL•FO
RenderX

33. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825-2830 [FREE Full text]
34. Pölsterl S. scikit-survival: a library for time-to-event analysis built on top of scikit-learn. J Mach Learn Res 2020;21(212):1-6 [FREE Full text]
35. Harrell FEJ, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA 1982 May 14;247(18):2543-2546. [Medline: 7069920]
36. Longato E, Vettoretti M, Di Camillo B. A practical perspective on the concordance index for the evaluation and selection of prognostic time-to-event models. J Biomed Inform 2020 Aug;108:103496. [doi: 10.1016/j.jbi.2020.103496] [Medline: 32652236]
37. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med 2011 May 10;30(10):1105-1117. [doi: 10.1002/sim.4154] [Medline: 21484848]
38. Schmid M, Wright MN, Ziegler A. On the use of Harrell's C for clinical risk prediction via random survival forests. Expert Syst Appl 2016 Nov;63:450-459. [doi: 10.1016/j.eswa.2016.07.018]
39. Arnold M, Park JY, Camargo MC, Lunet N, Forman D, Soerjomataram I. Is gastric cancer becoming a rare disease? A global assessment of predicted incidence trends to 2035. Gut 2020 May;69(5):823-829. [doi: 10.1136/gutjnl-2019-320234] [Medline: 32001553]
40. Li J. Gastric cancer in young adults: a different clinical entity from carcinogenesis to prognosis. Gastroenterol Res Pract 2020;2020:9512707. [doi: 10.1155/2020/9512707] [Medline: 32190044]
41. Heer EV, Harper AS, Sung H, Jemal A, Fidler-Benaoudia MM. Emerging cancer incidence trends in Canada: the growing burden of young adult cancers. Cancer 2020 Oct 15;126(20):4553-4562. [doi: 10.1002/cncr.33050] [Medline: 32770762]
42. Song P, Wu L, Jiang B, Liu Z, Cao K, Guan W. Age-specific effects on the prognosis after surgery for gastric cancer: a SEER population-based analysis. Oncotarget 2016 Jul 26;7(30):48614-48624. [doi: 10.18632/oncotarget.9548] [Medline: 27224925]
43. Zhang H, Cheng X, Guo W, et al. Metastasis patterns and prognosis in young gastric cancer patients: a propensity score-matched SEER database analysis. PLoS One 2024;19(4):e0301834. [doi: 10.1371/journal.pone.0301834]
44. Fryan LH, Alazzam MB. Survival analysis of oncological patients using machine learning method. Healthcare (Basel) 2022 Dec 27;11(1):80. [doi: 10.3390/healthcare11010080] [Medline: 36611540]
45. Tizi W, Berrado A. Machine learning for survival analysis in cancer research: a comparative study. Sci Afr 2023 Sep;21:e01880. [doi: 10.1016/j.sciaf.2023.e01880]
46. Wu M, Yang X, Liu Y, et al. Development and validation of a deep learning model for predicting postoperative survival of patients with gastric cancer. BMC Public Health 2024 Mar 6;24(1):723. [doi: 10.1186/s12889-024-18221-6] [Medline: 38448849]
47. Afrash MR, Mirbagheri E, Mashoufi M, Kazemi-Arpanahi H. Optimizing prognostic factors of five-year survival in gastric cancer patients using feature selection techniques with machine learning algorithms: a comparative study. BMC Med Inform Decis Mak 2023 Apr 6;23(1):54. [doi: 10.1186/s12911-023-02154-y] [Medline: 37024885]
48. Zhang C, Zhang Y, Yang YH, et al. Machine learning models for predicting one-year survival in patients with metastatic gastric cancer who experienced upfront radical gastrectomy. Front Mol Biosci 2022;9:937242. [doi: 10.3389/fmolb.2022.937242] [Medline: 36533072]
49. Koh B, Tan DJ, Ng CH, et al. Patterns in cancer incidence among people younger than 50 years in the US, 2010 to 2019. JAMA Netw Open 2023 Aug 1;6(8):e2328171. [doi: 10.1001/jamanetworkopen.2023.28171] [Medline: 37585204]
50. Sadeghi Z, Alizadehsani R, Cifci MA, et al. A review of explainable artificial intelligence in healthcare. Comput Electr Eng 2024 Aug;118:109370. [doi: 10.1016/j.compeleceng.2024.109370]
51. Zhou H, Ren D, Xia H, Fan M, Yang X, Huang H. AST-GNN: an attention-based spatio-temporal graph neural network for Interaction-aware pedestrian trajectory prediction. Neurocomputing 2021 Jul;445:298-308. [doi: 10.1016/j.neucom.2021.03.024]
52. Fan M, Zhang X, Hu J, Gu N, Tao D. Adaptive data structure regularized multiclass discriminative feature selection. IEEE Trans Neural Netw Learn Syst 2022;33(10):5859-5872. [doi: 10.1109/TNNLS.2021.3071603]
53. Li X, Huang H, Zhao H, Wang Y, Hu M. Learning a convolutional neural network for propagation-based stereo image segmentation. Vis Comput 2020 Jan;36:39-52. [doi: 10.1007/s00371-018-1582-y]
54. Korea-Clinical data Utilization network for Research Excellence. URL: https://k-cure.mohw.go.kr [accessed 2024-05-22]
55. Predicting mortality in young patients with gastric cancer. GitHub. URL: https://github.com/KwangSun-Ryu/gastric_cancer_mortality [accessed 2026-05-10]

Abbreviations

AJCC: American Joint Committee on Cancer
CPSD: Cancer Public Staging Database
eGFR: estimated glomerular filtration rate

XSL•FO
RenderX

EST: extra survival tree
GBSA: gradient boosting survival analysis
GC: gastric cancer
KCCR: Korea Central Cancer Registry
ML: machine learning
NCDC: National Cancer Data Center
PH: proportional hazards
RSF: random survival forest
TNM: tumor, node, metastasis

Edited by M Balcarras; submitted 23.Oct.2025; peer-reviewed by H Chen, YH Lai; revised version received 01.Apr.2026; accepted 02.Apr.2026; published 26.May.2026.

Please cite as:
Kang HYJ, Jang W, Ko M, Ryu KS
Machine Learning–Based Survival Prediction Models for Young Patients With Gastric Cancer: Model Development and Validation Study
JMIR Cancer 2026;12:e86418
URL: https://cancer.jmir.org/2026/1/e86418
doi:10.2196/86418

© Ha Ye Jin Kang, Wooyeong Jang, Minsam Ko, Kwang Sun Ryu. Originally published in JMIR Cancer (https://cancer.jmir.org), 26.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Prognostic Value of Dynamic Lactate Dehydrogenase Trends in Immunotherapy for Advanced Esophageal Squamous Cell Carcinoma: Retrospective Cohort Study

Baidong Zhang[1*], MM; Zhichao Kang[1*], MM; Yi Ding[1], MM; Wang Jing[1], MD; Alei Feng[1], MD; Renya Zeng[1], MD; Jianan Li[1], MM; Yi Zhao[1], MM; Yuanliu Nie[2], MM; Wentao Zhang[2], MM; Lu Sun[1], MD; Zhe Yang[1], MD

[1]Tumor Research and Therapy Center, Shandong Provincial Hospital Affiliated to Shandong First Medical University, 324 Jingwu Road, Jinan, China

[2]Department of Radiation Oncology, Shandong Provincial Hospital, Jinan, China

[*]these authors contributed equally

Corresponding Author:
Zhe Yang, MD
Tumor Research and Therapy Center, Shandong Provincial Hospital Affiliated to Shandong First Medical University, 324 Jingwu Road, Jinan, China

Abstract

Background: Immune checkpoint inhibitors (ICIs) have emerged as a pivotal treatment for advanced esophageal squamous cell carcinoma (ESCC). However, their efficacy can significantly differ among patients, highlighting the need for reliable prognostic markers to enhance treatment outcomes. Lactate dehydrogenase (LDH) plays a key regulatory role in the complex relationship between cancer metabolism and the immune system, suggesting that monitoring LDH levels may provide valuable insights into treatment efficacy and inform personalized therapeutic strategies for advanced ESCC.

Objective: This study aimed to explore the prognostic significance of dynamic changes in LDH levels during ICI therapy in predicting treatment outcomes.

Methods: We retrospectively analyzed the clinical data of 126 patients with advanced ESCC who received first-line ICI therapy at the Department of Radiation Oncology, Cancer Center, Shandong Provincial Hospital, between April 2018 and November 2022. Serum LDH levels were measured after every 3 cycles of combined immunotherapy and chemotherapy. Receiver operating characteristic curve analysis determined the optimal LDH reduction threshold. Kaplan-Meier survival curves and Cox regression models assessed progression-free survival (PFS) and overall survival.

Results: Among the 126 patients, 55 (43.6%) were classified into the LDH-increased group, while 71 (56.4%) belonged to the LDH-decreased group. Within the LDH-increased group, 78.2% (43/55) of the patients were male, compared to 90.1% (64/71) in the LDH-decreased group. The median age of patients in the LDH-increased group was 59 (range 55 - 68) years, whereas the median age in the LDH-decreased group was 65 (range 58 - 65) years. LDH decrease following first-line ICI therapy was associated with improved outcomes compared to LDH increases (median PFS 13.4, IQR 8.1 - 24.3 mo vs median 10.8, IQR 4.8 - 20.6 mo; P= .03). Patients with a posttreatment LDH decrease of more than 14.4% had a median PFS of 11.1 (IQR 7.2 - 24.3) months, whereas those with an LDH decrease between 0% and 14.4% had a median PFS of 21.7 (IQR 9.4 - 34.5) months. Conversely, an increase in LDH resulted in a median PFS of 10.8 (IQR 4.8 - 20.6) months. Patients with tumor reduction exhibited a significantly greater decrease in LDH levels compared with those without tumor reduction (P<.001). Multivariate analysis identified LDH decrease as an independent predictor of a 41% lower mortality risk (hazard ratio 0.59, 95% CI 0.36 - 0.96; P=.04).

Conclusions: In patients with advanced ESCC, a decrease in serum LDH levels ranging from 0% to 14.4% after treatment initiation was significantly associated with prolonged PFS. Notably, an early decrease in LDH levels observed after 3 cycles of immunotherapy further correlated with improved clinical outcomes. These results highlight the potential of LDH as a valuable biomarker for risk stratification and personalized treatment optimization in advanced ESCC.

(JMIR Cancer 2026;12:e73576) doi:10.2196/73576

KEYWORDS

lactate dehydrogenase; esophageal squamous cell carcinoma; immune checkpoint inhibitors; prognostic value; biomarker

XSL•FO
RenderX

Introduction

Background

Esophageal squamous cell carcinoma (ESCC) is a globally prevalent oncological challenge marked by high incidence and mortality rates, particularly in regions such as Asia, Eastern Europe, and Africa [1]. The typically asymptomatic nature of early-stage ESCC often results in late-stage diagnoses, which limit therapeutic options and contribute to high mortality rates [2]. Current management strategies for ESCC include a combination of surgery, radiation, and chemotherapy. However, the 5-year survival rate remains dismally low at approximately 20%, emphasizing the aggressive nature of this malignancy and the complexities surrounding its treatment [3].

Recent advancements in immunotherapy, particularly immune checkpoint inhibitors (ICIs) targeting the programmed death-1 (PD-1) and programmed death-ligand 1 (PD-L1) axis, such as nivolumab and pembrolizumab, have revolutionized the treatment landscape for advanced ESCC. Evidence from pivotal studies such as KEYNOTE-181 and ATTRACTION-3 has demonstrated the superior efficacy of ICIs compared to conventional chemotherapy [4]. Despite these breakthroughs, not all patients with advanced ESCC derive significant benefits from ICIs [5], highlighting the critical need for reliable, accessible, and cost-effective biomarkers to predict treatment response and patient prognosis more effectively.

Lactate dehydrogenase (LDH), a key enzyme in cancer cell metabolism, has gained attention as a potential prognostic biomarker in oncology due to its role in metabolic reprogramming and tumor adaptation to hypoxic microenvironments [6]. Elevated LDH levels have been linked to increased tumor aggressiveness, metabolic stress, and poor clinical outcomes across various malignancies, including melanoma, lung cancer, and breast cancer, underscoring its potential as a universal biomarker of malignancy [7,8]. Emerging research highlights the intricate relationship between cancer metabolism and the immune system, with LDH functioning as a critical mediator of this interplay [9]. LDH plays a pivotal role in metabolic reprogramming, enabling tumor survival in hypoxic conditions by facilitating anaerobic glycolysis and lactate production. These processes not only support tumor growth but also contribute to the creation of an immunosuppressive tumor microenvironment. This suggests that dynamic changes in LDH levels could significantly influence the efficacy of immunotherapeutic approaches in ESCC [10]. Preliminary studies have shown that variations in LDH levels following ICI treatment correlate with improved survival outcomes, pointing to the potential of LDH as a valuable prognostic tool in ESCC immunotherapy [11].

Objectives

To identify a more accessible and precise prognostic marker for advanced ESCC, this study enrolled patients receiving first-line immunotherapy combined with chemotherapy. The primary objective was to investigate the prognostic value of dynamic changes in LDH by assessing LDH kinetics. This research aims to provide new insights into the potential of LDH as a cost-effective and practical biomarker for guiding personalized treatment strategies and improving outcomes in patients with advanced ESCC.

Methods

Ethical Considerations

This study was approved by the Ethics Committee of Biomedical Research at Shandong Provincial Hospital (SWYX NO 2023 - 595). Given the retrospective nature of the study, which involved secondary analysis of existing electronic health records, no prospective informed consent specific to this research was required. At the time of admission or outpatient registration, all patients provided a general consent permitting the use of their medical records for research purposes, in accordance with institutional policy. All data in this study were deidentified to ensure participant anonymity and protect confidentiality. This was a retrospective study based on the analysis of pre-existing patient medical records. As such, no participants were actively recruited for this research, and therefore, no compensation was involved.

Patient Selection

This retrospective cohort study was conducted at Shandong Provincial Hospital between April 2018 and November 2022. The study included patients diagnosed with stage IV ESCC based on the Tumor-Node-Metastasis classification system. Eligible patients were those who received ICIs in combination with chemotherapy and who had a Karnofsky Performance Status score greater than 80.

The exclusion criteria were as follows: history of other malignancies, presence of autoimmune disorders, previous immunotherapy with agents such as Carrelizumab or Tislelizumab, and incomplete clinical or laboratory data.

The follow-up period ended in September 2023. The medical records were retrieved from the Shandong Provincial Hospital database. Data collected included patient age, sex, comorbidities, tumor location, treatment response evaluations, TNM stages, stage at initial diagnosis, baseline LDH levels, posttreatment LDH levels, and dynamic changes in LDH levels.

Treatment and Evaluation Criteria

All participants in the study received a treatment regimen consisting of PD-1 inhibitors combined with chemotherapy agents, such as 5-fluorouracil, cisplatin, taxanes, and irinotecan, administered either as monotherapy or as part of polychemotherapy protocols. The PD-1 inhibitors were administered intravenously at a standard dose of 200 mg, with infusions performed every 2 to 3 weeks for a total of 4 to 6 cycles, or until disease progression, unacceptable toxicity, or death.

Serum LDH levels were measured at 2 key time points: at baseline (before the initiation of immunotherapy) and after the final immunotherapy cycle. LDH quantification was performed by the Department of Laboratory Medicine at Shandong Provincial Hospital. The study evaluated treatment efficacy longitudinally, with key end points including progression-free survival (PFS) and overall survival (OS). Progressive disease was defined as a ≥20% increase in the sum of the longest

XSL•FO
RenderX

diameters of target lesions from the nadir, the appearance of new measurable disease foci, or other substantial indicators of disease progression, according to the Response Evaluation Criteria in Solid Tumors (version 1.1) [12]. PFS was measured from the initiation of anti–PD-1 therapy to the point of documented disease progression or death from any cause, while OS was defined as the time from the initiation of immunotherapy to death from any cause.

Statistical Analysis

Overview

Descriptive statistics were used to summarize baseline characteristics, using means (SD) for normally distributed variables, medians (IQR) for non–normally distributed variables (as assessed by the Shapiro-Wilk test), and frequencies (%) for categorical variables. Serum LDH levels were measured both before treatment and after immunotherapy (median of 3 cycles). The LDH reduction rate was calculated as follows:

$$[\text{Baseline LDH} - \text{Post-treatment LDH}]/\text{Baseline LDH} \times 100\%$$

Critical Methodological Note on Cutoff Selection

The initial receiver operating characteristic analysis for OS prediction yielded suboptimal performance (area under the curve=0.602; data not shown), likely due to the heterogeneity inherent in retrospective cohorts. Consequently, we adopted a distribution-driven approach, selecting the 25th percentile of observed LDH reduction values (−14.4%) as the primary cutoff. This threshold, representing a substantial decline in LDH, demonstrated significant discriminative power for PFS in subsequent analyses.

Kaplan-Meier curves, analyzed using log-rank tests, were used to compare survival between different LDH trajectory groups. Cox regression analysis was used to identify prognostic factors, with a two-step process: (1) univariate screening with a $P<.10$ entry threshold and (2) multivariate adjustment for age, Karnofsky Performance Status, and TNM stage. Results were reported as adjusted hazard ratios (HRs) with 95% CIs. Missing data were minimal (<1% for LDH, PFS, and OS) and were handled using complete-case analysis, confirmed to be missing completely at random; sensitivity analyses with multiple imputation yielded consistent results. The sample size (n=126) provided 80% power ($\alpha=.05$) to detect an HR of 0.60 for PFS improvement in LDH reducers versus nonreducers, based on KEYNOTE-181 data with a 15% attrition adjustment. Analyses were conducted using SPSS (version 25.0; IBM Corp), with statistical significance set at $P<.05$ (2-tailed).

Results

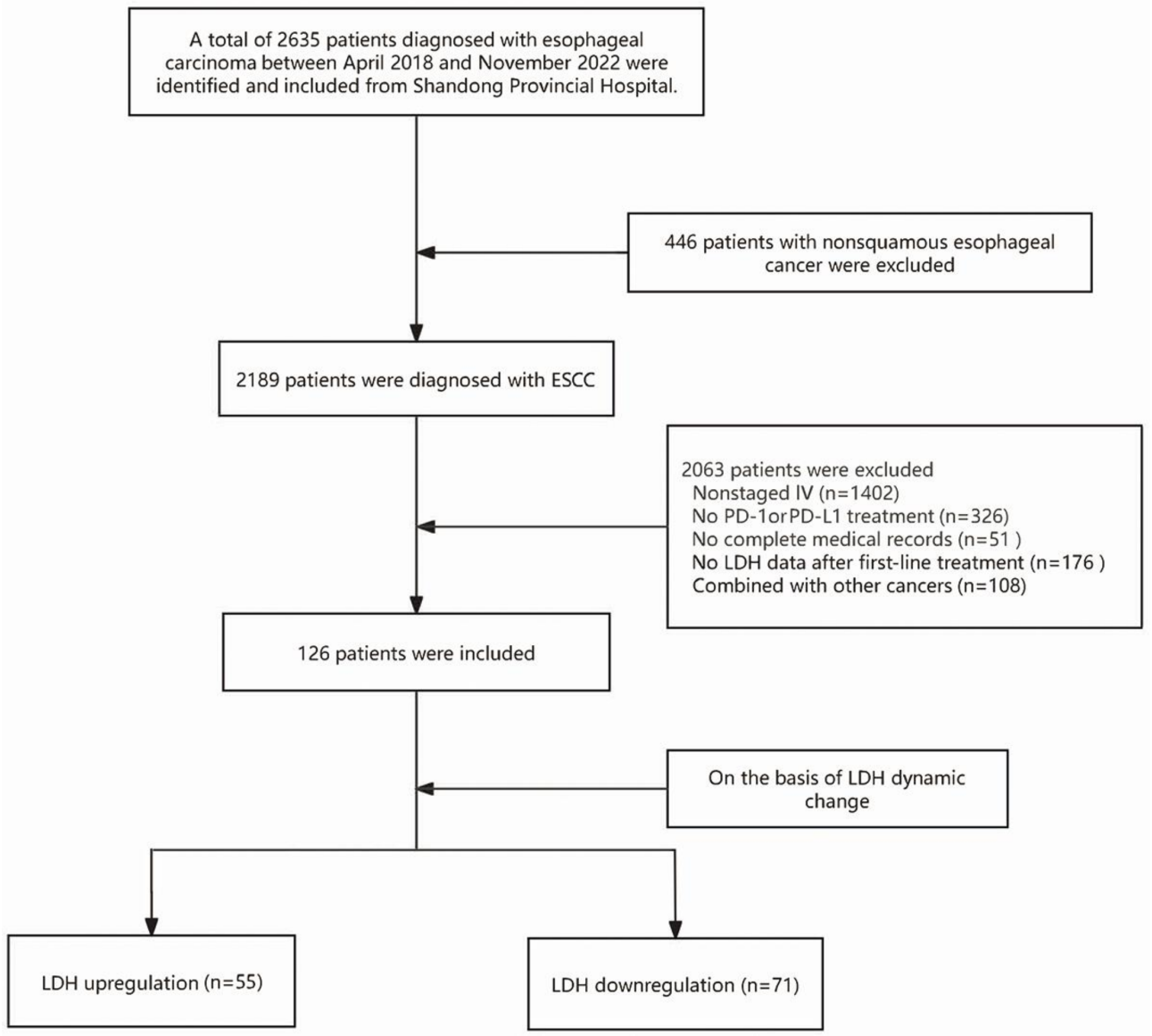
Patient Characteristics

As shown in Figure 1, from an initial cohort of 2635 patients diagnosed with ESCC, 446 (16.9%) patients with non-ESCC diagnoses, 1402 (53.2%) patients at non-IV stages of ESCC, 326 (12.4%) patients who did not receive PD-L1 and PD-1 treatment, 51 (1.94%) patients with incomplete medical records, 176 (6.68%) patients lacking LDH data during first-line treatment, and 108 (4.1%) patients diagnosed with other cancers were excluded. The final study cohort consisted of 126 patients.

The follow-up period concluded with a median follow-up duration of 13.8 (range 8.00 - 18.75) months. From the 126 patients, 9 (7.14%) were alive at the time of follow-up. The median PFS for all patients was 11.8 (95% CI 6.1 - 22.9) months, while the median OS was 28 (95% CI 14.6 - 43.9) months.

XSL•FO
RenderX

Figure 1. Patient selection criteria and process. ESCC: esophageal squamous cell carcinoma; LDH: lactate dehydrogenase; PD-1: programmed death-1; PD-L1: programmed death-ligand 1.



Patient characteristics are summarized in Table 1. Among the 126 patients, 55 (43.6%) were classified into the LDH-increased group, while 71 (56.4%) belonged to the LDH-decreased group. Within the LDH-increased group, 78.2% (43/55) of the patients were male, compared to 90.1% (64/71) in the LDH-decreased group. The median age of patients in the LDH-increased group was 59 (range 55 - 68) years, whereas the median age in the LDH-decreased group was 65 (range 58 - 65) years. In addition, in the LDH-increased group, 38.2% (21/55) of patients had other comorbidities, whereas in the LDH-decreased group, the proportion was 43.7% (31/71).

The LDH ratio was calculated by comparing the posttreatment LDH levels to the baseline LDH levels. On the basis of this ratio, patients were subsequently classified into 2 categories: those with elevated LDH levels and those with decreased LDH levels.

XSL•FO
RenderX

Table . Patient characteristics—comparison of lactate dehydrogenase (LDH) levels after treatment and at baseline.

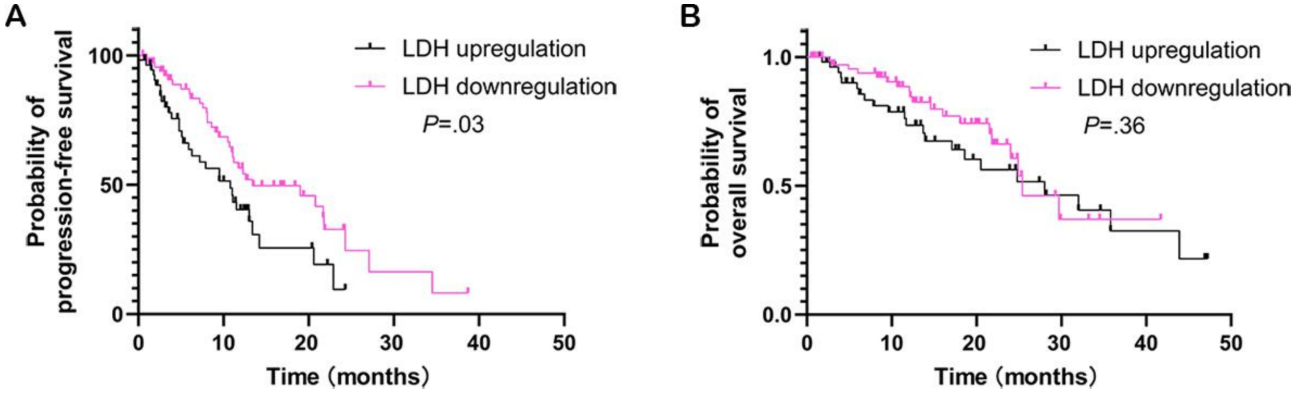
	LDH upregulation (n=55), n (%)	LDH downregulation (n=71), n (%)	*P* value
Age (years), median (IQR)	59 (55 - 68)	65 (58 - 65)	—[a]
Age (years), n (%)			.07
≤63	33 (60.0)	30 (42.3)	
>63	22 (40.0)	41 (57.7)	
Gender, n (%)			.08
Male	43 (78.2)	64 (90.1)	
Female	12 (21.8)	7 (9.9)	
Tumor location, n (%)			.81
Upper	5 (9.1)	5 (7.0)	
Middle	28 (50.9)	40 (56.3)	
Lower	22 (40.0)	26 (36.6)	
T stage, n (%)			.59
T1-3	25 (45.4)	28 (39.4)	
T4	30 (54.6)	43 (60.6)	
N stage, n (%)			.24
N0-1	8 (14.5)	5 (7.0)	
N2-3	47 (85.5)	66 (93.0)	
Clinical stage, n (%)			.57
IVA	16 (29.1)	25 (35.2)	
IVB	39 (70.9)	46 (64.8)	
Comorbidity, n (%)			.07
Yes	21 (38.2)	31 (43.7)	
No	34 (61.8)	40 (56.3)	

[a]Not applicable.

Kaplan-Meier Survival Analysis

To assess dynamic changes in LDH levels, collection points were established at intervals of every 3cycles of chemotherapy combined with immunotherapy. Dichotomous analysis revealed that patients exhibiting downregulated LDH levels showed a significant improvement in PFS (*P*=.03; Figure 2A). However, no statistically significant difference was observed in OS (*P*=.36; Figure 2B).

Figure 2. Kaplan-Meier curves for (A) progression-free survival and (B) overall survival based on lactate dehydrogenase (LDH) dichotomization (downregulation vs upregulation). Patients with downregulated LDH levels exhibited a significantly better prognosis.

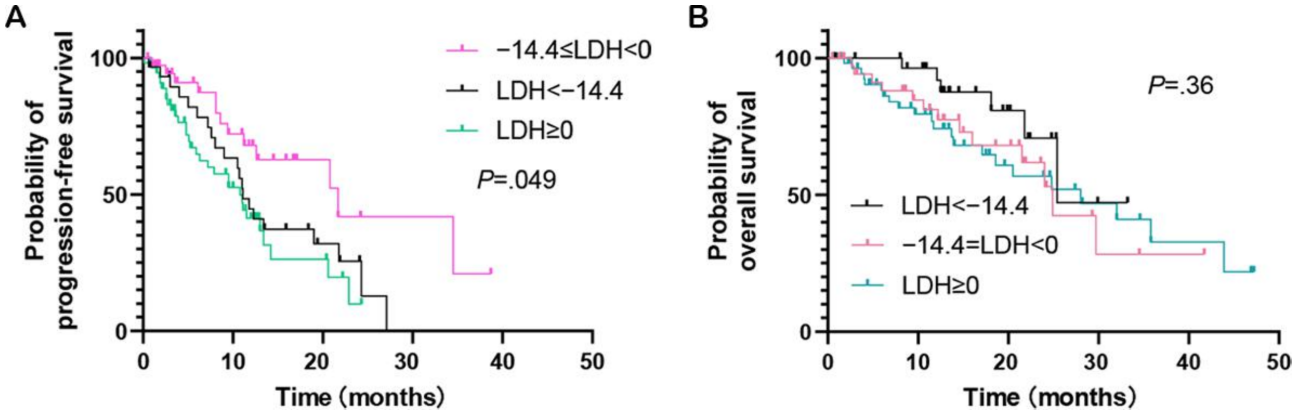


It was noted that a reduction in LDH levels may also reflect a decline in the patient's physical condition or nutritional status, necessitating further stratified analysis. Receiver operating characteristic curve methodology was applied, with OS used as the end point criterion for analysis. On the basis of this methodology, an appropriate threshold for LDH reduction was

XSL•FO
RenderX

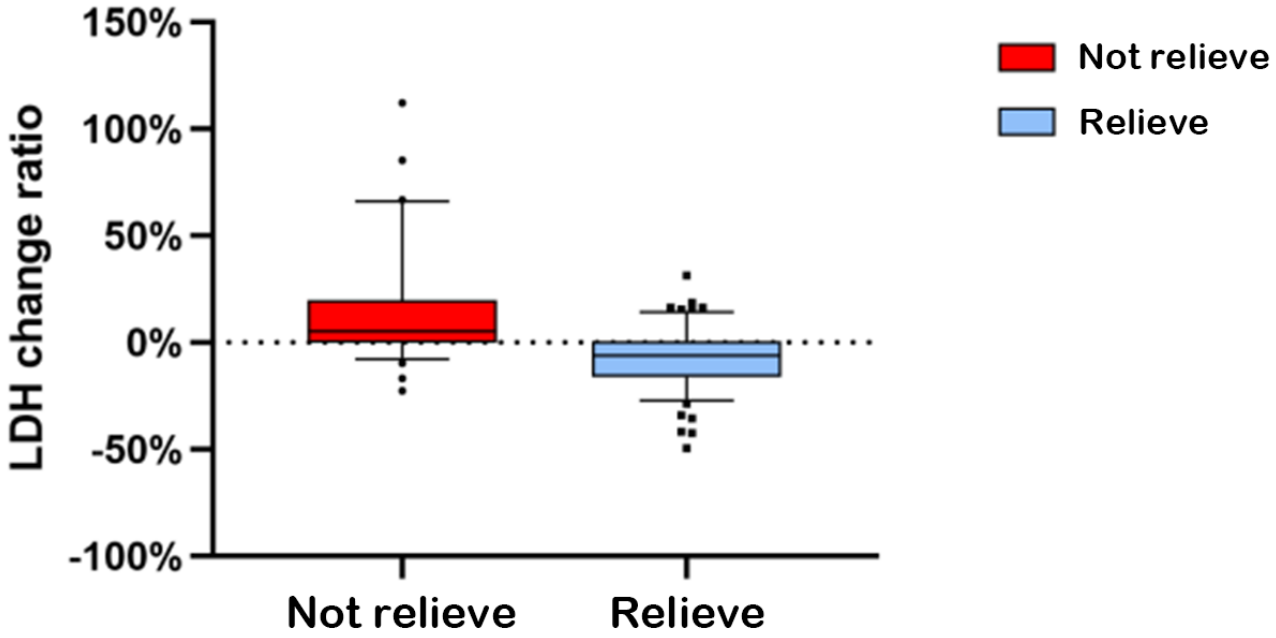
identified as greater than −14%, closely approximating the −14.4% mark, which represents the boundary at the 25th percentile of the data distribution. At the threshold of −14.4%, a significant difference in PFS was observed (Figure 3A), while no significant difference was detected for OS (Figure 3B).

Figure 3. Kaplan-Meier curves of (A) progression-free survival and (B) overall survival stratified by lactate dehydrogenase (LDH) change into 3 groups: an increase (LDH≥0%), a moderate decrease (−14.4%≤LDH<0%), and a marked decrease (LDH<−14.4%) following treatment. A decrease in LDH levels exceeding 14.4% after treatment was associated with a significantly improved prognosis.



Further we analyzed the relationship between LDH dynamics and radiological tumor response. As shown in Figure 4, patients in the tumor reduction group experienced a significantly greater decline in LDH levels than those in the non-reduction group (P<.001; Figure 4). This indicates that a decrease in LDH levels is consistent with tumor shrinkage.

Figure 4. Comparison of LDH change rates between the tumor reduction group and the non-reduction group after 3 cycles of treatment. The tumor reduction group exhibited a significantly greater decrease in LDH levels compared with the non-reduction group (P < 0.05). (P<.001). “Relieve” indicates radiological evidence of reduction in the size of the primary tumor or affected lymph nodes after treatment, as assessed by the radiology department. “Not relieve” indicates no measurable reduction on follow-up imaging.



Multivariate and Univariate Analyses

In the univariate analysis performed using the Cox proportional hazards model (Table 2), it was observed that patients with decreased LDH levels (threshold=200 U/L) following treatment demonstrated a 41% reduction in the risk of death compared with those exhibiting increased LDH levels (HR 0.59, 95% CI 0.37 - 0.98; P=.04).

In addition to LDH changes, given that other clinical factors, including age, tumor length, tumor stage, tumor location, and the presence of comorbidities, were associated with PFS, we included these factors in our multivariate analysis (Table 2). The results of the multivariate analysis identified a reduction in LDH levels as an independent prognostic factor for improved outcomes (HR 0.59, 95% CI 0.36 - 0.96; P=.04; Table 2).

Compared with patients with elevated LDH levels, LDH downregulation was an independent prognostic factor for a 41% reduction in the risk of death (HR 0.59, 95% CI 0.36 - 0.96; P=.04).

XSL•FO
RenderX

Table . Univariate and multivariate analyses of factors influencing progression-free survival in all patients.

Variables	Univariate analysis		Multivariate analysis	
	HR[a] (95% CI)	*P* value	HR (95% CI)	*P* value
Gender (male vs female)	0.73 (0.37 - 1.43)	.35	—[b]	—
Age (years; <63 vs >63)	1.10 (0.67 - 1.80)	.70	—	—
Comorbidity (yes vs no)	1.08 (0.65 - 1.82)	.76	—	—
T stage (1 - 3 vs 4)	0.6 (0.37 - 0.97)	.04	0.67 (0.38 - 1.20)	.18
N stage (0 - 1 vs 2 - 3)	0.78 (0.33 - 1.80)	.56	—	—
Clinical stage (IVA[c] vs IVB[d])	0.60 (0.34 - 1.08)	.09	0.75 (0.37 - 1.48)	.40
Tumor length (<5 cm vs >5 cm)	0.98 (0.58-1.65)	.93	—	—
Tumor location				
Upper	Reference	—	—	—
Middle	0.69 (0.29 - 1.65)	.40	—	—
Lower	0.74 (0.30 - 1.82)	.51	—	—
LDH[e] change (down vs up)	0.59 (0.37 - 0.98)	.04	0.59 (0.36 - 0.96)	.04

[a]HR: hazard ratio.

[b]Not available.

[c]IVA: cancer stage IVA.

[d]IVB: cancer stage IVB.

[e]LDH: lactate dehydrogenase.

Discussion

Principal Findings

This study provides valuable insights into the prognostic significance of dynamic changes in LDH levels among patients with ESCC undergoing ICI therapy. The findings demonstrate that a reduction in LDH levels is associated with improved PFS. Specifically, patients with a moderate LDH decrease (0% to 14.4%) exhibited a median PFS of 21.7 months, compared to 10.8 months in those with increased LDH. Multivariate analysis further confirmed that an LDH decrease independently predicted a 41% reduction in mortality risk. These results underscore the potential utility of dynamic LDH monitoring as an accessible, cost-effective, and valuable prognostic tool for informing individualized treatment strategies in advanced ESCC.

Comparison to Previous Work

LDH, a key enzyme involved in cellular metabolism, facilitates the interconversion of lactate and pyruvate, which are critical to both aerobic and anaerobic metabolic pathways [13]. Elevated LDH levels have long been recognized as a prognostic marker in various malignancies, reflecting increased tumor burden and metabolic stress [14,15]. However, this study expands on existing knowledge by focusing on the dynamic changes in LDH levels during treatment rather than static measurements. This novel approach enables a more precise evaluation of treatment efficacy by capturing real-time fluctuations in tumor metabolism and therapeutic response. While traditional biomarkers for ICI therapy, such as PD-L1 expression, tumor mutational burden, and immune cell infiltration, provide valuable prognostic information, they often require complex, resource-intensive assessments [16-18]. In contrast, dynamic LDH monitoring offers a similarly predictive yet simpler and more cost-effective alternative. This is particularly advantageous in resource-limited settings, where such a practical tool can guide clinical decisions and optimize treatment strategies.

Role of LDH in Tumor Metabolism and Therapeutic Implications

Elevated LDH activity in cancer cells reflects reliance on anaerobic glycolysis, leading to increased lactate production [19]. Accumulation of lactate acidifies the tumor microenvironment, disrupts immune cell function, promotes tumor progression, and creates conditions favorable for immune evasion [20,21]. Dynamic monitoring of LDH levels provides crucial insights into these metabolic shifts, illustrating the impact of therapy on tumor biology. Specifically, elevated lactate levels impair T-cell function, reduce immune cell cytotoxicity, and promote an immunosuppressive phenotype in tumor-associated macrophages [22,23]. Conversely, a reduction in LDH levels, indicative of decreased lactate production, may alleviate these

XSL•FO
RenderX

immunosuppressive effects and enhance the efficacy of immune therapies. ICIs are specifically designed to counteract immune suppression in the tumor microenvironment [24]. Monitoring dynamic changes in LDH levels can therefore provide valuable insights into the evolving tumor state and its microenvironment during therapy. A decrease in LDH levels may signal effective tumor cell apoptosis or reduced glycolytic activity, both of which can relieve metabolic stress, enhance immune cell function, and ultimately improve therapeutic outcomes [25].

Clinical Utility

Dynamic monitoring of LDH facilitates personalized treatment strategies by allowing real-time assessment of therapeutic efficacy and metabolic changes. Patients with significant reductions in LDH levels are more likely to benefit from continued ICI therapy, while those with stable or increasing LDH levels may require alternative therapeutic approaches [26]. This approach allows clinicians to optimize treatment strategies tailored to individual patient responses, enhancing treatment efficacy and patient outcomes. However, integrating LDH monitoring into routine clinical practice necessitates the establishment of standardized testing protocols and data interpretation guidelines to ensure consistency and reliability across clinical settings [27]. Regular LDH assessments should be incorporated into treatment plans to provide ongoing feedback on therapeutic response. Furthermore, combining LDH data with other biomarkers could enhance the predictive accuracy of treatment outcomes, supporting individualized decision-making and more precise treatment strategies [28].

Limitations

Despite its potential, the retrospective design of this study introduces certain limitations, including selection bias and inconsistencies in data recording, which may impact the generalizability of the findings. Although efforts were made to minimize these biases through standardized procedures and data validation, the inherent limitations of retrospective analyses remain. In addition, the relatively small sample size may have influenced the statistical power and generalizability of the results. Moreover, variations in chemotherapy regimens could influence treatment outcomes, thereby affecting the conclusions of this study. In the future, additional samples should be collected to conduct more detailed subgroup analyses of various highly specific chemotherapy regimens, and prospective studies are needed to validate these findings and confirm the reliability of dynamic LDH monitoring as a prognostic marker. The applicability of these results to diverse populations and clinical settings also requires further exploration. Large-scale, multicenter studies are essential to evaluate the consistency and generalizability of LDH dynamics as a prognostic tool across diverse patient cohorts. Furthermore, variations across ethnicities, genders, and regions should be investigated to ensure broader applicability and relevance.

Future Directions

Future research should include prospective cohort studies and randomized controlled trials to validate the prognostic value of dynamic LDH changes and identify optimal monitoring strategies for enhancing treatment efficacy. Exploring the combination of LDH with other biomarkers, such as genomic features and immune cell subsets, could further refine prognostic accuracy. Integrating multiple biomarkers may offer a more comprehensive assessment of patient status, informing more precise treatment strategies. Moreover, studies should investigate how to tailor immune therapies based on dynamic LDH changes, including potential treatment modifications and individualized approaches. Evaluating the impact of these strategies on long-term patient outcomes through clinical trials will be critical. The integration of LDH monitoring with other clinical data could lead to the development of comprehensive treatment plans, ultimately improving patient prognosis and quality of life.

Conclusions

Dynamic changes in LDH levels, specifically a reduction within the 0% to 14.4% range, represent novel and significant predictors of PFS in patients with advanced ESCC. After 2 cycles of immunotherapy, patients with decreased LDH dynamic ratios demonstrated significantly better prognoses. These findings highlight the potential of LDH as a metabolic biomarker that reflects tumor biology and therapeutic response in real time. Future investigations are warranted to validate the utility of this specific LDH dynamic threshold as a prognostic tool. Such efforts could enhance personalized treatment strategies, ultimately improving clinical outcomes in patients with advanced ESCC.

Acknowledgments

This work was supported by the Beijing Life Oasis Public Service Center (grant CPHCF-2023-233).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

BZ was responsible for the study design, data analysis, and manuscript writing and revision. ZK contributed to experimental implementation, data collection, and manuscript writing and revision. YD, WJ, AF, RZ, JL, and YZ were responsible for experimental implementation and data collection. LS and ZY contributed to the study design and revised the manuscript. All authors reviewed and approved the final manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Morgan E, Soerjomataram I, Rumgay H, et al. The Global Landscape of Esophageal Squamous Cell Carcinoma and Esophageal Adenocarcinoma incidence and mortality in 2020 and projections to 2040: new estimates from GLOBOCAN 2020. Gastroenterology 2022 Sep;163(3):649-658. [doi: 10.1053/j.gastro.2022.05.054] [Medline: 35671803]
2. He S, Xu J, Liu X, Zhen Y. Advances and challenges in the treatment of esophageal cancer. Acta Pharm Sin B 2021 Nov;11(11):3379-3392. [doi: 10.1016/j.apsb.2021.03.008] [Medline: 34900524]
3. Chen MF, Chen PT, Lu MS, Lee CP, Chen WC. Survival benefit of surgery to patients with esophageal squamous cell carcinoma. Sci Rep 2017 Apr 6;7:46139. [doi: 10.1038/srep46139] [Medline: 28383075]
4. Kato K, Cho BC, Takahashi M, et al. Nivolumab versus chemotherapy in patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy (ATTRACTION-3): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol 2019 Nov;20(11):1506-1517. [doi: 10.1016/S1470-2045(19)30626-6] [Medline: 31582355]
5. Yap DWT, Leone AG, Wong NZH, et al. Effectiveness of immune checkpoint inhibitors in patients with advanced esophageal squamous cell carcinoma: a meta-analysis including low PD-L1 subgroups. JAMA Oncol 2023 Feb 1;9(2):215-224. [doi: 10.1001/jamaoncol.2022.5816] [Medline: 36480211]
6. Ždralević M, Brand A, Di Ianni L, et al. Double genetic disruption of lactate dehydrogenases A and B is required to ablate the "Warburg effect" restricting tumor growth to oxidative metabolism. J Biol Chem 2018 Oct 12;293(41):15947-15961. [doi: 10.1074/jbc.RA118.004180] [Medline: 30158244]
7. Xie H, Valera VA, Merino MJ, et al. LDH-A inhibition, a therapeutic strategy for treatment of hereditary leiomyomatosis and renal cell cancer. Mol Cancer Ther 2009 Mar;8(3):626-635. [doi: 10.1158/1535-7163.MCT-08-1049] [Medline: 19276158]
8. Jerby-Arnon L, Shah P, Cuoco MS, et al. A cancer cell program promotes T cell exclusion and resistance to checkpoint blockade. Cell 2018 Nov 1;175(4):984-997. [doi: 10.1016/j.cell.2018.09.006] [Medline: 30388455]
9. Petrelli F, Ardito R, Merelli B, et al. Prognostic and predictive role of elevated lactate dehydrogenase in patients with melanoma treated with immunotherapy and BRAF inhibitors: a systematic review and meta-analysis. Melanoma Res 2019 Feb;29(1):1-12. [doi: 10.1097/CMR.0000000000000520] [Medline: 30308577]
10. Apostolova P, Pearce EL. Lactic acid and lactate: revisiting the physiological roles in the tumor microenvironment. Trends Immunol 2022 Dec;43(12):969-977. [doi: 10.1016/j.it.2022.10.005] [Medline: 36319537]
11. Pouysségur J, Marchiq I, Parks SK, Durivault J, Ždralević M, Vucetic M. "Warburg effect" controls tumor growth, bacterial, viral infections and immunity - genetic deconstruction and therapeutic perspectives. Semin Cancer Biol 2022 Nov;86(Pt 2):334-346. [doi: 10.1016/j.semcancer.2022.07.004] [Medline: 35820598]
12. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009 Jan;45(2):228-247. [doi: 10.1016/j.ejca.2008.10.026] [Medline: 19097774]
13. Sharma D, Singh M, Rani R. Role of LDH in tumor glycolysis: regulation of LDHA by small molecules for cancer therapeutics. Semin Cancer Biol 2022 Dec;87:184-195. [doi: 10.1016/j.semcancer.2022.11.007] [Medline: 36371026]
14. Dercle L, Ammari S, Roblin E, et al. High serum LDH and liver metastases are the dominant predictors of primary cancer resistance to anti-PD(L)1 immunotherapy. Eur J Cancer 2022 Dec;177:80-93. [doi: 10.1016/j.ejca.2022.08.034] [Medline: 36332438]
15. Lv Y, Tang W, Xu Y, et al. Apolipoprotein L3 enhances CD8+ T cell antitumor immunity of colorectal cancer by promoting LDHA-mediated ferroptosis. Int J Biol Sci 2023;19(4):1284-1298. [doi: 10.7150/ijbs.74985] [Medline: 36923931]
16. Schoenfeld AJ, Rizvi H, Bandlamudi C, et al. Clinical and molecular correlates of PD-L1 expression in patients with lung adenocarcinomas. Ann Oncol 2020 May;31(5):599-608. [doi: 10.1016/j.annonc.2020.01.065] [Medline: 32178965]
17. Chan TA, Yarchoan M, Jaffee E, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. Ann Oncol 2019 Jan 1;30(1):44-56. [doi: 10.1093/annonc/mdy495] [Medline: 30395155]
18. Baharom F, Ramirez-Valdez RA, Khalilnezhad A, et al. Systemic vaccination induces CD8+ T cells and remodels the tumor microenvironment. Cell 2022 Nov 10;185(23):4317-4332. [doi: 10.1016/j.cell.2022.10.006] [Medline: 36302380]
19. Miholjcic TBS, Halse H, Bonvalet M, et al. Rationale for LDH-targeted cancer immunotherapy. Eur J Cancer 2023 Mar;181:166-178. [doi: 10.1016/j.ejca.2022.11.032] [Medline: 36657325]
20. Colegio OR, Chu NQ, Szabo AL, et al. Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. Nature New Biol 2014 Sep 25;513(7519):559-563. [doi: 10.1038/nature13490] [Medline: 25043024]
21. Chen L, Huang L, Gu Y, Cang W, Sun P, Xiang Y. Lactate-lactylation hands between metabolic reprogramming and immunosuppression. IJMS 2022;23(19):11943. [doi: 10.3390/ijms231911943]
22. Fischer K, Hoffmann P, Voelkl S, et al. Inhibitory effect of tumor cell-derived lactic acid on human T cells. Blood 2007 May 1;109(9):3812-3819. [doi: 10.1182/blood-2006-07-035972] [Medline: 17255361]

XSL•FO
RenderX

23. Zhang Q, Zhao Q, Li T, et al. Lactobacillus plantarum-derived indole-3-lactic acid ameliorates colorectal tumorigenesis via epigenetic regulation of CD8+ T cell immunity. Cell Metab 2023 Jun 6;35(6):943-960. [doi: 10.1016/j.cmet.2023.04.015] [Medline: 37192617]
24. Bagchi S, Yuan R, Engleman EG. Immune checkpoint inhibitors for the treatment of cancer: clinical impact and mechanisms of response and resistance. Annu Rev Pathol 2021 Jan 24;16:223-249. [doi: 10.1146/annurev-pathol-042020-042741] [Medline: 33197221]
25. Cascardo F, Anselmino N, Páez A, et al. HO-1 modulates aerobic glycolysis through LDH in prostate cancer cells. Antioxidants (Basel) 2021 Jun 16;10(6):966. [doi: 10.3390/antiox10060966] [Medline: 34208670]
26. van Not OJ, Wind TT, Ismail RK, et al. A survival tree of advanced melanoma patients with brain metastases treated with immune checkpoint inhibitors. Cancers (Basel) 2023 May 26;15(11):2922. [doi: 10.3390/cancers15112922] [Medline: 37296885]
27. Serganova I, Rizwan A, Ni X, et al. Metabolic imaging: a link between lactate dehydrogenase A, lactate, and tumor phenotype. Clin Cancer Res 2011 Oct 1;17(19):6250-6261. [doi: 10.1158/1078-0432.CCR-11-0397] [Medline: 21844011]
28. Claps G, Faouzi S, Quidville V, et al. The multiple roles of LDH in cancer. Nat Rev Clin Oncol 2022 Dec;19(12):749-762. [doi: 10.1038/s41571-022-00686-2] [Medline: 36207413]

Abbreviations

ESCC: esophageal squamous cell carcinoma
HR: hazard ratio
ICI: immune checkpoint inhibitor
LDH: lactate dehydrogenase
OS: overall survival
PD-1: programmed death-1
PD-L1: programmed death-ligand 1
PFS: progression-free survival

Edited by N Cahill; submitted 06.Mar.2025; peer-reviewed by W Xue, YS Wang; revised version received 02.Sep.2025; accepted 04.Sep.2025; published 14.Jan.2026.

Please cite as:
Zhang B, Kang Z, Ding Y, Jing W, Feng A, Zeng R, Li J, Zhao Y, Nie Y, Zhang W, Sun L, Yang Z
Prognostic Value of Dynamic Lactate Dehydrogenase Trends in Immunotherapy for Advanced Esophageal Squamous Cell Carcinoma: Retrospective Cohort Study
JMIR Cancer 2026;12:e73576
URL: https://cancer.jmir.org/2026/1/e73576
doi:10.2196/73576

© Baidong Zhang, Zhichao Kang, Yi Ding, Wang Jing, Alei Feng, Renya Zeng, Jianan Li, Yi Zhao, Yuanliu Nie, Wentao Zhang, Lu Sun, Zhe Yang. Originally published in JMIR Cancer (https://cancer.jmir.org), 14.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Explainable AI for Predicting Mortality Risk in Metastatic Cancer: Retrospective Cohort Study Using the Memorial Sloan Kettering-Metastatic Dataset

Polycarp Nalela[1]; Deepthi Rao[1]; Praveen Rao[1]

The University of Missouri, Columbia, MO, United States

Corresponding Author:
Praveen Rao
The University of Missouri
225 Naka Hall
Columbia, MO
United States
Phone: 1 9132317241
Email: praveen.rao@missouri.edu

Abstract

Background: Metastatic cancer remains one of the leading causes of cancer-related mortality worldwide. Yet, the prediction of survivability in this population remains limited by heterogeneous clinical presentations and high-dimensional molecular features. Advances in machine learning (ML) provide an opportunity to integrate diverse patient- and tumor-level factors into explainable predictive ML models. Leveraging large real-world datasets and modern ML techniques can enable improved risk stratification and precision oncology.

Objective: This study aimed to develop and interpret ML models for predicting overall survival in patients with metastatic cancer using the Memorial Sloan Kettering-Metastatic (MSK-MET) dataset and to identify key prognostic biomarkers through explainable artificial intelligence techniques.

Methods: We performed a retrospective analysis of the MSK-MET cohort, comprising 25,775 patients across 27 tumor types. After data cleaning and balancing, 20,338 patients were included. Overall survival was defined as deceased versus living at last follow-up. Five classifiers (extreme gradient boosting [XGBoost], logistic regression, random forest, decision tree, and naive Bayes) were trained using an 80/20 stratified split and optimized via grid search with 5-fold cross-validation. Model performance was assessed using accuracy, area under the curve (AUC), precision, recall, and F_1-score. Model explainability was achieved using Shapley additive explanations (SHAP). Survival analyses included Kaplan-Meier estimates, Cox proportional hazards models, and an XGBoost-Cox model for time-to-event prediction. The positive predictive value and negative predictive value were calculated at the Youden index–optimal threshold.

Results: XGBoost achieved the highest performance (accuracy=0.74; AUC=0.82), outperforming other classifiers. In survival analyses, the XGBoost-Cox model with a concordance index (C-index) of 0.70 exceeded the traditional Cox model (C-index=0.66). SHAP analysis and Cox models consistently identified metastatic site count, tumor mutational burden, fraction of genome altered, and the presence of distant liver and bone metastases as among the strongest prognostic factors, a pattern that held at both the pan-cancer level and recurrently across cancer-specific models. At the cancer-specific level, performance varied; prostate cancer achieved the highest predictive accuracy (AUC=0.88), while pancreatic cancer was notably more challenging (AUC=0.68). Kaplan-Meier analyses demonstrated marked survival separation between patients with and without metastases (80-month survival: approximately 0.80 vs 0.30). At the Youden-optimal threshold, positive predictive value and negative predictive value were approximately 70% and 80%, respectively, supporting clinical use for risk stratification.

Conclusions: Explainable ML models, particularly XGBoost combined with SHAP, can strongly predict survivability in metastatic cancers while highlighting clinically meaningful features. These findings support the use of ML-based tools for patient counseling, treatment planning, and integration into precision oncology workflows. Future work should include external validation on independent cohorts, integration with electronic health records via Fast Healthcare Interoperability Resources–based dashboards, and prospective clinician-in-the-loop evaluation to assess real-world use.

(JMIR Cancer 2026;12:e74196) doi:10.2196/74196

XSL•FO
RenderX

KEYWORDS

explainable artificial intelligence; machine learning in oncology; metastatic cancer; survivability prediction

Introduction

Cancer remains one of the foremost global health challenges, with 611,720 deaths projected in the United States alone for 2024, and metastatic disease accounting for the overwhelming majority of these fatalities [1]. Metastasis is particularly vexing because disseminated tumor cells frequently acquire distinct genomic and phenotypic profiles that render them resistant to standard therapies. Contemporary oncology research, therefore, acknowledges intra- and intertumor heterogeneity as fundamental obstacles to curative treatment, therapeutic resistance, and accurate prognosis [2,3]. Reliable, early-stage prediction of patient survivability is pivotal not merely for counseling patients but also for tailoring aggressive interventions, prioritizing scarce health care resources, and designing adaptive clinical trials aimed at improving long-term outcomes.

Historically, clinicians have relied on the tumor-node-metastasis staging system, Kaplan-Meier life tables, and the Cox proportional hazards (CPH) regression model to stratify survival risk. Although statistically rigorous, these techniques impose proportional hazards and linearity assumptions that rarely hold across the dynamic, nonlinear biology of metastatic cancers. When violated, CPH models can yield biased hazard ratios, suffer from time-dependent covariate effects, and perform poorly on censored or highly imbalanced datasets [4].

The emergence of artificial intelligence (AI) and machine learning (ML) has enabled the development of sophisticated models that can uncover previously hidden patterns in heterogeneous clinical and multiomics datasets, thereby providing new insights into cancer biology, diagnosis, prognosis, and treatment outcomes. Over the past decade, ML models have repeatedly outperformed traditional statistical approaches. A deep-learning algorithm developed by Esteva et al [5] achieved a sensitivity of 97% and a specificity of 78% in classifying skin lesions as benign or malignant, while Liu et al [6] reported an area under the receiver operating characteristic curve (AUROC) of 0.94 for lung cancer risk prediction using computed tomography images. Similar gains have been demonstrated for breast cancer survival [7], lymph node metastasis [8], colorectal and soft tissue sarcoma outcomes [9], lung cancer survival [10], and prostate cancer prognosis [11,12].

Among contemporary ML pipelines, gradient-boosted decision trees, particularly extreme gradient boosting (XGBoost) [13], have emerged as a workhorse because they natively handle missing data, nonlinear feature interactions, and mixed data types. Recent examples include a non–small cell lung cancer microwave-ablation study where XGBoost achieved an area under the curve (AUC) of 0.89 [14]; a 2025 *Scientific Reports* analysis integrating survival models for breast cancer recurrence (concordance index [C-index]=0.82) [15]; a large colorectal cancer cohort where boosted trees yielded the highest 5-year survival accuracy [16]; and a thyroid cancer study that constructed a 10-year overall survival nomogram using Surveillance, Epidemiology, and End Results data [17].

Biomarkers are critical for early detection, diagnosis, prognosis, and monitoring. Traditional biomarker-discovery approaches often suffer from low sensitivity, limited reproducibility, and dependence on prior biological hypotheses. ML circumvents many of these limitations by integrating diverse data types and identifying complex nonlinear relationships. Algorithms such as random forests [18], support vector machines [19], and neural networks [20] have successfully identified biomarkers from gene expression [21-23], microRNA expression [24-27], DNA methylation [28,29], and imaging modalities [30-33].

Despite this progress, significant gaps persist at the intersection of scale, interpretability, and clinical use. First, while large datasets like the Memorial Sloan Kettering-Metastatic (MSK-MET) cohort [34] provide unprecedented scale, their analysis has largely relied on traditional statistics, failing to harness state-of-the-art ML for predictive modeling. Second, most ML survival-prediction studies focus on single tumor types, use modest sample sizes, or omit high-dimensional genomic features, limiting their generalizability to the pan-cancer reality of metastatic disease. Third, and most critically, interpretability remains a bottleneck; oncologists are understandably reluctant to incorporate opaque “black box” risk scores into clinical workflows. Explainable AI methods like Shapley additive explanations (SHAP) provide a mechanism for transparency. For instance, SHAP has revealed previously unknown drivers of prostate cancer mortality [35]. Although explainable AI frameworks offer a solution, large-scale, pan-cancer implementations that jointly optimize predictive performance and model explainability remain scarce. Consequently, the field lacks an interpretable, cross-tumor framework capable of ranking metastasis-specific risk factors at a scale that reflects real-world heterogeneity.

To address these critical gaps, we designed a methodological framework that moves beyond standard, single-model architectures. We leverage the MSK-MET dataset, a pan-cancer cohort of genomic and clinical data from 25,775 patients spanning 27 tumor types, as an ideal test bed for this purpose due to its scale and diversity. Our approach is conceived as a hierarchical and explainable pipeline specifically to tackle the challenges of data heterogeneity, the need for clinical trust, and biological discovery. It integrates rigorous pan-cancer benchmarking with targeted, tumor-specific submodels and crucially unifies ML classification with traditional survival analysis. This ensures that the predictive performance of our models is directly coupled with transparent, clinically actionable insights.

Driven by this methodology, our work addresses persistent gaps at the intersection of scale, interpretability, and clinical use in metastatic cancer prognosis. Leveraging a large pan-cancer cohort that captures real-world heterogeneity, we focus on whether explainable ML approaches can yield clinically useful

XSL•FO
RenderX

survivability predictions while providing transparent, biologically and clinically coherent insights across tumor types.

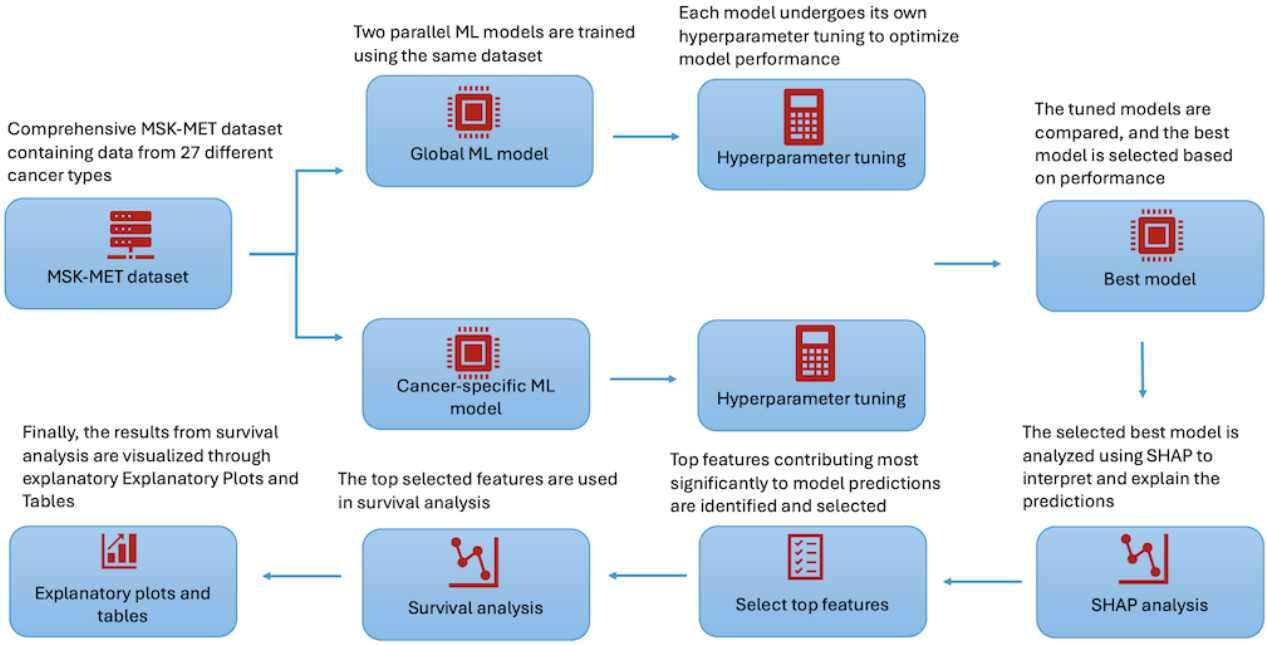
The primary aim of this study is to develop and validate an interpretable ML framework for predicting overall survival in patients with metastatic cancer. We hypothesize that (1) ML models will achieve clinically useful discrimination and calibration for survivability prediction, (2) explainable AI techniques will identify a core set of prognostic biomarkers that are consistently important across diverse tumor types, and (3) integrating ML predictions with established survival analysis techniques will yield a transparent and clinically actionable tool for risk stratification.

Methods

Overview

Figure 1 illustrates the steps followed to predict cancer survivability using explainable AI. Raw data were used to initially train the ML models, followed by SHAP analysis. Top features identified by SHAP were then further used in the survival analysis. The subsequent steps below detail how this was implemented. All analysis and visualization were carried out in Python (version 3.12; Python Software Foundation) with relevant packages and libraries such as *pandas*, *numpy*, *scikit-learn*, *shap*, etc.

Figure 1. Overview of the explainable machine learning (ML) pipeline for metastatic cancer survivability prediction. The Memorial Sloan Kettering-Metastatic (MSK-MET) cohort is preprocessed and balanced, then split into stratified training and test sets before training and tuning 5 candidate classifiers (extreme gradient boosting [XGBoost], random forest, logistic regression, decision tree, and naive Bayes). The best-performing XGBoost model is subsequently interrogated with Shapley additive explanations (SHAP) to identify key prognostic clinical and genomic features, which are then carried forward into downstream survival analyses (Kaplan-Meier curves, Cox proportional hazards models, and XGBoost-Cox) to generate time-to-event estimates and clinically interpretable risk stratification.



Data Preprocessing

We first performed a thorough exploratory data analysis on MSK-MET that contained information from 25,775 patients with cancer. Our exploratory data analysis process began with a comprehensive analysis of the dataset, including the distribution of different cancer types, stages, and other relevant features. This helped us gain a deeper understanding of the underlying patterns and structures, which informed subsequent preprocessing steps. We preprocessed the input dataset and dropped columns (such as patient ID) and rows having large proportions of missing data. The target variable had 2 classes, that is, living (positive class, coded as 1) and deceased (negative class, coded as 0). In the final stage of preprocessing, the data were balanced using the target classes. Using resampling, we down-sampled the minority class (0). The final set contained 20,338 patients (10,169 living and 10,169 deceased) with 39 variables for each patient. In total, there were 27 cancer types (Table 1). The overall survival status was the target variable for prediction. Categorical variables were encoded using label encoding, and features were scaled using minimum-maximum scaling to ensure that variables with larger magnitudes did not unduly influence model outcomes. The resulting preprocessed data were then split into training and testing sets for further analysis.

XSL•FO
RenderX

Table 1. Distribution of primary cancer types in the Memorial Sloan Kettering-Metastatic cohort. Frequency counts are reported for all cancer types represented in the dataset; the 5 largest groups (non–small cell lung, colorectal, breast, pancreatic, and prostate cancer) provided the primary strata for cancer-specific extreme gradient boosting models and downstream survival analyses.

Cancer type	Frequency count, n
Non–small cell lung cancer	3790
Colorectal cancer	2696
Breast cancer	2043
Pancreatic cancer	1738
Prostate cancer	1596
Endometrial cancer	988
Ovarian cancer	923
Melanoma	882
Bladder cancer	870
Hepatobiliary cancer	790
Esophagogastric cancer	738
Soft tissue sarcoma	420
Head and neck cancer	362
Thyroid cancer	319
Renal cell carcinoma	318
Gastrointestinal stromal tumor	286
Small cell lung cancer	277
Germ cell tumor	241
Mesothelioma	219
Appendiceal cancer	160
Uterine sarcoma	133
Salivary gland cancer	123
Gastrointestinal neuroendocrine tumor	115
Skin cancer (nonmelanoma)	87
Cervical cancer	80
Small bowel cancer	76
Anal cancer	68

Stratified Random Sampling of Training and Testing Sets

We used a stratified random sampling approach to create the training and test sets. First, we randomized the complete dataset to eliminate any inherent order or sequence. Then, we implemented stratification to ensure that the distribution of specific cancer types or stages in our training and testing sets mirrored that of the entire dataset. This is paramount to avoid potential biases and to ensure that our models have a representative sample of the different cancer types and stages present in the entire dataset. Following stratification, we allocated 80% of the data (16,270 patient records) to the training set while reserving the remaining 20% (4068 patient records) for the test set. This approach provides a robust foundation for model development and validation, ensuring both broad and deep representation of the dataset in our training and testing phases.

Selection and Screening of ML Models for Cancer Survival Prediction

This study used 5 ML algorithms—XGBoost, naive Bayes, decision tree, logistic regression, and random forest—to predict cancer survival rates using the MSK-MET dataset. XGBoost was selected for its efficiency in handling sparse data and combining models to improve accuracy through ensemble learning. Naive Bayes, a simple classifier applying Bayes theorem, was chosen for its efficiency in high-dimensional datasets. The decision tree, known for its easy visualization and handling of nonlinear relationships, was included for its interpretability. Logistic regression was used for binary classification, predicting survival probabilities, while random forest, an ensemble method using multiple decision trees, was

XSL•FO
RenderX

chosen for its accuracy and control over overfitting in large datasets.

Hyperparameter Optimization Via Grid Search and Model Training

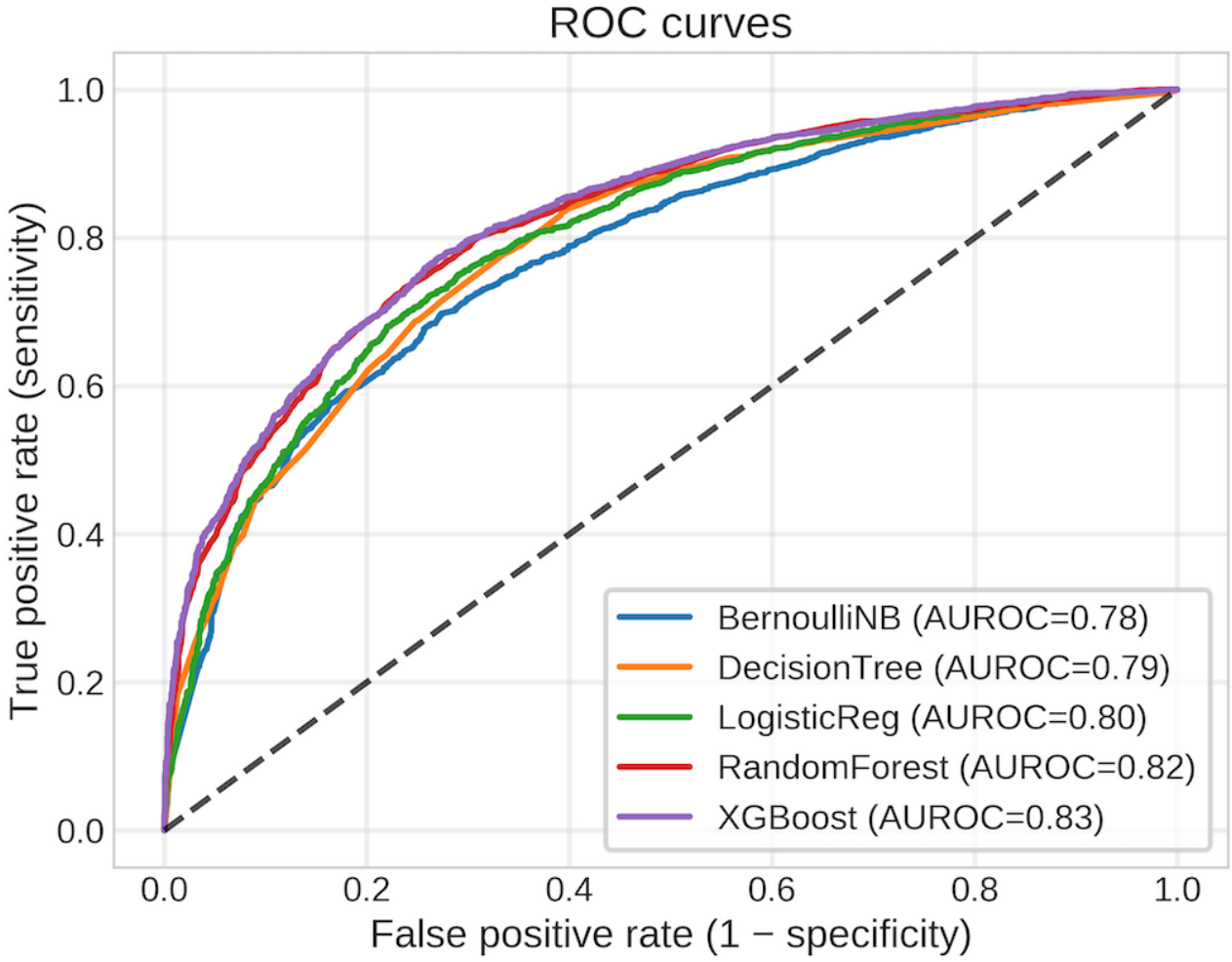
Grid search with hyperparameter tuning was applied to all 5 ML models. For XGBoost, parameters "n_estimators" (50-1000), "max_depth" (1-10), and "learning_rate" (0.01-0.3) were adjusted to optimize the number of trees, tree depth, and learning speed. Naive Bayes was tuned by varying "alpha" (0.01-10.0), "binarize" (0.0, 0.5, 1.0), and "fit_prior" (True/False). The decision tree's grid search adjusted "max_depth" (None-10), "min_samples_split" (2, 5, 10), "min_samples_leaf" (1, 2, 4), and "criterion" ("gini" or "entropy"). Logistic regression was optimized with "C" (0.001-1000), "penalty" ("l1," "l2," "elasticnet," "none"), and "solver" ("newton-cg," "lbfgs," "liblinear," "sag," "saga"). Random forest explored "n_estimators" (50, 100, 200) and "max_features" ("auto," "sqrt").

Using 5-fold cross-validation on the MSK-MET dataset, we trained the 5 classifiers to identify a robust, interpretable predictor and to derive a cohort-wide view of metastatic patterns.

Model Evaluation

After training and testing the ML models on the MSK-MET dataset, we assessed their performance using 2 key metrics: the classification report (Table S2 in Multimedia Appendix 1) and AUROC (Figure 2). The AUROC measures the model's ability to distinguish between classes, with higher AUC indicating better prediction. A score of 1 represents perfect predictions, 0.5 indicates random guessing, and below 0.5 suggests worse than random predictions. These metrics provide a comprehensive evaluation, ensuring the models not only predict accurately but also effectively identify positive cancer cases. This approach helps in selecting the best model for predicting cancer survival, balancing the need to detect true cases while minimizing false diagnoses.

Figure 2. Receiver operating characteristic (ROC) curves for the 5 tuned classifiers on the held-out Memorial Sloan Kettering-Metastatic test set. Each curve shows the trade-off between sensitivity and 1 – specificity across decision thresholds, with corresponding area under the curve values summarized in Table 2. The extreme gradient boosting (XGBoost) model achieves the steepest ROC trajectory and highest area under the receiver operating characteristic curve (AUROC=0.82), indicating the strongest discrimination between surviving and deceased patients, while random forest and logistic regression (LogisticReg) form an intermediate tier, and decision tree and Bernoulli naive Bayes (BernoulliNB) exhibit comparatively weaker performance.



To characterize operating-point behavior, we computed threshold-dependent trade-offs of true positive (TP), false positive (FP), true negative (TN), and false negative (FN); sensitivity, specificity, precision, negative predictive value (NPV), F_1-score, balanced accuracy, Matthews correlation coefficient (MCC), Cohen κ, and accuracy at 3 decision

XSL•FO
RenderX

thresholds: a fixed 0.50, the Youden J optimum, and a clinically constrained point targeting sensitivity ≥0.85. Scalar metrics were accompanied by 95% CI obtained via a 1000-sample stratified bootstrap.

To obtain unbiased estimates and avoid tuning leakage, we used nested repeated cross-validation on the training data. The outer loop applied RepeatedStratifiedKFold (5 folds × 10 repeats); within each outer training split, an inner 5-fold GridSearchCV tuned hyperparameters per model. Inner-tuned models were then scored on their corresponding outer validation splits, yielding a matrix of outer cross-validation (CV) results. The model with the highest mean outer-CV performance was refit on the full training set using the selected hyperparameters and evaluated once on the held-out test set (reporting accuracy, AUROC, and a confusion matrix).

For statistical comparison across the 5 models, we ran a Friedman omnibus test on the outer-CV score matrix. As a prespecified confirmatory analysis, we tested the directional hypothesis that random forest ≥ each baseline using 1-sided Wilcoxon signed-rank tests with Holm correction; we also reported effect sizes (mean and median paired differences) with 95% paired bootstrap CI. For rank-based visualization and exploratory post hoc inference, we computed average ranks with a Nemenyi critical-difference diagram and performed 1-sided Wilcoxon tests comparing the top-ranked model to the remaining models (Holm-adjusted). Where paired per-example test predictions were available, we used the McNemar test to compare error patterns between 2 models.

Cancer Survival Prediction With XGBoost

XGBoost was therefore chosen for the rest of the analyses, involving both pan-cancer and cancer-specific models, metastatic cancer survival prediction, and survival analysis. To enhance transparency and clinical interpretability, we adopted a 2-tier XGBoost design. First, a unified XGBoost model was trained without the “Cancer Type” variable, leveraging clinical and demographic features to capture signals that generalize across diseases. Second, we trained cancer-specific XGBoost models for the 5 largest cancer groups identified in the cohort to capture within-disease interactions that can be diluted in a single global model. This combination provides both a cross-cancer perspective and disease-tailored insights that reflect differences in metastatic behavior and treatment context. In the final evaluation for both global and cancer-specific analysis, we focused only on measuring accuracy and AUC score.

Model Interpretation and Explanation

To enhance the understandability and transparency of our predictions, we used XGBoost and SHAP for model explainability. SHAP, based on game theory, provides a detailed and consistent measure of feature importance by computing each feature’s contribution to the prediction. SHA*P* values represent a feature’s responsibility for a change in the model output, ensuring local accuracy, missingness, and consistency. This method quantifies the impact of each feature on predictions and explains how the presence or absence of a feature affects the outcome. In our SHAP analysis, survival is the positive class. Beeswarm plots are particularly useful for visualizing SHA*P* values, showing features’ influence and variability in a nuanced manner.

Survival Analysis

Following the training of the XGBoost ML model and SHAP analysis, the most important features, such as metastatic site count, tumor mutational burden (TMB), and specific organ metastases (eg, liver, bone, and lung) influencing patient survival, were identified and used in the survival analysis. The primary goal was to examine the duration from cancer diagnosis to patient death, assessing how clinical and genomic variables impact survival times. We used Kaplan-Meier survival analysis, CPH modeling, log-rank tests for comparing survival distributions, and XGBoost survival analysis to deepen our understanding of patient outcomes. All of the original data was used in the survival analysis.

Kaplan-Meier Survival Analysis

The Kaplan-Meier estimator was used to evaluate survival probabilities over time across different patient subgroups. Patients were stratified based on key features identified from SHAP analysis. Survival curves were compared using the log-rank test to assess statistically significant differences between groups. A *P* value of <.05 was considered statistically significant. The analysis was first performed on all the data to assess survival of patients with metastatic versus nonmetastatic disease and then on subgroups of the top 5 cancer types including non–small cell lung cancer, colorectal cancer, breast cancer, pancreatic cancer, and prostate cancer.

CPH Model

The CPH model was applied to assess the influence of multiple covariates on patient survival while controlling for potential confounders. Key covariates included metastatic site count, fraction of genome altered, TMB, and distant metastases in specific organs. The proportional hazards assumption was evaluated using Schoenfeld residuals, and any violations were addressed through stratification or inclusion of time-varying covariates. Hazard ratios with corresponding 95% CI were reported to quantify risk associations.

Log-Rank Test

To further compare survival distributions between different patient cohorts, the log-rank test was applied. This test was used to determine whether survival differences observed between patient subgroups (eg, metastatic vs nonmetastatic) were statistically significant. The resulting *P* values guided the identification of meaningful clinical predictors. Furthermore, a plot was generated for the Kaplan-Meier survival curves with the overall survival (months) on the x-axis and survival probability on the y-axis.

XGBoost Survival Analysis

To capture complex, nonlinear relationships and interactions among variables, XGBoost survival analysis was implemented. This adaptation of XGBoost used a Cox-based loss function to accommodate censored survival data. Hyperparameter tuning was conducted using grid search, optimizing parameters such as “n_estimators,” “max_depth,” and “learning_rate.” The model’s C-index was used to evaluate predictive performance.

XSL•FO
RenderX

SHA*P* values were also applied to the survival model to interpret feature importance and explore individual risk predictions.

Ethical Considerations

This study analyzed secondary, noninterventional data from the MSK-MET cohort obtained via a publicly accessible repository (cBioPortal for Cancer Genomics). All records used for modeling and statistical analysis were anonymous and deidentified prior to access; no direct identifiers (eg, names, street addresses, full dates of birth, medical record numbers) or indirect reidentification keys were available to the research team. Because only deidentified data were used and no contact with human participants occurred, the work was considered non–human participants research and did not require informed consent or additional institutional review.

Data handling procedures followed best-practice privacy safeguards. Working datasets were stored on access-controlled systems. We did not attempt any record linkage or reidentification. To support responsible AI, model development incorporated transparent methods (eg, SHAP explanations) and prespecified subgroup evaluations to screen for potential performance disparities. All code and evaluation protocols are shared to enable reproducibility without exposing any protected information.

Results

Model Performance

The evaluation of the 5 distinct models on the MSK-MET dataset yielded a spectrum of performances. The overall classification performance for the 5 classifiers is summarized in Table 2. Receiver operating characteristic (ROC) curves are shown in Figure 2, and precision-recall curves are shown in Figure S2A, calibration in Figure S2B, decision-curve analysis in Figure S2C, and threshold-dependent metrics in Figure S2D in Multimedia Appendix 1.

Table 2. Discrimination and calibration of tuned classifiers on the same test set. Reported are area under the curve (AUC), area under the precision-recall curve (AUPRC), and Brier score, each with 95% CIs; the outcome prevalence was 0.50. Extreme gradient boosting (XGBoost; lr=0.01; depth=5; n=500) showed the strongest overall performance, with the highest AUC and AUPRC and the lowest Brier score, followed by random forest, while logistic regression, decision tree, and naive Bayes exhibited progressively lower discrimination and less favorable calibration.

Model	AUC (95% CI)	AUPRC (95% CI)	Brier score (95% CI)
XGBoost (lr=0.01; depth=5; n=500)	0.82 (0.81-0.84)	0.83 (0.81-0.85)	0.17 (0.16-0.18)
Random forest (n_estimators=200; max_features=auto)	0.80 (0.78-0.81)	0.81 (0.79-0.82)	0.18 (0.18-0.19)
Logistic regression (C=10; l2; liblinear)	0.79 (0.78-0.81)	0.79 (0.77-0.81)	0.19 (0.18-0.19)
Decision tree (gini; depth=8; min_leaf=1; min_split=10)	0.78 (0.77-0.80)	0.78 (0.76-0.80)	0.19 (0.18-0.20)
Naive Bayes (α=10.0; binarize=0.5; fit_prior=False)	0.78 (0.77-0.80)	0.77 (0.75-0.79)	0.21 (0.20-0.22)

XGBoost demonstrated the strongest discrimination and probability quality. It achieved an AUC of 0.82 (95% bootstrap CI 0.81-0.84) and the highest area under the precision-recall curve (AUPRC) at approximately 0.83 against a baseline precision equal to the prevalence (0.50; Table 2 and Table S2 in Multimedia Appendix 1). In accuracy terms, XGBoost reached 0.74 (3010/4068) and, for our summary counts, corresponded to 3335 out of 4068 test patients correctly stratified at the chosen operating point. Random forest (AUC=0.80; AUPRC=0.81) and logistic regression (AUC=0.79; AUPRC=0.79; accuracy=0.72; 2929/4068) formed a consistent middle tier, while decision tree and Bernoulli naive Bayes trailed slightly (both AUC=0.78; accuracy=0.72; 2929/4068; correctly stratified=3173/4068 for AUC-aligned counts; see Table 3 and Table S3 in Multimedia Appendix 1 for detailed counts and metrics). These rankings were concordant across ROC and precision-recall analyses and remained stable across threshold-dependent operating points (fixed threshold=0.50, Youden J optimum, and a clinically constrained sensitivity ≥0.85; Figure 2; Figures S2A and S2D and Tables S2 and S3 in Multimedia Appendix 1).

Calibration analyses reinforced this ordering. XGBoost produced the lowest Brier score (approximately 0.17; Table 2; Table S2 and Figure S2B in Multimedia Appendix 1) and a reliability curve that closely tracked the 45° line, with logistic regression comparably well-calibrated, whereas Bernoulli naive Bayes deviated most at the extremes (Table 2; Table S2 and Figure S2B in Multimedia Appendix 1).

XSL•FO
RenderX

Table 3. Compact comparison of 5 classifiers across 3 operating points, that is, default probability cutoff (0.50), Youden J (threshold maximizing sensitivity + specificity – 1), and a high-sensitivity setting (constrained to sensitivity ≥0.85). For each model-threshold pair, we report sensitivity, specificity, Matthews correlation coefficient (MCC), and accuracy; thresholds are applied to the predicted positive-class probability. This summary emphasizes decision-relevant trade-offs: balanced performance at Youden J and the specificity cost of prioritizing high sensitivity. In this cohort, extreme gradient boosting (XGBoost) yields the strongest balanced performance (highest MCC and accuracy) at Youden J while retaining the best specificity among the high-sensitivity operating points.

Model and operating point	Threshold	Sensitivity	Specificity	MCC	Accuracy
Naive Bayes					
Probability cutoff 0.5	0.50	0.66	0.78	0.44	0.72
Youden J	0.50	0.65	0.78	0.44	0.72
Sensitivity ≥0.85	0.15	0.85	0.52	0.39	0.69
Decision tree					
Probability cutoff 0.5	0.50	0.72	0.71	0.43	0.72
Youden J	0.56	0.64	0.80	0.45	0.72
Sensitivity ≥0.85	0.32	0.86	0.49	0.37	0.67
Logistic regression					
Probability cutoff 0.5	0.50	0.73	0.73	0.45	0.73
Youden J	0.53	0.69	0.77	0.46	0.73
Sensitivity ≥0.85	0.38	0.85	0.54	0.41	0.69
Random forest					
Probability cutoff 0.5	0.50	0.69	0.76	0.45	0.72
Youden J	0.49	0.70	0.75	0.46	0.73
Sensitivity ≥0.85	0.32	0.85	0.54	0.41	0.69
XGBoost					
Probability cutoff 0.5	0.50	0.72	0.78	0.50	0.75
Youden J	0.50	0.72	0.78	0.50	0.75
Sensitivity ≥0.85	0.33	0.85	0.58	0.45	0.72

Threshold-dependent performance, summarized from confusion matrix–derived metrics at 3 operating points, made the error trade-offs explicit (Table 2; Table S3 and Figure S2D in Multimedia Appendix 1). At a fixed 0.50 threshold essentially indistinguishable from the Youden J optimum on this balanced test set, XGBoost balanced sensitivity and specificity most effectively (sensitivity=0.72; specificity=0.78; positive predictive value (PPV)=0.77; NPV=0.74; F_1-score=0.74; MCC=0.50; accuracy=0.75). Under a clinically constrained operating point prioritizing case-finding (sensitivity ≥0.85), XGBoost required a threshold of approximately 0.33 and achieved a sensitivity of 0.85 (specificity=0.58; PPV=0.67; NPV=0.80; F_1-score=0.75; MCC=0.45; accuracy=0.72). The corresponding counts at this setting were TP=1731, FP=846, TN=1188, and FN=303 (n=4068). Comparator models met the same sensitivity target with lower specificity and weaker composite indices; for example, random forest and logistic regression both settled near a specificity of 0.54 with MCC=0.41, while decision tree and Bernoulli naive Bayes lost additional specificity and MCC. These results indicate that, when sensitivity is held high, XGBoost preserves more TNs and maintains stronger global agreement (MCC, balanced accuracy).

Decision-curve analysis supported the same ordering of clinical use across a broad range of threshold probabilities (approximately 0.15-0.70), with XGBoost yielding the highest net benefit, random forest next, and logistic regression close behind (Figure S2C and Table S2 in Multimedia Appendix 1). Together with the calibration findings, these analyses suggest that XGBoost not only separates cases from controls most effectively but also produces usable risk estimates for thresholding and shared decision-making.

Furthermore, when the global XGBoost model was compared to cancer-specific XGBoost models for the top 5 cancer types (non–small cell lung, colorectal, breast, pancreatic, and prostate cancer; Tables 4 and 5), the prostate cancer model emerged as the most accurate, with an AUC of 0.88 and an accuracy of 0.84. (268/319 test patients) and an AUC of 0.88 (281/319 test patients correctly stratified). Meanwhile, pancreatic cancer posted a lower AUC of 0.68 (236/348 test patients correctly stratified), reflecting greater challenges in classification for that subgroup.

XSL•FO
RenderX

Table 4. Condensed Cox proportional hazards model for overall survival in the Memorial Sloan Kettering-Metastatic cohort. Hazard ratios (HRs), 95% CIs, and *P* values are reported for the most influential covariates identified in the global model, showing that metastatic status, higher metastatic site count, increased tumor mutational burden (TMB), greater fraction of genome altered, and distant metastases to liver, bone, and lung are all associated with significantly elevated mortality risk, whereas primary sample type is modestly protective.

Risk factor	HR (95% CI)	*P* value
Patient with metastatic disease	2.18 (1.97-2.42)	<.001
Metastatic site count	1.03 (1.02-1.04)	<.001
TMB (nonsynonymous)	1.00 (0.99-1.00)	<.001
Fraction of genome altered	1.32 (1.19-1.46)	<.001
Sample type (primary=ref)	0.87 (0.83-0.90)	<.001
Distant metastasis: liver	1.81 (1.73-1.90)	<.001
Distant metastasis: bone	1.43 (1.37-1.50)	<.001
Distant metastasis: lung	1.16 (1.11-1.22)	<.001

Table 5. Comparison of classification performance between the unified or global extreme gradient boosting (XGBoost) model and cancer-specific XGBoost models. Overall accuracy and area under the curve (AUC) are reported for each classifier, showing that while the global model achieves strong baseline discrimination (accuracy=0.74; AUC=0.82), several cancer-specific models, particularly prostate and breast cancer, attain even higher AUCs (0.88 and 0.85, respectively), whereas the pancreatic cancer model lags behind, reflecting underlying heterogeneity in predictability across tumor types.

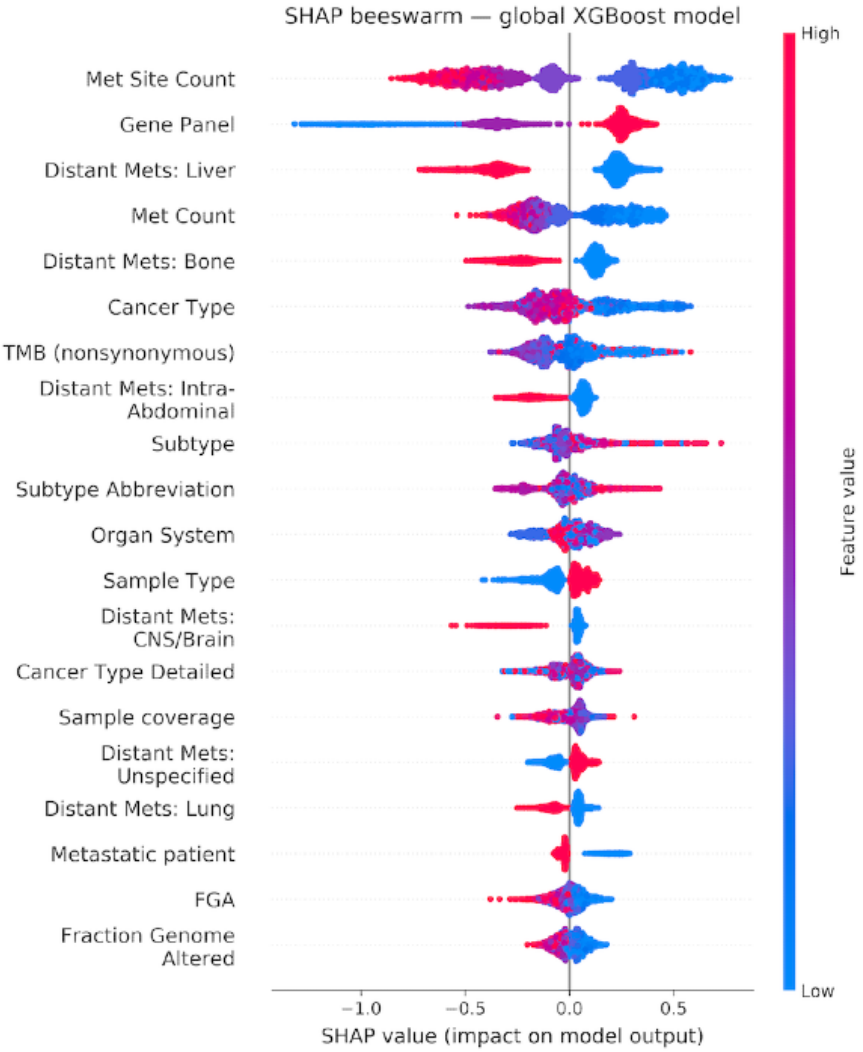
Classifier	Accuracy score	AUC score
Unified or global model	0.74	0.82
Non–small cell lung model	0.71	0.79
Colorectal cancer model	0.73	0.81
Breast cancer model	0.76	0.85
Pancreatic cancer model	0.72	0.68
Prostate cancer model	0.84	0.88

Model Explainability

The SHAP analysis of the global XGBoost model established a clear hierarchy of feature importance, identifying metastatic site count, TMB, fraction of genome altered, and distant metastases to the liver and bone as the predominant prognostic factors (Figure 3 and Table S6 in Multimedia Appendix 1). The beeswarm plots for these top features demonstrated a robust predictive power, with high and low feature values cleanly separating along the SHA*P* value axis, indicating a consistent and strong directional impact on model output (Figures S3A-S3E in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 3. Shapley additive explanations (SHAP) beeswarm plot for the global extreme gradient boosting (XGBoost) mortality classifier in the Memorial Sloan Kettering-Metastatic cohort. Features are ordered by mean absolute SHAP value, highlighting metastatic (Met) site count, tumor mutational burden (TMB), fraction of genome altered (FGA), and distant metastases to the liver and bone as the strongest drivers of model predictions. Each point represents a patient, with horizontal position indicating the direction and magnitude of impact on predicted mortality risk and color denoting low (blue) to high (red) feature values, illustrating how extreme values systematically shift risk estimates. Central nervous system is denoted by CNS.

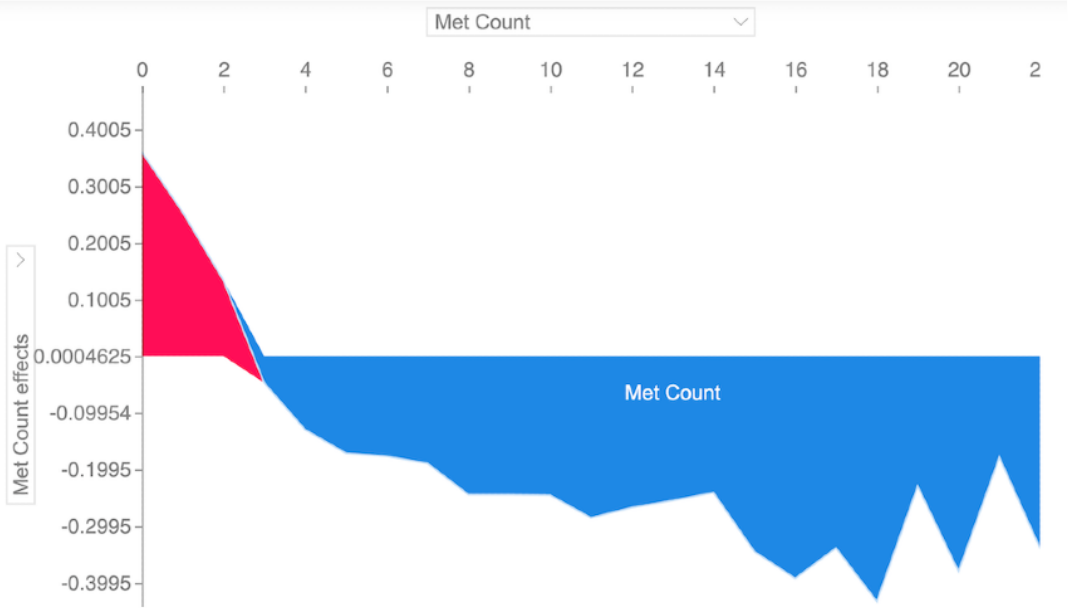


This analysis was extended to cancer-specific models, revealing a critical dual perspective, that is, the core features identified in the global model recurrently ranked among the most important across individual cancer types, while disease-specific features also emerged. For instance (Table S6 and Figures S3A-S3E in Multimedia Appendix 1), the global top features like metastatic burden and liver metastases remained highly influential in specific models such as colorectal and prostate cancer. Concurrently, the models identified context-specific predictors, such as distant metastasis in the lung for non–small cell lung cancer, sample type for breast cancer, and distant metastasis in the male genital for prostate cancer (Figure 3; Figures S3D and S3E in Multimedia Appendix 1). This underscores that while a common set of pan-cancer drivers exists, the models successfully capture nuanced, disease-specific metastatic behaviors.

The consistent importance of features like metastatic site count was further validated by the SHAP force plot (Figure 4), which visually confirmed that an increasing number of metastatic sites directly correlated with a higher model output for mortality risk, reinforcing the clinical and biological plausibility of the model's predictions.

XSL•FO
RenderX

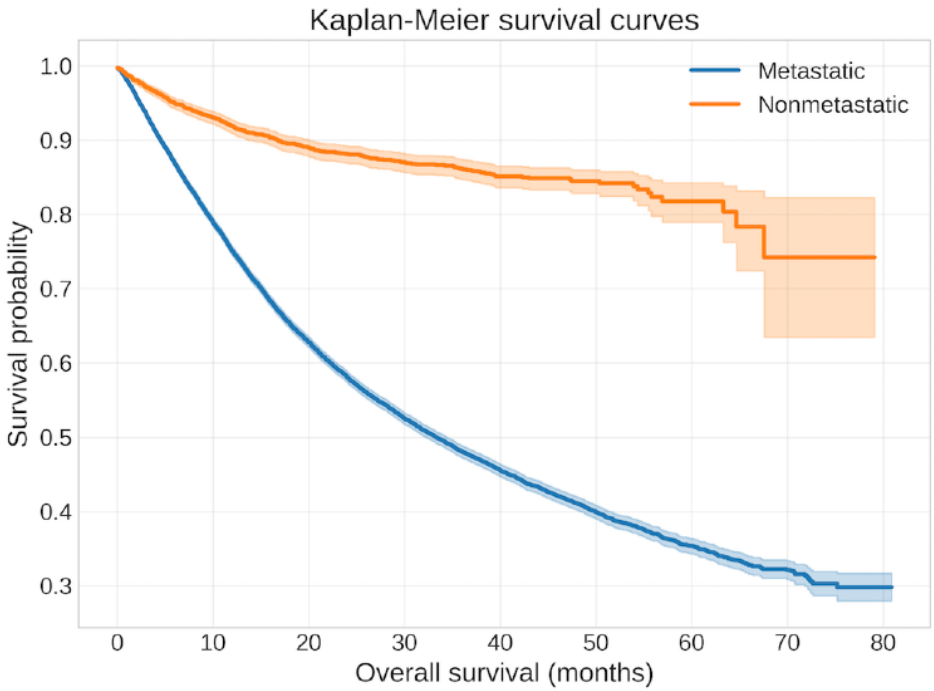
Figure 4. Shapley additive explanations force plot illustrating the impact of metastatic (Met) site count on predicted mortality risk for the first 1000 patients in the Memorial Sloan Kettering-Metastatic test set. Each horizontal bar represents an individual patient, with the baseline prediction shown at the center and shifts toward higher (right, red) or lower (left, blue) mortality risk driven by the number of metastatic sites. The consistent rightward push associated with increasing metastatic site count visually confirms its strong, monotonic contribution to higher predicted risk, reinforcing the biological and clinical plausibility of the extreme gradient boosting model's behavior.



Survival Analysis

In the Kaplan-Meier analysis (Figure 5 and Table S7 in Multimedia Appendix 1), patients were stratified into "metastatic" and "nonmetastatic" groups to compare differences in overall survival. The survival probability of patients in the metastatic group was notably lower than that of the nonmetastatic group, as seen in the pronounced separation of their survival curves. By approximately 80 months, patients with metastatic disease exhibited a survival probability of 0.30 (3041/10,169 patients with metastatic disease) versus 0.80 (8135/10,169 patients without metastatic disease), underscoring the substantial impact of metastatic status on long-term survival outcomes.

Figure 5. Kaplan-Meier survival curves comparing overall survival in patients with vs without metastatic disease in the Memorial Sloan Kettering-Metastatic cohort. Patients with metastases demonstrate markedly lower survival probabilities and earlier median survival times than patients without metastases, underscoring the substantial impact of metastatic status on long-term outcomes.



Subgroup survival results for the top 5 cancer types were as follows: in non–small cell lung cancer (n=4686), 43.75% (n=2050) of patients experienced an event and 82.5% (n=3866) presented with metastatic disease; the median overall survival

XSL•FO
RenderX

was 17.7 (IQR 8.4-33.2) months. In colorectal cancer (n=3541), the event rate was 33.52% (n=1187), the metastatic proportion was 90.82% (n=3216), and the median survival was 17.5 (IQR 10-30) months.

In breast cancer (n=2601), 35.64% (n=927) experienced an event and 76.47% (n=1989) had metastatic disease, with a median survival of 33 (IQR 14.4-43.4) months. In pancreatic cancer (n=1981), the event rate was 60.27% (n=1194), the metastatic proportion was 93.54% (n=1853), and the median survival was 11.6 (IQR 4.2-21.8) months. Finally, in prostate cancer (n=2166), 26.45% (n=573) experienced an event and 81.12% (n=1757) had metastatic disease, with a median survival of 21.6 (IQR 11.6-38.2) months (Figure S5A-S5E in Multimedia Appendix 1).

In the CPH model (Table 4), factors like metastatic site count, TMB, fraction of genome altered, and distant metastases displayed hazard ratios above 1.0, indicating an increased risk of mortality. These relationships attained statistical significance, with *P* values under the established threshold. The proportional hazards assumption was checked through Schoenfeld residuals, and only minor deviations were noted, which did not substantially affect the covariate estimates. The model's C-index reached approximately 0.66, reflecting moderate predictive power in distinguishing survival outcomes among different patient subgroups.

An XGBoost survival model, fitted with a Cox-based loss function, achieved a higher C-index (0.7) than the standard Cox model. In Table S7 in Multimedia Appendix 1, the model's important features are displayed, with distant metastasis, TMB, and fraction of genome altered, among others, listed as the most important features that influence the prediction.

Discussion

Principal Findings

Using a large, multitumor metastatic cohort (MSK-MET), we developed explainable ML models for survivability prediction and complemented them with time-to-event modeling. Among conventional classifiers, XGBoost delivered the best overall performance, achieving an accuracy of 0.74 and AUC=0.82 on the held-out test set. In parallel, an XGBoost-Cox variant improved time-to-event concordance over a standard Cox model (C-index=0.70 vs 0.66, with key hazard ratios from the Cox model). Model explanations (SHAP) and hazard-based analyses converged on clinically recognizable prognostic factors: the number of metastatic sites, TMB, fraction of genome altered, and the presence of liver and bone metastases, while cancer-specific performance varied in biologically plausible ways (eg, higher AUC in prostate and lower in pancreatic). At the Youden-optimal threshold, PPV was approximately 70% and NPV approximately 80%, indicating practical use for risk stratification in balanced decision contexts. We expand the discussion of these findings in the following subsections.

Model Performance, Robustness, and Benchmarking

Our evaluation moved beyond accuracy to provide a comprehensive assessment across discrimination, calibration, operating-point trade-offs, and clinical use. XGBoost consistently emerged as the most dependable model, with its incremental gains in AUC and AUPRC translating into more favorable confusion-matrix profiles at clinically relevant thresholds. For instance, when sensitivity is constrained to be high, XGBoost retains more specificity and a higher MCC, reducing FPs without sacrificing case finding. Its well-calibrated probability estimates are crucial for decision support, enabling rational threshold selection and clear communication of absolute risk.

The model's practical use is further evidenced by its performance across different decision contexts. For balanced decision-making, a default threshold near 0.50 (coinciding with the Youden J optimum) provides a sensible starting point. In triage-like scenarios demanding high sensitivity, a lower threshold around 0.33 yields sensitivity near 0.85 with tolerable specificity losses and an NPV around 0.80. Decision-curve analysis confirmed that XGBoost provides a larger net benefit across a wide band of threshold probabilities, suggesting robustness to varying clinical preferences.

These performance results harmonize with model explanations and survival evidence. SHAP analyses validate that predictions are driven by biologically sensible covariates, while survival curves and hazard ratios show coherent, directionally consistent effects. This triangulation across discrimination, calibration, interpretability, and survival analysis adds credibility that the learned signal reflects underlying disease biology rather than being an artifact of the classifier or data split.

Parameter Sensitivity and Model Robustness

During extensive grid-search tuning, we found that XGBoost hyperparameters, especially tree depth, learning rate, and number of estimators, greatly influenced AUC and classification accuracy. Shallow trees underfit, while deeper ones improved performance but risked overfitting in smaller cancer-specific cohorts. Learning rates below 0.05 caused slow convergence, while overly high rates destabilized training. Despite this variability, certain predictive features, particularly metastatic burden and genomic alterations, remained consistently impactful, underscoring the robustness of our model. Future research could explore adaptive optimization techniques such as Bayesian optimization or reinforcement learning for enhanced generalizability.

Is the Performance "Good Enough"?

The discriminatory performance of prognostic models is central to their clinical use. In oncology, C-index or AUC of at least 0.70 is generally regarded as the minimum threshold for clinical usefulness, while values exceeding 0.80 are considered strong and often necessary for clinical translation [36,37]. These interpretive standards are rooted in established methodological work on ROC analysis, where AUC values between 0.7 and 0.8 are typically described as "acceptable" and those above 0.8 as "excellent" [36,37].

Evidence from systematic reviews of head and neck cancer prognostic models demonstrates how these thresholds translate into practice. Philip et al [38] reported that most radiomics-based prognostic models achieved C-indices below 0.70, underscoring their weak discriminatory power and limited clinical

XSL•FO
RenderX

applicability. In contrast, externally validated models that approached or exceeded the 0.80 mark, as summarized by Dretzke et al [39], were identified as more robust and clinically relevant, supporting their potential integration into patient management. These findings emphasize that models below the 0.70 threshold may remain academically interesting but lack sufficient reliability for routine use, whereas those meeting or surpassing 0.80 provide stronger grounds for clinical adoption.

Comparisons with deep learning approaches further reinforce this interpretation. Gouthamchand et al [40] found that the most competitive deep learning models in head and neck oncology consistently achieved AUCs of 0.80 or higher, placing them above many handcrafted radiomics models. This pattern suggests that 0.80 represents not merely an aspirational benchmark but an increasingly practical requirement for prognostic tools aiming to match or surpass the performance of advanced ML methods already being piloted in clinical contexts.

Regulatory precedents also align with this threshold. Analyses of Food and Drug Administration (FDA)–cleared AI- and ML-enabled medical devices, such as the LumineticsCore software, indicate that most systems demonstrate discriminatory performance in the high 0.7 to 0.9 range, with pivotal studies for devices such as IDx-DR reporting AUCs above 0.80. Joshi et al [41] provide a comprehensive overview of this landscape, confirming that AUCs in this range are characteristic of AI systems deemed safe and effective for clinical use. FDA guidance similarly emphasizes the importance of demonstrating robust discriminatory performance in supporting claims of clinical effectiveness [42].

However, the thresholds discussed here should not be treated as rigid cutoffs. Clinical acceptability is not defined by a single number but by the context of the decision, the potential harms and benefits, and the baseline risk of the population. In high-stakes oncology decisions such as treatment intensification that carries significant morbidity, a well-calibrated model with an AUC of 0.78 but demonstrable net benefit at clinically relevant thresholds may be more valuable than a model with an AUC of 0.84 that is poorly calibrated or unstable. Thus, performance thresholds should be interpreted as contextual benchmarks rather than universal standards.

It is also important to recognize that discrimination does not equal usefulness. AUC or C-index quantifies a model's ability to rank patients by risk, but it does not measure whether the absolute risk estimates are accurate. For clinical practice, calibration is equally critical whether a patient predicted to have a 30% risk actually experiences that outcome about 30% of the time. Calibration-in-the-large, calibration slope, Brier score, and visual calibration plots provide this essential information. Moreover, calibration at clinically meaningful cut points (eg, deciles of predicted risk) is necessary to ensure predictions are trustworthy for patient counseling and decision-making.

Finally, prognostic end points in oncology are often time-to-event outcomes. In such settings, the C-index is commonly used, but time-dependent AUCs provide more clinically interpretable information by specifying the prediction horizon (eg, 12-month or 36-month survival). This allows clinicians to understand how well the model performs over the actual time frames relevant to follow-up and treatment decisions. Where applicable, handling of censoring and competing risks should also be reported, particularly if the terminal outcome is mortality.

Taken together, these considerations emphasize that while an AUC ≥0.80 is encouraging and aligns with many published benchmarks, true clinical use depends on calibration, decision-curve analysis, and performance at specific decision thresholds relevant to patient care.

Clinically Actionable Metrics, Implications, and Use

The performance metrics of our optimal XGBoost model are translated into clinically actionable insights through decision-curve analysis. At the threshold maximizing the Youden index, the model achieves a PPV of 70.3% and an NPV of 79.8%. This indicates that a high-risk prediction from the model would correspond to an actual terminal outcome in approximately 7 out of 10 cases, supporting its use in justifying intensified monitoring or treatment. Conversely, the high NPV means 8 out of 10 low-risk predictions correctly identify patients with a more favorable prognosis, providing a quantitative basis for discussions about de-escalating care and reducing treatment-related morbidity.

Critically, the decision-curve analysis confirms the model's practical use across a spectrum of clinical decision-making preferences. The analysis demonstrates that the XGBoost model provides a superior net benefit compared to both alternative models and the default strategies of treating all or no patients across a wide range of threshold probabilities. This indicates that using the model to guide decisions is clinically advantageous regardless of whether the clinician prioritizes avoiding FPs (overtreatment) or FNs (missed interventions). The point at which the net benefit of the model crosses the "treat all" strategy is particularly important, as it defines the minimum probability at which the model's prediction becomes more useful than intervening in every case. The sustained positive net benefit of our model underscores its robustness and potential to improve patient outcomes by aligning interventions with individualized risk.

The combination of strong predictive performance and model explainability supports several immediate clinical use cases: (1) patient triage and counseling via individualized risk summaries, (2) treatment planning and shared decision-making aided by feature-level rationales, and (3) integration into electronic health record (EHR)–embedded dashboards for longitudinal monitoring. The top predictive features are routinely available in most cancer centers, facilitating adoption with minimal workflow disruption.

Model Explainability and Biological Plausibility

Our model selection was guided by metrics that reflect clinical reality, prioritizing both accuracy and AUROC. The AUROC is especially critical in medical settings, as it captures the essential balance between TPs and FPs, where the cost of FNs (missed cases) is high [43]. This evaluation consistently identified XGBoost as the top performer. Its superior ability to capture complex, nonlinear relationships within high-dimensional clinical and genomic data, as evidenced by

XSL•FO
RenderX

its significant lead over other models, aligns with established literature on gradient boosting for cancer predictions [44,45]. The model's discriminative power is substantial, with an AUC of 0.82 representing an excess over chance of 0.32. This can be intuitively communicated as a "number needed to screen" of approximately 3-4 patients to identify 1 additional correct classification relative to chance, providing a tangible sense of clinical yield. This strong performance is operationalized at the Youden-optimal threshold, where PPV of approximately 70% and NPV of approximately 80% offer concrete use for risk stratification, effectively ruling in or ruling out high-risk status to guide downstream clinical actions.

Crucially, this high performance is rendered transparent and trustworthy through SHAP-based interpretation. The model's decisions are not black box outputs but are demonstrably driven by biologically plausible features that align with oncological principles. SHAP analysis consistently identified metastatic site count, liver and bone metastases, TMB, and fraction of genome altered as the top predictors; moreover, these findings are well-supported in the literature [46]. This biological plausibility was further refined and validated in our cancer-specific models. For example, the prominence of lung metastases in non–small cell lung cancer and prostate cancer models, and the major role of sample type in the breast cancer model, demonstrate a nuanced understanding of disease-specific pathophysiology as pointed out in other studies [47]. The distinct, right-skewed SHAP distribution for "metastatic count" and its direct correlation with risk in force plots further reinforced the importance of total lesion burden in driving high-risk predictions. This coherence between model explanations and established clinical knowledge across both pan-cancer and disease-specific contexts is fundamental for building clinician trust and facilitating the integration of this tool into point-of-care decision support.

Survival Modeling and Risk Stratification

Time-to-Event Insights

Survival analysis confirmed classification results. Kaplan-Meier curves revealed steep survival drops in patients with versus without metastatic disease (0.3 vs 0.8 at 80 months). Traditional Cox models identified high hazard ratios for metastatic site count, TMB, and fraction of genome altered but struggled with their linearity assumptions (C-index=0.66). In contrast, XGBoost-based survival modeling better captured nonlinearity, achieving a higher C-index of 0.70.

Subgroup Survival Analysis

The 5 disease-specific results highlight substantial heterogeneity in survival outcomes across cancer types within the same analytic framework. Pancreatic cancer exhibited the shortest median survival alongside the highest event rate and metastatic proportion, underscoring its aggressive clinical course even within a predominantly metastatic cohort. By contrast, breast cancer showed the longest median survival (as visualized in the respective Kaplan-Meier curves) despite a sizable metastatic share, suggesting comparatively slower disease trajectories and greater effectiveness of available therapies in this subgroup. Non–small cell lung cancer and colorectal cancer shared similar median survivals near 17-18 months, though colorectal cancer carried the highest metastatic proportion among the five. This juxtaposition implies that crude metastatic prevalence alone does not fully account for survival differences, motivating disease-specific modeling of covariates and sites of spread. Prostate cancer combined the lowest event rate with a midrange metastatic proportion and an intermediate median survival, indicating a slower accumulation of events over time relative to the other cancers. Together, these patterns justify presenting separate Kaplan-Meier curves and parsimonious Cox summaries per cancer type, while enabling a consistent cross-cancer narrative that focuses on differences in event rates, metastatic burden, and median survival.

Novelty and Comparison to Prior Work

Although significant studies [48-52] (Table S8 in Multimedia Appendix 1) have been conducted in the realm of predicting cancer survivability, our work stands out by bringing new enhancements that significantly contribute to better prediction of cancer survivability, particularly by the thorough comparison of ML models, the strategic use of both global and cancer-specific models, in-depth model explainability using SHA*P* values, and detailed survival analysis.

First, we begin by comparing 5 different ML models (XGBoost, naive Bayes, decision tree, logistic regression, and random forest), each rigorously tuned using exhaustive grid search for hyperparameters. This approach ensures that each model is thoroughly tested for the task of predicting cancer survivability, a detail that is often overlooked in existing literature. Many studies tend to focus on one or two models, without ensuring that the models are fully evaluated for comparison. For example, prior work done by Zhao et al [48], Tapak et al [49], and Nicolò et al [50] evaluate models but lack the thoroughness in hyperparameter tuning that our study provides. We believe that this rigorous approach enhances the reliability of our findings and provides a more comprehensive understanding of which model performs best under specific conditions.

Moreover, we use a methodology that is designed to first use a global model to gain a general overview of the most important patterns and predictors for metastatic cancer survivability, followed by a deeper dive into cancer-specific models. This 2-tiered approach is critical because it allows us to identify broad patterns while also uncovering nuances that might be missed or misinterpreted in a global model. Many published studies, such as Kourou et al [51] and Zhao et al [48], predominantly dwell on global accuracy metrics without taking this crucial next step to explore more specific patterns within subgroups of the data. By contrast, our approach provides a dual perspective, that is, broad insights from the global model and detailed, cancer-specific insights that we believe are essential for advancing personalized medicine.

The use of SHA*P* values in our study is particularly noteworthy. We did not just stop at model performance but delved deep into explainability, first for the global model and then for the cancer-specific models. This process allowed us to generate refined explainability that highlights not just which features are important, but how their importance varies across different types of cancer. The use of SHAP in both global and specific contexts

XSL•FO
RenderX

is a novel approach that we believe adds substantial value to the study. While many studies, such as Maouche et al [52], use SHAP or similar methods, they often do so at a surface level, without the comprehensive, model-specific analysis that we provide. This depth of analysis is crucial for understanding the true implications of the model's predictions and for making informed clinical decisions.

Finally, we conduct survival analysis after predicting cancer survival with explainable ML to translate model predictions into clinically meaningful insights about patient outcomes over time. While ML classifiers can distinguish between patients likely to survive or not, survival analysis provides a time-to-event perspective, capturing not just "if" but "when" an event such as death occurs. This temporal dimension allows researchers and clinicians to estimate hazard rates, median survival times, and differences between risk groups. Coupled with explainable ML (eg, SHA*P* values), survival analysis also helps validate whether the features driving predictions correspond to biologically and clinically relevant risk factors. Together, this integration strengthens both the predictive performance and interpretability of the model, supporting its potential use as a trustworthy decision-support tool in oncology practice.

Limitations and Biases

We acknowledge several limitations that must be considered alongside the model's strong performance. First, generalizability may be constrained by inherent biases in the MSK-MET dataset, which is a single-institution cohort from a tertiary care center. The sample underrepresents socioeconomically disadvantaged and racially diverse populations, with non-White ethnicities comprising less than 15%, and genomic profiling was preferentially performed in advanced-stage cases, potentially inflating the importance of features like TMB. These biases may partly explain the performance disparities observed across cancer types. Furthermore, technical limitations exist; for instance, linear and naive Bayes baselines may have been disadvantaged by the use of label encoding for categorical variables.

These factors underscore the necessity for several future steps before clinical deployment. External validation using diverse, multicenter cohorts with stratified sampling is crucial to improve equity, fairness, and generalizability. Prospective calibration should be rechecked under local prevalence shifts, with techniques like Platt scaling or isotonic regression applied, if necessary. Additionally, prospective benchmarking against oncologist-estimated prognoses and real-world deployment outcomes will be critical to establish clinical noninferiority or superiority. Finally, while this study provides a comprehensive classification and survival analysis, future work should focus on integrating these outputs; for example, by generating risk-stratified Kaplan-Meier curves at chosen thresholds to unify the decision framework across binary and time-to-event end points.

Future Directions and Deployment Considerations

Improving Rare Cancer Predictions

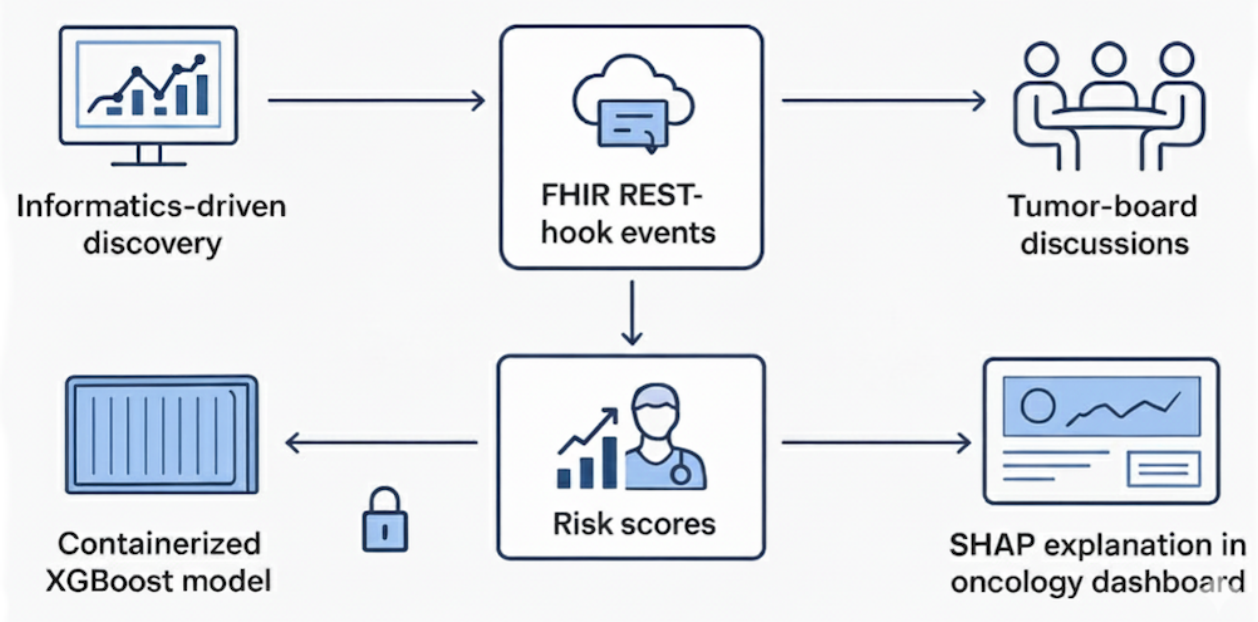
Model performance was limited for rare cancers (eg, anal cancer: n=68; AUC=0.61), due to data scarcity. To address this, we suggest (1) using transfer learning to initialize models with global XGBoost weights, (2) using synthetic oversampling (eg, synthetic minority oversampling technique, adaptive synthetic sampling) during training, and (3) leveraging federated learning to aggregate data across institutions while preserving privacy. These techniques can bolster performance in underrepresented malignancies.

EHR Integration and Clinical Deployment

While the primary aim of this study was informatics-driven discovery, the robustness of our model supports a clear pathway for clinical translation through EHR integration. We envision a real-time, standards-based system where automated risk alerts are seamlessly integrated into clinical workflows, particularly within tumor-board discussions (Figure 6). By leveraging Fast Healthcare Interoperability Resources Representational State Transfer hook events, patient risk scores could be dynamically updated and surfaced directly within the EHR as cases are reviewed. To ensure both privacy and computational efficiency, the XGBoost model would be deployed as a containerized service within the hospital's secure analytics infrastructure (eg, Epic Cogito), keeping protected health information on premises while using available hardware acceleration.

XSL•FO
RenderX

Figure 6. Electronic health record (EHR) integration and clinical deployment. Framework illustrating how patient risk scores and Shapley additive explanations (SHAP)–based explanations are integrated into the EHR and surfaced during tumor-board discussions to enable transparent, real-time clinical decision support. FHIR: Fast Healthcare Interoperability Resources; REST: Representational State Transfer; XGBoost: extreme gradient boosting.



A critical differentiator of this approach is the coupling of each risk alert with a SHAP-based explanation embedded directly into the oncology dashboard. This transforms the model from a "black box" into an "explain-and-act" tool, providing clinicians with immediate, interpretable rationale by highlighting the top clinical and genomic features contributing to an individual's risk score. Alert protocols would be tiered and threshold-aware, directly translating the model's predictive values into actionable clinical guidance. For instance, a low-risk classification (supported by an NPV of 79.8%) could justify lengthening follow-up intervals, while a high-risk flag (PPV of 70.3%) would prompt rapid biomarker reassessment and discussion of treatment intensification.

An initial rollout would prioritize safety and reliability through a prospective pilot study, continuous performance monitoring for model drift, and an implementation that minimally disrupts existing workflows. The validation evidence and explainable framework presented in this study provide a solid foundation for a potential regulatory submission as a Class II clinical decision support tool, paving the way for a new generation of transparent, AI-augmented oncology care.

Conclusion

In this large-scale, pan-cancer study, we developed and validated an interpretable ML framework for predicting survivability in patients with metastatic cancer. By leveraging the comprehensive MSK-MET cohort, we demonstrated that an XGBoost classifier robustly predicts overall survival (AUC=0.82, accuracy=0.74), outperforming other conventional ML models. Crucially, the integration of SHAP explainability illuminated the model's decision-making process, consistently identifying metastatic site count, TMB, fraction of genome altered, and the presence of liver and bone metastases as major prognostic features across diverse tumor types. This biological plausibility is fundamental for building clinical trust.

Our 2-tiered modeling approach, combining a unified pan-cancer perspective with targeted, cancer-specific submodels, provided both broad generalizability and nuanced, disease-tailored insights. This was evidenced by the varying performance across cancer types, such as the high predictive accuracy for prostate cancer (AUC=0.88) contrasted with the greater challenges in pancreatic cancer (AUC=0.68). Furthermore, the survival analysis corroborated the classification findings, with the XGBoost-Cox model (C-index=0.70) capturing nonlinear relationships more effectively than the traditional Cox model (C-index=0.66), and Kaplan-Meier curves starkly illustrating the significant survival disadvantage associated with metastatic disease.

The clinical use of our model is underscored by its strong predictive values (PPV=70% and NPV=80% at the Youden-optimal threshold) and its demonstrated net benefit across a range of decision thresholds. By reconciling high performance with transparent, actionable explanations, this work provides a foundational framework for the next generation of clinical decision-support tools in oncology. Future efforts should focus on external validation in multicenter cohorts, prospective evaluation integrated within EHR systems, and addressing performance gaps in rare cancer subtypes to ensure equitable and widespread clinical adoption.

Acknowledgments

We thank cBioPortal for Cancer Genomics for freely making available the data that were used in this study. Finally, we are grateful to the FABRIC [53] team for their support.

XSL•FO
RenderX

There was no generative artificial intelligence tool used. Responsibility for the final manuscript lies entirely with the authors.

Funding

This work was supported by the National Science Foundation under Grant No. 2201583. The sponsor had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The datasets used in this study are available in the Memorial Sloan Kettering-Metastatic repository at cBioPortal [54]. The code is available at GitHub [55].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional tables and figures.
[DOCX File , 2147 KB - cancer_v12i1e74196_app1.docx]

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024;74(1):12-49 [FREE Full text] [doi: 10.3322/caac.21820] [Medline: 38230766]
2. Li L, Xiong F, Wang Y, Zhang S, Gong Z, Li X, et al. What are the applications of single-cell RNA sequencing in cancer research: a systematic review. J Exp Clin Cancer Res 2021;40(1):163 [FREE Full text] [doi: 10.1186/s13046-021-01955-1] [Medline: PMC8111731]
3. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. Biochim Biophys Acta 2010;1805(1):105-117 [FREE Full text] [doi: 10.1016/j.bbcan.2009.11.002]
4. Li L, Xiong F, Wang Y, Zhang S, Gong Z, Li X, et al. What are the applications of single-cell RNA sequencing in cancer research: a systematic review. J Exp Clin Cancer Res 2021;40(1):163-193 [FREE Full text] [doi: 10.1186/s13046-021-01955-1] [Medline: 33975628]
5. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542(7639):115-118 [FREE Full text] [doi: 10.1038/nature21056] [Medline: 28117445]
6. Liu Y, Kim J, Balagurunathan Y, Hawkins S, Stringfield O, Schabath MB, et al. Prediction of pathological nodal involvement by CT-based radiomic features of the primary tumor in patients with clinically node-negative peripheral lung adenocarcinomas. Med Phys 2018;45(6):2518-2526 [FREE Full text] [doi: 10.1002/mp.12901] [Medline: 29624702]
7. Montazeri M, Montazeri M, Montazeri M, Beigzadeh A. Machine learning models in breast cancer survival prediction. Technol Health Care 2016;24(1):31-42. [doi: 10.3233/THC-151071] [Medline: 26409558]
8. Zhou L, Wu X, Huang S, Wu G, Ye H, Wei Q, et al. Lymph node metastasis prediction from primary breast cancer US images using deep learning. Radiology 2020;294(1):19-28. [doi: 10.1148/radiol.2019190372] [Medline: 31746687]
9. Foersch S, Eckstein M, Wagner DC, Gach F, Woerl A, Geiger J, et al. Deep learning for diagnosis and survival prediction in soft tissue sarcoma. Ann Oncol 2021;32(9):1178-1187 [FREE Full text] [doi: 10.1016/j.annonc.2021.06.007] [Medline: 34139273]
10. Doppalapudi S, Qiu RG, Badr Y. Lung cancer survival period prediction and understanding: Deep learning approaches. Int J Med Inform 2021;148:104371. [doi: 10.1016/j.ijmedinf.2020.104371] [Medline: 33461009]
11. Zhang G, Shen X, Zhang YD, Luo Y, Luo J, Zhu D, et al. Cross-modal prostate cancer segmentation via self-attention distillation. IEEE J Biomed Health Inform 2022;26(11):5298-5309. [doi: 10.1109/JBHI.2021.3127688] [Medline: 34767517]
12. Huang Y, Wei L, Hu Y, Shao N, Lin Y, He S, et al. Multi-parametric MRI-based radiomics models for predicting molecular subtype and androgen receptor expression in breast cancer. Front Oncol 2021;11:706733 [FREE Full text] [doi: 10.3389/fonc.2021.706733] [Medline: 34490107]
13. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. : Association for Computing Machinery; 2016 Presented at: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16); 2016 Aug 13; San Francisco, CA. New York, NY URL: https://doi.org/10.1145/2939672.2939785 [doi: 10.1145/2939672.2939785]
14. Liu Y, Wang Z, Cao X, Liu M, Zhong L. Machine learning models for predicting survival in lung cancer patients undergoing microwave ablation. Front Med (Lausanne) 2025;12:1561083 [FREE Full text] [doi: 10.3389/fmed.2025.1561083] [Medline: 40400639]
15. Noman SM, Fadel YM, Henedak MT, Attia NA, Essam M, Elmaasarawii S, et al. Leveraging survival analysis and machine learning for accurate prediction of breast cancer recurrence and metastasis. Sci Rep 2025;15(1):3728 [FREE Full text] [doi: 10.1038/s41598-025-87622-3] [Medline: 39880868]

XSL•FO
RenderX

16. Buk Cardoso L, Cunha Parro V, Verzinhasse Peres S, Curado MP, Fernandes GA, Wünsch Filho V, et al. Machine learning for predicting survival of colorectal cancer patients. Sci Rep 2023;13(1):8874 [FREE Full text] [doi: 10.1038/s41598-023-35649-9] [Medline: 37264045]
17. Ben Kridis W, Wang Z, Guven DC, Dharmarajan A. Editorial: advances in the treatment of hormonal receptor positive (HR+) breast cancer. Front Oncol 2024;14:1449566 [FREE Full text] [doi: 10.3389/fonc.2024.1449566] [Medline: 39386197]
18. Breiman L. Random Forests. Mach Learn 2001;45(1):5-32 [FREE Full text] [doi: 10.1023/a:1010933404324]
19. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20(3):273-297 [FREE Full text] [doi: 10.1007/bf00994018]
20. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436-444 [FREE Full text] [doi: 10.1038/nature14539] [Medline: 26017442]
21. Liu Z, Liu L, Weng S, Guo C, Dang Q, Xu H, et al. Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. Nat Commun 2022;13(1):816 [FREE Full text] [doi: 10.1038/s41467-022-28421-6] [Medline: 35145098]
22. He B, Bergenstråhle L, Stenbeck L, Abid A, Andersson A, Borg Å, et al. Integrating spatial gene expression and breast tumour morphology via deep learning. Nat Biomed Eng 2020;4(8):827-834. [doi: 10.1038/s41551-020-0578-x] [Medline: 32572199]
23. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. Cancer Genomics Proteomics 2018;15(1):41-51 [FREE Full text] [doi: 10.21873/cgp.20063] [Medline: 29275361]
24. Nguyen TB, Do DN, Nguyen-Thi M, Hoang-The H, Tran T, Nguyen-Thanh T. Identification of potential crucial genes and key pathways shared in inflammatory bowel disease and cervical cancer by machine learning and integrated bioinformatics. Comput Biol Med 2022;149:105996. [doi: 10.1016/j.compbiomed.2022.105996] [Medline: 36049413]
25. Stefanou IK, Dovrolis N, Gazouli M, Theodorou D, Zografos GK, Toutouzas KG. miRNAs expression pattern and machine learning models elucidate risk for gastric GIST. Cancer Biomark 2022;33(2):237-247 [FREE Full text] [doi: 10.3233/CBM-210173] [Medline: 35213356]
26. Pawelka D, Laczmanska I, Karpinski P, Supplitt S, Witkiewicz W, Knychalski B, et al. Machine-learning-based analysis identifies miRNA expression profile for diagnosis and prediction of colorectal cancer: a preliminary study. Cancer Genomics Proteomics 2022;19(4):503-511 [FREE Full text] [doi: 10.21873/cgp.20336] [Medline: 35732322]
27. Ghobadi MZ, Emamzadeh R, Afsaneh E. Exploration of mRNAs and miRNA classifiers for various ATLL cancer subtypes using machine learning. BMC Cancer 2022;22(1):433 [FREE Full text] [doi: 10.1186/s12885-022-09540-1] [Medline: 35449091]
28. Ding W, Chen G, Shi T. Integrative analysis identifies potential DNA methylation biomarkers for pan-cancer diagnosis and prognosis. Epigenetics 2019;14(1):67-80 [FREE Full text] [doi: 10.1080/15592294.2019.1568178] [Medline: 30696380]
29. Liang N, Li B, Jia Z, Wang C, Wu P, Zheng T, et al. Ultrasensitive detection of circulating tumour DNA via deep methylation sequencing aided by machine learning. Nat Biomed Eng 2021;5(6):586-599. [doi: 10.1038/s41551-021-00746-5] [Medline: 34131323]
30. Avanzo M, Stancanello J, Pirrone G, Sartor G. Radiomics and deep learning in lung cancer. Strahlenther Onkol 2020;196(10):879-887. [doi: 10.1007/s00066-020-01625-9] [Medline: 32367456]
31. Janssen BV, Verhoef S, Wesdorp NJ, Huiskens J, de Boer OJ, Marquering H, et al. Imaging-based machine-learning models to predict clinical outcomes and identify biomarkers in pancreatic cancer: a scoping review. Ann Surg 2022;275(3):560-567. [doi: 10.1097/SLA.0000000000005349] [Medline: 34954758]
32. Stanzione A, Verde F, Romeo V, Boccadifuoco F, Mainenti PP, Maurea S. Radiomics and machine learning applications in rectal cancer: current update and future perspectives. World J Gastroenterol 2021;27(32):5306-5321 [FREE Full text] [doi: 10.3748/wjg.v27.i32.5306] [Medline: 34539134]
33. Wang R, Dai W, Gong J, Huang M, Hu T, Li H, et al. Development of a novel combined nomogram model integrating deep learning-pathomics, radiomics and immunoscore to predict postoperative outcome of colorectal cancer lung metastasis patients. J Hematol Oncol 2022;15(1):11 [FREE Full text] [doi: 10.1186/s13045-022-01225-3] [Medline: 35073937]
34. Nguyen B, Fong C, Luthra A, Smith SA, DiNatale RG, Nandakumar S, et al. Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. Cell 2022;185(3):563-575.e11 [FREE Full text] [doi: 10.1016/j.cell.2022.01.003] [Medline: 35120664]
35. Ramírez-Mena A, Andrés-León E, Alvarez-Cubero MJ, Anguita-Ruiz A, Martinez-Gonzalez LJ, Alcala-Fdez J. Explainable artificial intelligence to predict and identify prostate cancer tissue by gene expression. Comput Methods Programs Biomed 2023;240:107719 [FREE Full text] [doi: 10.1016/j.cmpb.2023.107719] [Medline: 37453366]
36. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 2010;5(9):1315-1316 [FREE Full text] [doi: 10.1097/JTO.0b013e3181ec173d] [Medline: 20736804]
37. Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. Korean J Anesthesiol 2022;75(1):25-36 [FREE Full text] [doi: 10.4097/kja.21209] [Medline: 35124947]
38. Philip MM, Welch A, McKiddie F, Nath M. A systematic review and meta-analysis of predictive and prognostic models for outcome prediction using positron emission tomography radiomics in head and neck squamous cell carcinoma patients. Cancer Med 2023;12(15):16181-16194 [FREE Full text] [doi: 10.1002/cam4.6278] [Medline: 37353996]

XSL•FO
RenderX

39. Dretzke J, Abou-Foul AK, Albon E, Hillier B, Scandrett K, Price MJ, et al. Systematic review of prognostic models for predicting recurrence and survival in patients with treated oropharyngeal cancer. BMJ Open 2024;14(12):e090393 [FREE Full text] [doi: 10.1136/bmjopen-2024-090393] [Medline: 39638589]
40. Gouthamchand V, Fonseca LAF, Hoebers FJP, Fijten R, Dekker A, Wee L. Prognostic modeling in headneck cancer: deep learning or handcrafted radiomics? BJR Artif Intell 2025;2(1):ubaf008 [FREE Full text] [doi: 10.1093/bjrai/ubaf008]
41. Joshi G, Jain A, Araveeti SR, Adhikari S, Garg H, Bhandari M. FDA-approved artificial intelligence and machine learning (AI/ML)-Enabled medical devices: an updated landscape. Electronics 2024;13(3):498. [doi: 10.3390/electronics13030498]
42. Artificial intelligence and machine learning (AI/ML)-enabled medical devices. US Food and Drug Administration. 2025. URL: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices [accessed 2025-11-24]
43. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett 2006;27(8):861-874 [FREE Full text] [doi: 10.1016/j.patrec.2005.10.010]
44. Wolpert DH. Stacked generalization. Neural Netw 1992;5(2):241-259 [FREE Full text] [doi: 10.1016/s0893-6080(05)80023-1]
45. Kuncheva L, Whitaker C. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. Mach Learn 2003;51(2):181-207 [FREE Full text] [doi: 10.1023/a:1022859003006]
46. Davoli T, Uno H, Wooten EC, Elledge SJ. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. Science 2017;355(6322):eaaf8399 [FREE Full text] [doi: 10.1126/science.aaf8399] [Medline: 28104840]
47. Tsimberidou A. Targeted therapy in cancer. Cancer Chemother Pharmacol 2015;76(6):1113-1132 [FREE Full text] [doi: 10.1007/s00280-015-2861-1] [Medline: 26391154]
48. Zhao B, Gabriel RA, Vaida F, Lopez NE, Eisenstein S, Clary BM. Predicting overall survival in patients with metastatic rectal cancer: a machine learning approach. J Gastrointest Surg 2020;24(5):1165-1172 [FREE Full text] [doi: 10.1007/s11605-019-04373-z] [Medline: 31468331]
49. Tapak L, Shirmohammadi-Khorram N, Amini P, Alafchi B, Hamidi O, Poorolajal J. Prediction of survival and metastasis in breast cancer patients using machine learning classifiers. Clin Epidemiol Glob Health 2019;7(3):293-299. [doi: 10.1016/j.cegh.2018.10.003]
50. Nicolò C, Périer C, Prague M, Bellera C, MacGrogan G, Saut O, et al. Machine learning and mechanistic modeling for prediction of metastatic relapse in early-stage breast cancer. JCO Clin Cancer Inform 2020;4:259-274. [doi: 10.1200/CCI.19.00133] [Medline: 32213092]
51. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 2015;13:8-17 [FREE Full text] [doi: 10.1016/j.csbj.2014.11.005] [Medline: 25750696]
52. Maouche I, Terrissa LS, Benmohammed K, Zerhouni N. An explainable AI approach for breast cancer metastasis prediction based on clinicopathological data. IEEE Trans Biomed Eng 2023;70(12):3321-3329. [doi: 10.1109/TBME.2023.3282840] [Medline: 37276094]
53. Baldin I, Nikolich A, Griffioen J, Monga I, Wang K, Lehman T, et al. FABRIC: a national-scale programmable experimental network infrastructure. IEEE Internet Comput 2019;23(6):38-47. [doi: 10.1109/MIC.2019.2958545]
54. cBioPortal. URL: https://www.cbioportal.org/study/summary?id=msk_met_2021 [accessed 2026-01-02]
55. MU-Data-Science/GAF. GitHub. URL: https://github.com/MU-Data-Science/GAF/tree/main/metastatic [accessed 2026-01-02]

Abbreviations

AI: artificial intelligence
AUC: area under the curve
AUPRC: area under the precision-recall curve
AUROC: area under the receiver operating characteristic curve
C-index: concordance index
CPH: Cox proportional hazards
CV: cross validation
EHR: electronic health record
FDA: US Food and Drug Administration
FN: false negative
FP: false positive
MCC: Matthews correlation coefficient
ML: machine learning
MSK-MET: Memorial Sloan Kettering-Metastatic
NPV: negative predictive value
PPV: positive predictive value
ROC: receiver operating characteristic

XSL•FO
RenderX

SHAP: Shapley additive explanations
TMB: tumor mutational burden
TN: true negative
TP: true positive
XGBoost: extreme gradient boosting

Edited by N Cahill; submitted 19.Mar.2025; peer-reviewed by AMYD Alahakoon, AH Hounye, D Patel; comments to author 28.May.2025; revised version received 18.Nov.2025; accepted 19.Nov.2025; published 13.Jan.2026.

Please cite as:
Nalela P, Rao D, Rao P
Explainable AI for Predicting Mortality Risk in Metastatic Cancer: Retrospective Cohort Study Using the Memorial Sloan Kettering-Metastatic Dataset
JMIR Cancer 2026;12:e74196
URL: https://cancer.jmir.org/2026/1/e74196
doi:10.2196/74196
PMID:

©Polycarp Nalela, Deepthi Rao, Praveen Rao. Originally published in JMIR Cancer (https://cancer.jmir.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessment of Predictive Factors That Shorten Duration of Treatment in Patients With Multiple Myeloma Using AI: Real-World Longitudinal Study Using Data From Medical Data Vision Claims Database

Hiroshi Handa[1], MD, PhD; Tadao Ishida[2], MD, PhD; Shuji Ozaki[3], MD, PhD; Shinsuke Iida[4], MD, PhD; Kittima Wattanakamolkul[5], BSc, MSc, PhD; Chika Sakai[6], BPharm; Kenichi Kato[6], BSIS; David Bin-Chia Wu[7,8,9], MSc, PhD, BSc; DaeYoung Yu[7], BSc, MPH, PhD; Shota Nemoto[10], MIB, BE; Yasuho Yamashita[11], PhD; Takuma Shibahara[11], PhD

[1]Department of Hematology, Gunma University Hospital, Gunma, Japan

[2]Department of Hematology, Japanese Red Cross Medical Center, Tokyo, Japan

[3]Department of Hematology, Tokushima Prefectural Central Hospital, Tokushima, Japan

[4]Department of Hematology and Oncology, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan

[5]Value, Evidence and Access Department, Integrated Market Access, Johnson & Johnson, Tokyo, Japan

[6]Medical Affairs, Johnson & Johnson, Tokyo, Japan

[7]Asia Pacific Regional Market Access, Johnson & Johnson, Singapore, Singapore

[8]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

[9]School of Pharmacy, Faculty of Health & Medical Sciences, Taylor's University Malaysia, Malaysia, Malaysia

[10]Industrial & Digital Business Unit, Hitachi (Japan), Tokyo, Japan

[11]Research and Development Group, Hitachi (Japan), Tokyo, Japan

Corresponding Author:
Kittima Wattanakamolkul, BSc, MSc, PhD
Value, Evidence and Access Department, Integrated Market Access
Johnson & Johnson
3-5-2, Nishi-Kanda, Chiyoda-ku
Tokyo, 101-0065
Japan
Phone: 81 8054 7753
Fax: 81 03 4411 5050
Email: KWattan1@its.jnj.com

Abstract

Background: With the availability of newer therapies, the duration of therapy (DoT) shortens with each increasing line of treatment in Japanese patients with multiple myeloma (MM).

Objective: This study aimed to identify factors that shorten DoT in patients with MM using a machine learning (ML) procedure from the Medical Data Vision (MDV) database.

Methods: This nationwide, retrospective observational real-world study was conducted using anonymized patient data from the MDV claims database from 2003 to 2022. Patients (≥18 y) with transplant-ineligible newly diagnosed MM (continued first line therapy), or relapsed or refractory MM (continued second or third line therapies) were included. To identify important predictive factors, an explainable deep learning model was created using 647 extracted variables (continuous, binary, and nominal categorical) from the MDV database, and the extracted data were used to train ML algorithms to build point-wise linear (PWL) models for predicting DoT. Predictive performance of the PWL model was compared with the elastic net (regularized logistic regression) and the extreme gradient boosting models, and calculated by area under the curve and evaluated by 10-fold double cross-validation. A clustering analysis (k-means method) of 4848 individual samples assessed the relationship between each sample and DoT (3, 6, and 12 months). The characteristics of clusters and sample features belonging to each cluster during and after treatment were studied.

XSL•FO
RenderX

Results: Overall, 2762 (4848 individual samples) patients were evaluated (mean age 69.6, SD 11.8 years; 1450/2762, 52.5% male). The area under the curve score of the PWL model to predict DoT at 3, 6, and 12 months was 0.61, 0.64, and 0.66, respectively. Based on the similarity of coefficients of regression models, samples were categorized into 2 clusters (clusters A and B) at DoT of 3 months, 3 clusters (clusters A, B, and C) at 6 months, and 12 months (clusters A, B, and C). Cluster B versus cluster A (at 3 months) and cluster C versus cluster A and B (at 6 and 12 months) had a significantly (P<.01) higher pretreatment Charlson Comorbidity Index. They also showed a lower median of prediction probability. At 3 months in cluster B and at 6 and 12 months in cluster C, the use of immunomodulatory drugs for MM treatment was significantly higher in patients who met predicted DoT at each threshold versus those who did not. Additionally, the use of aspirin was significantly higher in cluster B and cluster C at 3 and 6 months, respectively.

Conclusions: Applying ML techniques using the PWL model yielded efficient results to understand trends associated with treatment and characteristics of Japanese patients with MM whose DoT were shortened. The study demonstrated that patients' disease status and management-related factors, including use of immunomodulatory drugs and management of thromboprophylaxis, may be associated with DoT length.

(JMIR Cancer 2026;12:e75586) doi:10.2196/75586

KEYWORDS

duration of therapy; machine learning; artificial intelligence; multiple myeloma; prediction probability; database; Japan

Introduction

Multiple myeloma (MM), with an estimated annual incidence of 7.1 per 100,000 population, is the second most common hematologic neoplasm worldwide [1]. As per the Global Cancer Observatory report, there were 6988 new cases of MM and 4827 deaths in 2022 in Japan [2]. With a rapidly aging population, the incidence of MM is projected to increase [3-5].

The treatment choice for patients with newly diagnosed multiple myeloma (NDMM) is based on patient-based factors (age, frailty status, comorbidities, compliance, and lifestyle preference), disease-based factors (cytogenetics abnormalities, mutations, biochemical abnormalities, extramedullary disease, and International Staging System score), treatment-related factors (access to therapy, cost, and safety), and the eligibility to undergo autologous stem cell transplantation postinduction therapy [6]. In Japan, several key therapies have been approved for MM treatment over the past 2 decades: bortezomib (approved in 2006) [7], lenalidomide (approved in 2010) [8], pomalidomide (approved in 2014) [9], and thalidomide (approved in 2008) [10]. These agents have become standard treatment options, demonstrating clinically meaningful improvements in both progression-free and overall survival for patients with MM. For transplant-ineligible patients with NDMM and relapsed or refractory multiple myeloma (RRMM), novel therapies (such as panobinostat, elotuzumab, carfilzomib, ixazomib, daratumumab, and isatuximab) are commonly used treatment modalities [11,12].

Although patient survival is increasing with the advancement in treatment options and supportive care [4,5], MM is still considered an incurable disease as patient survival solely depends on the prescribed treatment regimen and the patient's demographic and clinical characteristics [1].

Treatment continuation for an appropriate duration plays an important role in maximizing treatment efficacy in patients with MM, as it is directly correlated with patient survival [12]. Our previous real-world study using patients' data from the Medical Data Vision (MDV) claims database in Japan found that the attrition rate was increased with each line of treatment (first line [transplant-conducted]-third line). Subsequently, the treatment duration became shorter in Japanese patients with MM from 2016 to 2020 (6.7-8.6 months) compared with 2003-2015 (8.8-10.0 months); the reduction was mostly seen in older adult patients (aged >80 y) [12]. Although the accurate reason for the shorter duration of therapy (DoT) was not established, it was hypothesized that older age, comorbid disease, availability of newer treatment regimens, safety and tolerability of treatments, and activities of daily living (ADLs) could have impacted the DoT in patients with MM [13-15]. Hence, accurate evaluation of the predictive factors that shorten the DoT could help in the treatment decision-making process.

Artificial intelligence (AI) using machine learning (ML) has been continuously used for the diagnosis and treatment of diseases. Using different tools and algorithms, AI eases the human experts' tasks by computerization of the same. Informed evaluations are generated by detecting relationships in information [16]. Therefore, several studies have used different AI and ML algorithms for the diagnosis and risk stratification of MM using patient-level data as well as to create decision-making models [16-18].

This analysis is intended to identify predictive factors that shorten DoT at a threshold of 3, 6, and 12 months in patients with MM who had received first, second, and third-line treatments by applying an ML model using the patients' data (type of first treatment and patient characteristics) from the MDV claims database.

Methods

Data Sources

This was a nationwide, retrospective observational cohort study conducted with secondary use of the Japan claims database over a 19-year period between 2003 and 2022. The study used longitudinal patient data from the MDV (MDV Co, Ltd) claims database, which is built over 450 Japanese hospital administrative data (covers approximately 26% of "Diagnosis Procedure Combination hospitals," ie, the hospitals that perform

XSL•FO
RenderX

higher than the lowest in the university hospital group), containing over 38 million patient data since the year 2003. The MDV database is a national, patient-level, anonymized, and longitudinal claims database including hospital claims, discharge summaries, hospitalization data, and outpatient and prescription data. The *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) was used for disease diagnosis, and the medical procedures were coded using Japanese Procedure Codes. Hospitalized death and other clinical evaluations at discharge were available from the "discharge summary" data.

The study period was divided into a look-back period, treatment (time-at-risk) period, and follow-up period. The look-back period for covariates was the period from the first record of MM in the MDV database till the index date. Time-at-risk was the period from the start of the treatment to the end of the treatment for MM, while follow-up time was the period from the end of the treatment for MM to the last record registered in the MDV database. The index date was the record at the start of the treatment line in patients with transplant-ineligible NDMM or RRMM between 2003 and June 2021.

Study Population

Male and female patients aged 18 years and older at index date with a confirmed diagnosis of MM (transplant-ineligible NDMM or RRMM), between 2003 and 2022, were included. Patients were required to have confirmed administration of any of the treatment regimens for first line for the transplant-ineligible NDMM or second line and third line for the RRMM. The unit of analysis (sample) in this study was an individual treatment of first line used for the patients who were transplant-ineligible NDMM or RRMM, second line, and third line. Patients whose DoT cannot be calculated were excluded from the ML analysis.

Variables and Assessment

This study was conducted on 4848 samples for MM using 647 variables extracted from the MDV database. Variables were selected based on the collected opinions of local clinical experts and publications, and the results of a feasibility analysis. Final variables were determined after consulting local clinical experts on the results of the feasibility analysis and were broadly classified as patient, MM treatment-related, and MM treatment-related comorbidity management variables (Table 1). The primary end point was to identify the factors that correlate with the short DoT at the threshold of 3, 6, and 12 months. The DoT was defined as the time from the initiation of a line of treatment to the last day of the line of treatment.

XSL•FO
RenderX

Table 1. Feature variables extracted from the Medical Data Vision database to identify the important predictors associated with the duration of therapy.

Patient variables	MM[a] treatment-related variables	MM treatment-related comorbidity management variables
• Older age at the MM diagnosis (≥75 y, ≥80 y)[b] • Gender[b] • Height, weight[b] • Comorbidity (Charlson Index)[b] • eGFR[c], blood pressure[b] • Leukocytes, neutrophils[b] • Lactate dehydrogenase[b] • Prothrombin time[b] • Independence (ADL[d]), Performance status[b] • Under nutrition[b] • Cardiac function[b] • Presence or absence of eating or swallowing dysfunction[b] • Frequency and history of depressive symptoms[b,e] • Frequency and history of use of antidepressants[b,e] • Frequency and history of use of hypnotics[b,e] • Year at diagnosis of MM[b]	• Use of Bortezomib, Lenalidomide in previous treatment regimen[b] • Number of previous lines of regimens[b] • Use of triplet therapy in second line[e] • Use of oral drugs in the next line of regimens[e] • Stem cell transplantation conducted after the treatment regimen[e] • Steroid dose, timing of dose reduction[e] • Type of drug formulations in each regimen (eg, oral, with IV[f], or with SC[g])[e] • Type of hospital (ie, oncology-focused site), hospital size[b] • Type of treatment at molecule level (eg, IMiDs[h]-based, PI[i]-based, PI or IMiDs-based, anti-CD[j]28 mono, anti-CD38 combination, anti-SLAMF7[k], anti-SLAMF7 combination, and other)[b]	• Frequency and history of neutropenia management treatment (history of use of granulocyte-colony stimulating factor, frequency of administration, etc)[b,e] • Frequency and history of antibiotic use (levofloxacin)[b,e] • Frequency and history of prophylactic antibiotics[b,e] • History and frequency of antibiotic use for moderate to severe infections[b,e] • Frequency and history of pneumonia[b,e] • Frequency and history of thrombotic events[b,e] • Frequency and history of aspirin or direct oral anticoagulant use[e]

[a]MM: multiple myeloma.

[b]Variables used to develop machine learning models and understand the characteristics of clusters.

[c]eGFR: estimated glomerular filtration rate.

[d]ADL: activity of daily living.

[e]Variables used only to understand the characteristics of clusters (ie, variables not used to develop machine learning model).

[f]IV: intravenous.

[g]SC: subcutaneous.

[h]IMiDs: immunomodulatory drugs.

[i]PI: proteasome inhibitors.

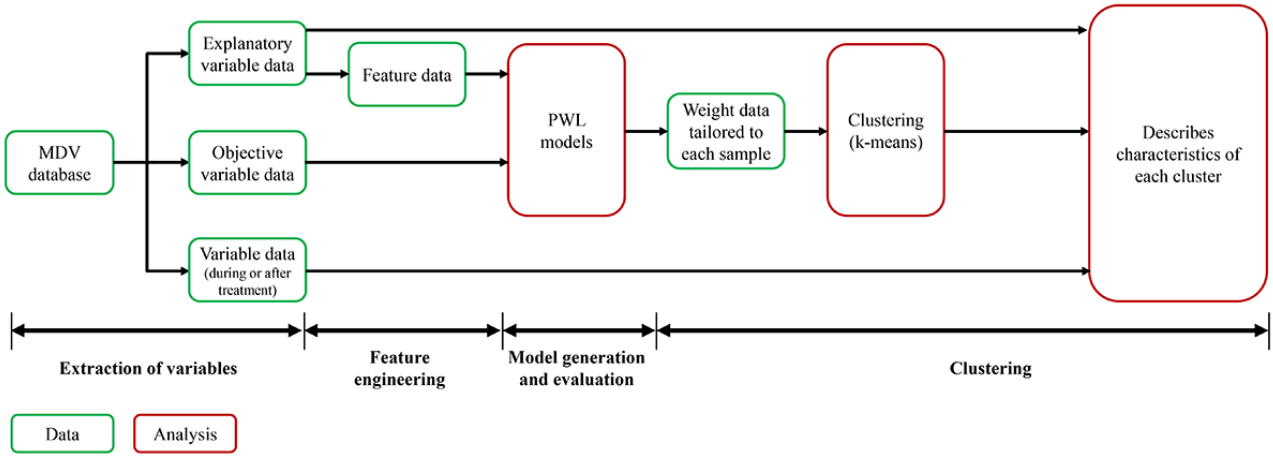
[j]CD: cluster of differentiation.

[k]SLAM7: Signaling Lymphocytic Activation Molecule Family Member 7.

Data Analysis

Data analysis included 5 steps as discussed in Figure 1.

Figure 1. The data analysis pipeline used in the study. MDV: Medical Data Vision; PWL: point-wise linear.



XSL•FO
RenderX

Variable Extraction

ML models were constructed using 647 variables extracted from the MDV database to identify important predictors associated with DoT. The deep learning used in this study was applied to the exploratory research of biomarkers that are associated with the response of the immune checkpoint inhibitors [19,20]. The binary variables that represented DoT within 3, 6, or 12 months were categorized as target variables, while variables used to develop ML models and understand the characteristics of clusters listed in Table 1 were categorized as explanatory variables.

Feature Engineering

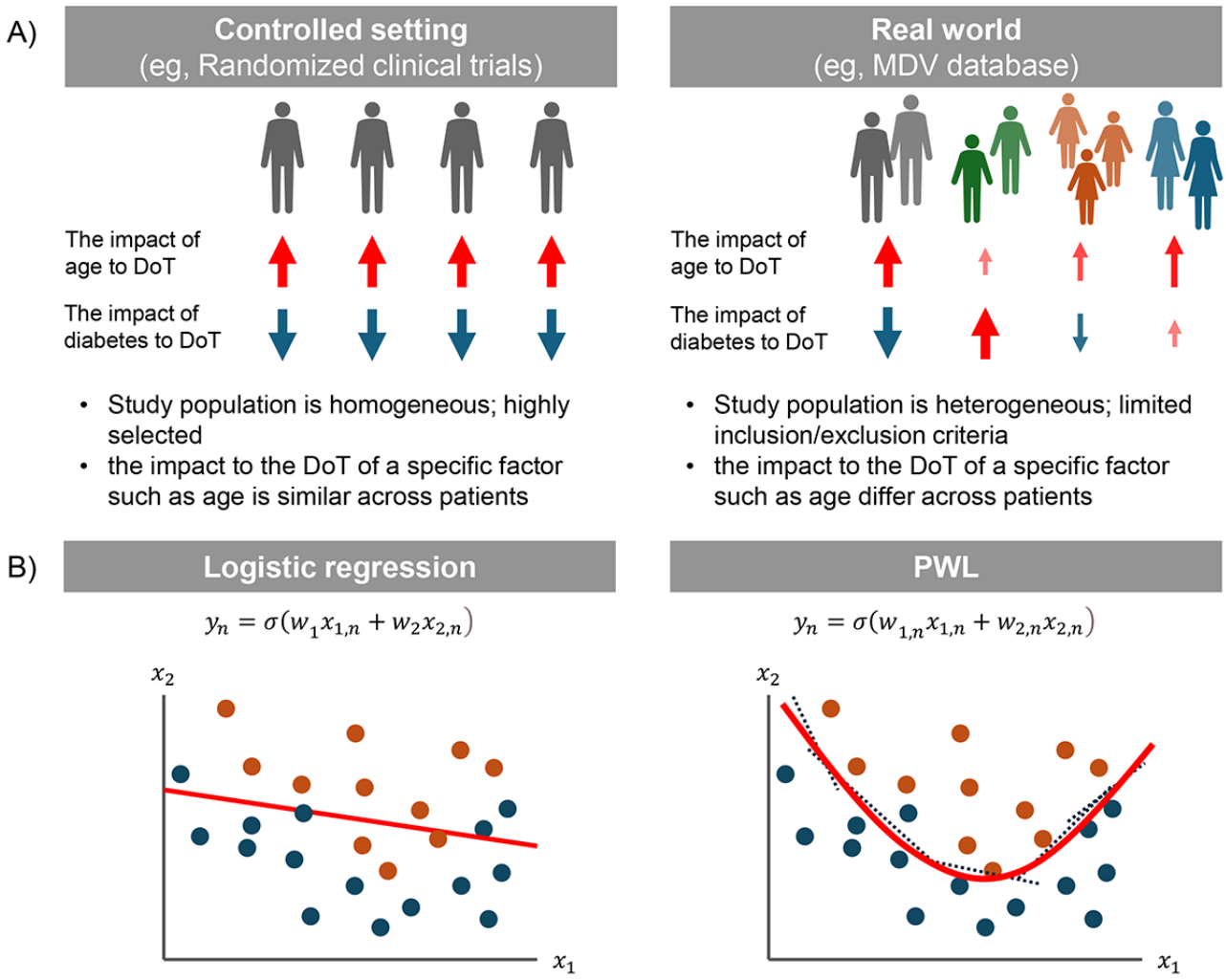
All the feature variables were classified into continuous variables, binary variables, and nominal categorical variables. Continuous variables are normalized to have a mean of zero and an SD of 1 to eliminate bias due to differences in scale between variables. Binary variables were encoded to −1 or 1, and one-hot encoding (1 and −1) was used for nominal categorical variables [21]. The variables with missing values were imputed with 0 because the 0-value input would not change the output in the neural network models [22]. The target variables were set as a binary class and presented in 3 threshold patterns (0 for DoT <3 months, <6 months, or <12 months; and 1 for DoT ≥3 months, ≥6 months, or ≥12 months).

Model Generation and Evaluation

Unlike the randomized controlled trials used in clinical trials, the patients enrolled in the MDV database used in this study are a heterogeneous patient population (Figure 2A). In other words, the impacts of each variable on DoT differ across patients. We used point-wise linear (PWL) models to predict DoT. The PWL model was used as the explainable deep learning model, which generated a custom-made logistic regression (LR) model for each sample by a meta-learning approach (Figure 2B) [23]. Especially, although the LR model has a weight vector $\boldsymbol{w}$ for the original variable vector $\boldsymbol{x}^{(n)}$ ([n] is the sample index), the PWL model has a point-wise weight vector $\boldsymbol{w}^{(n)}$ for the original variable vector $\boldsymbol{x}^{(n)}$ (Figure 2B). On the other hand, extreme gradient boosting (XGBoost) [24,25] is a decision-trees–based machine learning model capable of addressing patients' heterogeneity. Although XGBoost achieves as high a performance as deep learning models do, in general, XGBoost cannot calculate the weight, that is, the impact on the DoT of each variable. In the finalization of the DoT prediction model, the area under the receiver operating characteristic curve (AUROC), recall, precision, and F_1-score were calculated. The prediction performance of the PWL model was compared with the XGBoost (gradient-boosted decision trees) [24,25] and the LR model with a combination of L1 (Lasso) and L2 (Ridge) regularization terms (commonly known as elastic-net). The LR model is similar to PWL in that it explains feature importance by referencing the magnitude of weights (model parameters). The key difference between LR and PWL is that LR assigns a single weight coefficient to each feature across all samples, whereas PWL can assign different weights to the same feature for different samples. Consequently, PWL retains the capability to perform individualized factor analysis for each sample. The prediction performances of these models were calculated by the area under the curve (AUC) evaluated by 10-fold double cross-validation [26]. A "10-fold cross-validation" procedure was used, including 10 repeated attempts after classifying the data into 10 quadrants, with 1 quadrant for evaluation and the remaining 9 for learning. To avoid overlearning in hyperparameter tuning, the learning data were further divided into 10 pieces, 1 of which was used for tuning evaluation. The generalizability of the DoT prediction model was evaluated using the test set of the outer loop of 10-fold double cross-validation, and the datasets were reconstructed or hyperparameters were fine-tuned in case any significant difference (>0.1) was confirmed between the mean AUC in the test set and the training set.

XSL•FO
RenderX

Figure 2. The overview of real-world data and machine learning algorithms. (A) The difference for controlled settings and real world. (B) The comparison of logistic regression model and PWL model. The orange and blue circles represent positive and negative class, respectively. The red lines represent boundaries defined by each model. The dashed lines represent sample-made logistic regression generated by PWL model. DoT: duration of treatment; PWL: point-wise linear model.



Clustering

A clustering analysis of 4848 individual samples was performed based on the similarity of the coefficients of custom-made LR models calculated from the PWL model to understand the number of patterns of the relationship between each sample and DoT (3, 6, and 12 months). All the samples were classified into subgroups using the regression coefficient vector tailored to each sample in the PWL model. Cluster analysis was conducted using the k-means method [27,28], and the number of clusters (k) was determined by the Elbow method [29]. The characteristics of each cluster were analyzed by using all the variables listed in Table 1. The *P* values for continuous variables were assessed by *t* test and 1-way ANOVA, while the *P* values for categorical variables were assessed by chi-square test. To understand the characteristics within a particular cluster, we categorized all 4848 samples into 2 groups—those predicted not to continue treatment after the respective DoT threshold and those predicted to continue treatment after the respective DoT threshold. Continuous variables and categorical variables significantly different between the groups were assessed by 1-way ANOVA and chi-square test, respectively. The *P* values were adjusted by the Bonferroni correction method.

Ethical Considerations

All data used in this study were deidentified and anonymized patient records obtained from the MDV database for secondary research purposes. In accordance with the Ethical Guidelines for Epidemiological Research issued by the Japanese Ministry of Health, Labor, and Welfare, the study did not require informed consent or approval from an Institutional Review Board [30].

Results

Baseline Demographics and Clinical Characteristics

This analysis was conducted in 2762 patients (4848 individual samples) with MM. Of these, 1450 (52.5%) were male; 2079 (75.3%) patients were included during the index year 2016-2021. At the first MM diagnosis, the mean age, Charlson Comorbidity Index (CCI) score, and ADL score were 69.6 (SD 11.8) years, 0.5 (SD 1.1), and 14.6 (SD 7.2), respectively (Table 2).

XSL•FO
RenderX

Table 2. Baseline demographics and clinical characteristics of patients with multiple myeloma.

Characteristics	Overall (N=2762)
Age at first MM[a] diagnosis, mean (SD)	69.6 (11.8)
Age groups at first MM diagnosis (y), n (%)	
18-59	540 (19.6)
60-64	344 (12.5)
≥65	1878 (68)
Sex, n (%)	
Male	1450 (52.5)
Female	1312 (47.5)
Index year, n (%)	
2003-2015	683 (24.7)
2016-2021	2079 (75.3)
CCI[b] score at first MM diagnosis, mean (SD)	0.5 (1.1)
CCI score at first MM diagnosis, n (%)	
0	2121 (76.8)
1	281 (10.2)
2	177 (6.4)
3	99 (3.6)
≥4	84 (3)
ADL[c] score at first MM diagnosis, mean (SD)[d,e]	14.6 (7.2)
ADL score at first MM diagnosis, n (%)[d,e]	
0	111 (8.2)
1	38 (2.8)
2	31 (2.3)
3	18 (1.3)
4-19	466 (33.3)
20	694 (51.1)
Degree of care requirement at diagnosis of MM, mean (SD)[d,f]	0.4 (1.2)
Degree of care requirement at diagnosis of MM, n (%)[d,f]	
0	564 (86.5)
1	24 (3.7)
2	14 (2.1)
3	16 (2.5)
≥4	34 (5.2)
Tube and parenteral nutrition at first MM diagnosis, n (%)[d,g]	
0	214 (92.2)
1	17 (7.3)
2	1 (0.4)
Stage of MM at first MM diagnosis[d,h]	
Mean (SD)	5.0 (0.0)
5, n (%)	3 (100)

XSL•FO
RenderX

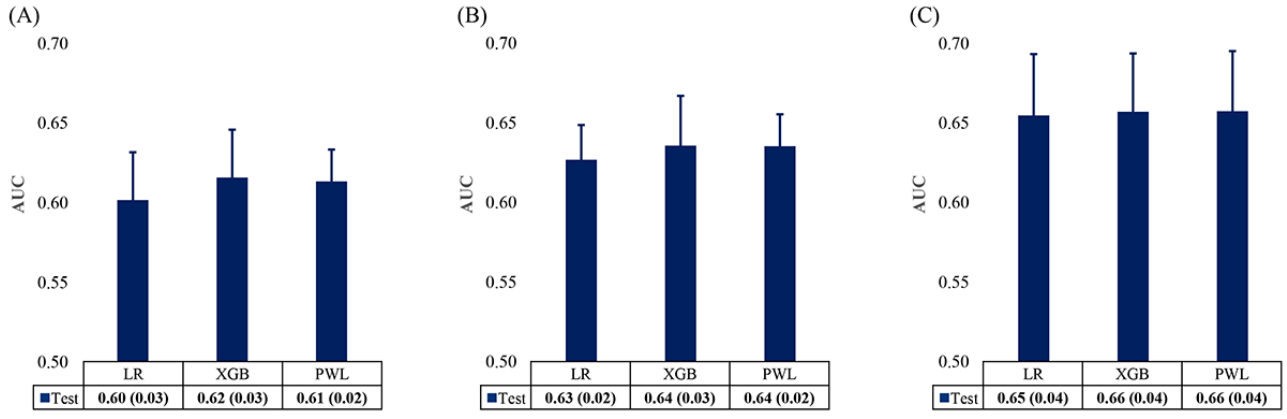
Characteristics	Overall (N=2762)
NYHA at first MM diagnosis, mean (SD)[d,i]	2.1 (1.1)
NYHA at first MM diagnosis, n (%)[d,i]	
0	2 (4.8)
1	15 (35.7)
2	7 (16.7)
3	14 (33.3)
4	4 (9.5)
Cancer at first MM diagnosis, n (%)[d,j]	
Primary	706 (97.8)
Recurrence	16 (2.2)

[a]MM: multiple myeloma.
[b]CCI: Charlson Comorbidity Index.
[c]ADL: activity of daily living.
[d]Missing data were present and these entries were not included in the table.
[e]The number of patients with missing values was 1404.
[f]The number of patients with missing values was 2110.
[g]The number of patients with missing values was 2530.
[h]The number of patients with missing values was 2759.
[i]The number of patients with missing values was 2720.
[j]The number of patients with missing values was 2040.

Prediction Performance

The AUROC score of the PWL model to predict DoT at 3, 6, and 12 months was 0.61, 0.64, and 0.66, respectively. The corresponding values for the XGBoost model were 0.62, 0.64, and 0.66, respectively. The PWL model with the target variable DoT at 12 months performed the best; however, no model reached an AUC>0.7 (Figure 3). Moreover, the slightly smaller SD in PWL suggested more consistent performance across different folds of cross-validation, making it the model of choice for this analysis. The recall, precision, and F_1-score of all models are presented in the supplementary material (Multimedia Appendix 1). The precision and F_1-score were decreased when the DoT length increased, and the thresholds shortened.

Figure 3. Predictive performance of different models to predict duration of therapy (DoT) thresholds. (A) DoT at 3 months, (B) DoT at 6 months, and (C) DoT at 12 months. The graphs represent mean (SD) of area under the curve in 10-fold double cross-validation. LR: regularized logistic regression; PWL: point-wise linear model; XGB: extreme gradient boosting.



Cluster Analysis

Overview

The PWL model generated regression models for each patient, allowing the clustering of patient populations based on the similarity of the coefficients of the regression models. When classified into different clusters, it was found that there were multiple populations with diverse groups of factors affecting the prediction. The samples were categorized into 2 clusters (cluster A and cluster B) at DoT of 3 months and 3 clusters (cluster A, cluster B, and cluster C) at each DoT of 6 months and DoT of 12 months. The baseline demographics and clinical characteristics of 4848 individual samples categorized as per different clusters at DoT of 3, 6, and 12 months are presented

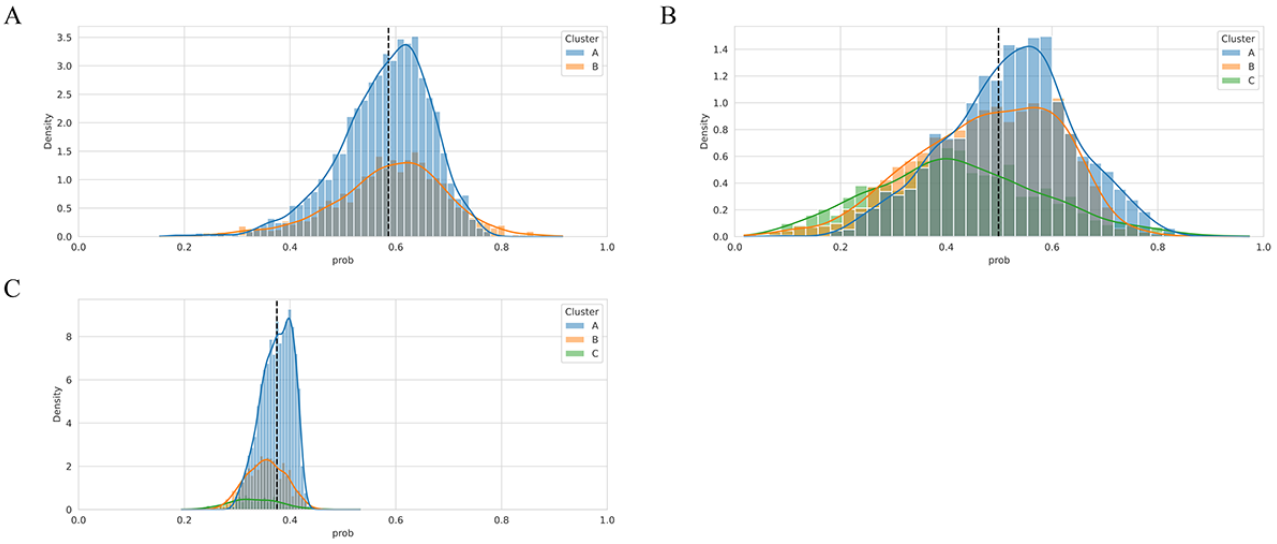
XSL•FO
RenderX

in Multimedia Appendix 2. In addition, we presented the distribution of each cluster in the 2 major components of the variables transformed by principal component analysis in the supplementary materials (Multimedia Appendix 3).

At DoT of 3 Months

Of the 4848 samples, 3287 (67.8%) samples were included in cluster A and 1561 (32.2%) were included in cluster B. Compared with cluster A, cluster B had more patients with age <75 years, male sex, index year 2003-2015, and CCI ≥4 (Multimedia Appendix 2). In the probability analysis, cluster B reported a lower median probability than cluster A, with pointed and extended left tails **(**Figure 4A). From Figure 4A, it can be estimated that many patients in cluster B might be in poor condition, but some were predicted to continue treatment for ≥3 months. Compared with cluster A, patients in cluster B had significantly more lines of previous therapy (0.5 vs 1.6; *P*<.001), higher pretreatment CCI (2.2 vs 2.5; *P*<.001), and a higher level of care requirement at the time of diagnosis of MM (0.4 vs 0.7; *P*=.001). Additionally, cluster B had significantly (*P*<.001) higher proportions of patients with pneumonia, depression, cancer, gastroesophageal reflux disease (GERD), other bowel dysfunctions, primary hypertension, pain, purine and pyrimidine metabolism disorders, and type 2 diabetes mellitus (T2DM) than cluster A (Multimedia Appendix 2).

Figure 4. Detailed cluster analysis. (A) Duration of treatment (DoT) at 3 months, (B) DoT at 6 months, (C) DoT at 12 months. The dashed lines indicate respective classification thresholds (3, 6, and 12 months, respectively, for A, B, and C). The group to the left of the dotted line indicates those predicted not to continue treatment after respective DoT threshold, and the group to the right of the dotted line indicates those predicted to continue treatment after respective DoT thresholds (3, 6, and 12 months, respectively, for A, B, and C).



At DoT of 6 Months

Of the 4848 samples, 2011 (41.5%) were included in cluster A, 1736 (35.8%) were included in cluster B, and 1101 (22.7%) were included in cluster C. Compared with cluster A and B, cluster C had more patients with age <75 years, female sex, the index year 2016-2021, and CCI ≥4 (Multimedia Appendix 2). In the probability analysis, cluster C reported a lower median probability than clusters A and B, with flat and extended right tails (Figure 4B). From Figure 4B, it can be estimated that many patients in cluster C had many comorbidities, but some were predicted to continue treatment for ≥6 months. Patients in cluster B had significantly lower estimated glomerular filtration rate (*P*<.001) and higher neutrophils (*P*<.001), while cluster C had significantly higher pretreatment CCI (*P*<.001) compared to other clusters. Additionally, cluster C had significantly (*P*<.001) higher proportions of patients with pneumonia, depression, cancer, GERD, other bowel dysfunction, primary hypertension, pain, purine and pyrimidine metabolism disorders, and T2DM than clusters A and B (Multimedia Appendix 2).

At DoT of 12 Months

Of the 4848 samples, 3425 (70.6%) were included in cluster A, 1123 (23.2%) were included in cluster B, and 300 (6.2%) were included in cluster C. Compared with cluster A and B, cluster C had more patients with male sex, index year 2016-2021, and CCI ≥4 (Multimedia Appendix 2). In the probability analysis, cluster C reported a lower median probability than cluster A and B, with peaked and extended right tails (Figure 4C). From Figure 4C, it can be estimated that many patients in cluster C were in poor condition, but some were predicted to continue treatment for ≥12 months. Cluster C had significantly higher neutrophils (*P*<.001), leukocytes (*P*<.001), lactate dehydrogenase (*P*<.001), previous treatment lines (*P*<.001), pretreatment CCI (*P*<.001), degree of care requirement at diagnosis of MM (*P*<.001), and pretreatment tube and parenteral nutrition (*P*<.001) among clusters (Multimedia Appendix 2). Additionally, cluster C had a statistically (*P*<.001) higher proportion of patients with comorbidities, including depression, pneumonia, cancer, GERD, other bowel dysfunction, primary hypertension, pain, purine and pyrimidine metabolism disorders, and T2DM, than cluster A and B.

Characteristics of Each Cluster Before and After the DoT Threshold

All 4848 samples were categorized before and after each DoT threshold (3, 6, and 12 months) to determine the factors responsible for shortened DoT. As cluster B at DoT of 3 months and cluster C at DoT of 6 and 12 months were predicted to

XSL•FO
RenderX

continue the treatment, patients from these clusters were further divided into those predicted to continue treatment after the DoT threshold, and those predicted to discontinue treatment within the DoT threshold, and the incidence of events during and after treatment was compared. There were a lesser number of treatments included in DoT <3 months than DoT ≥3 months (n=1859, 38.3% vs n=2989, 61.7%), while DoT <6 (n=2753, 56.8%) and <12 months (n=3672, 75.7%) had a higher number of treatments compared with DoT ≥6 (n=2095, 43.2%) and ≥12 months (n=1176, 24.3%), respectively. In cluster B, a significantly (*P*<.01) higher percentage of patients used immunomodulatory drugs (IMiDs) and used or maintained aspirin during MM therapy in DoT ≥3 than in DoT <3 months group. However, the rate of subsequent transplantation in the next treatment was significantly (*P*<.01) higher in DoT <3 than DoT ≥3 months group. Similarly, in cluster C, a significantly (*P*<.01) higher percentage of patients used IMiDs or aspirin during MM therapy in the DoT ≥6 months than in the DoT <6 months group. In cluster C, a significantly (*P*<.01) higher percentage of patients used IMiDs during MM therapy in DoT ≥12 months than in the DoT <12 months group (Table 3). A summary of the key clinical traits for each cluster after the DoT threshold analysis is presented in Table 4.

Table 3. The characteristics between before and after treatment thresholds, as per the duration of therapy.

Medications	Cluster B			Cluster C			Cluster C		
	<3 months, n/N (%)	≥3 months, n/N (%)	Adjusted *P* value	<6 months, n/N (%)	≥6 months, n/N (%)	Adjusted *P* value	<12 months, n/N (%)	≥12 months, n/N (%)	Adjusted *P* value
Samples	1859/4848 (38.3)	2989/4848 (61.7)	—[a]	2753/4848 (56.8)	2095/4848 (43.2)	—	3672/4848 (75.7)	1176/4848 (24.3)	—
Use of IMiDs[b]	261/714 (36.6)	624/847 (73.7)	<.01	311/763 (40.8)	238/338 (70.4)	<.01	91/240 (37.9)	43/60 (71.7)	<.01
Use of aspirin	206/714 (28.9)	431/847 (50.9)	<.01	244/763 (32.0)	170/338 (50.3)	<.01	69/240 (28.8)	32/60 (53.3)	.77
Continue aspirin	125/714 (17.5)	236/847 (27.9)	<.01	—	—	—	—	—	—
Subsequent transplant	90/714 (12.6)	19/847 (2.2)	<.01	—	—	—	—	—	—

[a]Not applicable.

[b]IMiD: immunomodulatory drug.

Table 4. A summary of the key clinical characteristics distinguishing clusters.

Cluster	DoT[a] threshold	Key characteristics	CCI[b]≥4, n/N (%)	IMiD[c] use (≥DoT), n/N (%)	Aspirin use (≥DoT), n/N (%)
Cluster B	3 months	Higher comorbidity, more prior therapy	393/1561 (25.2)	624/847 (73.7)	431/847 (50.9)
Cluster C	6 months	High comorbidity, more comorbid events	383/1101 (34.8)	238/338 (70.4)	170/338 (50.3)
Cluster C	12 months	High neutrophils, LD[d], leukocytes	127/300 (42.3)	43/60 (71.7)	32/60 (53.3)

[a]DoT: duration of therapy.

[b]CCI: Charlson Comorbidity Index.

[c]IMiD: immunomodulatory drug.

[d]LD: lactate dehydrogenase.

Discussion

Principal Findings

This retrospective real-world observational study using longitudinal patients' data from the nationwide MDV claims database analyzed the predictive factors that shorten DoT in Japanese patients with MM over a 19-year period (2003-2022). Additionally, this analysis used ML models to determine the factors responsible for shorter DoT using information, such as type of treatment and patient characteristics from the MDV claims database, categorizing all the data in different clusters at 3, 6, and 12 months. Finally, a clustering analysis was conducted to investigate the characteristics of the sample in each cluster and the relationship between predicted DoT and the comorbidity management before and after the treatment thresholds (3, 6, and 12 months).

Patients in cluster B than cluster A (at 3 months), and patients in cluster C than cluster A and B (at 6 and 12 months) were supposed to be in poor condition as the number of treatment lines and pretreatment CCI were high; patients had more comorbidities, and elevated laboratory parameters. Cluster B (3-month DoT) and cluster C (6- and 12-month DoT) consistently captured subgroups with high comorbidity burden and elevated pretreatment CCI, yet demonstrated longer

XSL•FO
RenderX

treatment durations. Additionally, the use of IMiDs for MM treatment was significantly higher in patients who met predicted DoT at each threshold versus those who did not. In contrast to the clinical findings, the data from this study indicate that the choice of drug, irrespective of whether it is used for induction or maintenance therapy, affects the expected DoT.

Comparison With Previous Work

From our recent real-world study in Japan, it was found that the DoT decreased continuously with increasing the line of treatment in patients with MM, especially after 2016 when several newer therapy options were in force [12]. In clinical practice, treatment may be discontinued for various reasons; however, for the desired prognosis, it is important to continue the medications for an appropriate duration. The DoT has a strong association with the clinical outcomes and survival in patients with MM [31,32]; therefore, it was inevitable to determine the factors affecting the shortened DoT. Although it was evident that the availability of new drugs might increase the treatment options, making it easier to change treatment regimens, the huge pattern of treatment makes it difficult to identify the cause of shortened DoT. Therefore, this analysis was conducted using the national MDV claims database to determine the predictive factors that could help clinicians in the decision-making process while considering the management of MM. Additionally, this study found the characteristics of patients (age, CCI, comorbidity, treatment lines, laboratory parameters, and care requirement) who had shorter DoT as well as the characteristics of patients with frequent treatment change, which reflects the actual picture of clinical practice. This information is critical to continually improving MM management, achieving optimal patient care, and understanding the burden of disease and health services available to Japanese patients with MM in the real world.

In line with the published evidence [33,34], our study compared the prediction performance of the PWL model with the XGBoost and the LR models. The AUROC values of constructed models were all between 0.6 and 0.7 (Figure 2). Values between 0.6 and 0.7 indicate satisfactory performance [35-37]. The ranges observed in this study are consistent with those reported in previous real-world data–based modeling studies [33,38,39]. Administrative claims data often face inherent data limitations, such as the absence of genomic or staging variables, which can constrain predictive performance. In real-world settings, such models can still provide clinically relevant insights for subgroup identification and hypothesis generation. Nonetheless, models built on such data can still yield clinically relevant insights—particularly for subgroup identification and hypothesis generation in real-world settings. In our study, we generated the models using the variables required in the claims database. For improvements in prediction performance, the use of MM-specific laboratory parameters might be required. Furthermore, in our study, the PWL model showed greater AUROC than the linear regression model at the DoT thresholds of 3, 6, and 12 months. However, these scores were the differences within the error bars between PWL and XGBoost models at each DoT threshold.

DoT Assessment Using PWL Models

DoT predictive models in our study were specifically designed to forecast the DoT in patients with MM, rather than to evaluate clinical effectiveness between specific treatment groups. Due to this focus, these models were not intended for direct application in clinical decision-making or disease diagnosis. Consequently, we did not perform propensity score matching or sensitivity analysis, which are typically used in studies evaluating treatment efficacy. Instead, our efforts were concentrated on carefully developing and comparing the PWL model with the XGBoost and LR models to thoroughly assess the factors influencing DoT. The PWL model differs from LR by generating a separate set of regression coefficients for each individual sample, rather than using a single global set. This approach allows identification of factors that are most relevant for each patient, which may enhance interpretability and personalization in real-world datasets where treatment decisions and outcomes are highly individualized. Unlike randomized controlled trials, the MDV claim database includes a diverse and unselected patient population, without adjustments for specific demographics or conditions. This makes the data more representative of real-world variability. Multiple patterns can thus be anticipated based on the combination of variables that affect DoT. The LR model cannot recognize multiple patterns because the model has only a single regression coefficient to a variable for the whole sample. On the other hand, since the PWL model has a custom-made regression coefficient, it is possible to perform individual analysis for each sample. Therefore, we performed the cluster analysis using the regression coefficient vector tailored to each sample in the PWL model to understand the number of patterns. The PWL model enabled case-level insights, highlighting which patient features most influenced predicted treatment duration. Such interpretability is especially valuable in clinical settings, where understanding the rationale behind risk stratification supports shared decision-making. In our study, the k-means method was used for cluster analysis to classify all the samples into subgroups, and the Elbow method was used to determine the number of clusters. The k-means method is an unsupervised ML approach widely used in research for clustering samples based on similarities [28,40]. Additionally, the k-means method could substantiate heterogeneous longitudinal data into distinct, homogeneous clusters [27,28]. The Elbow method is an established method to determine the appropriate number of clusters that rides the clustering algorithm many times with different values of k (number of clusters) [41], and defines the optimal number of clusters to efficiently categorize the data [42,43].

From the baseline demographics and predictors of cluster assignment, it was found that patients in cluster B at DoT of 3 months and cluster C at DoT of 6 and 12 months, compared with other clusters, had advanced disease and continued the treatment beyond the threshold duration. Interestingly, we observed that in these clusters, more patients had CCI ≥4, ≥1 line of treatment, and comorbidities (pneumonia, depression, cancer, GERD, and other bowel dysfunction) than in other clusters at each DoT threshold. Additionally, the laboratory parameters such as estimated glomerular filtration rate (MDV specific), leukocytes, neutrophils, and lactate dehydrogenase

XSL•FO
RenderX

were significantly higher in the cluster C than the cluster A and B at DoT 12 months. An all-inclusive assessment of comorbidities and overall status (with the help of laboratory parameters) may help to determine a patient's suitability for appropriate therapy as well as the adverse effects that may occur. Altogether, the results from this study support the understanding that the patient's condition and management-related factors may be associated with shortened DoT, consistent with clinical findings [12,44,45]. Evidence suggests that severe disease, older age, and comorbidities are the prognostic factors for poor survival in patients with MM [46,47]. Although it was not in the scope of this analysis to comment on the causal relationship between predictive factors and patient survival, and their effect on longitudinal outcomes and DoT, our findings validate a depiction of shortened DoT in those with reduced disease burden. This is consistent with the findings of the retrospective study conducted by Ailawadhi S et al [32], which reported that each additional month of therapy was associated with a reduced probability of disease progression and death at 2 years. Similar results were also observed in a pooled analysis of 3 phase III trials that evaluated the advantage of novel agent-based continuous therapy versus fixed DoT in patients with newly diagnosed myeloma [48].

To understand the characteristics of the clusters, we analyzed the relationship between the explanatory variables and the flag that indicated the samples belonged to each cluster. Additionally, we calculated the association between the variables during and the flag to investigate the tendency to the event during and after treatment. The significance was detected by the test for noncorrelation at a significance level false discovery rate of <0.05 using the Benjamini-Hochberg method. The Benjamini-Hochberg method is a validated tool used in multiple correlation analyses that could decrease the false discovery rate [19].

Significantly, more patients used IMiDs and aspirin beyond the treatment thresholds (both ≥3 [cluster B] and ≥6 months [cluster C]) compared with before treatment thresholds (<3 and <6 months), and in cluster C at DoT ≥12 months, a higher proportion of patients used IMiDs during MM therapy compared with DoT <12 months group. The use of IMiDs was significantly higher when comparing sample populations predicted as DoT of <3, <6, and <12 months with those who did not. However, this observational association may reflect confounding factors, such as patient eligibility, tolerability, or physician preference, and does not imply causation. In patients with MM, lenalidomide, one of the commonly used IMiDs, is sometimes used in maintenance therapy, and the results from this study suggest that IMiD therapies are associated with treatment continuation for MM in patients with high CCI, severe disease, and comorbidities. Furthermore, these results imply that the use of appropriate IMiD may impact DoT in patients with MM. Since the treatment choices for patients with MM depend on several factors, including the type of previous therapy and comorbidity, these data may assist the clinical decision-making on whether IMiDs should be used in these patients. Similar to IMiD use, the management of thromboprophylaxis was also associated with treatment continuation for MM in patients with high CCI.

Factors Identified for Shortened DoT

Our analysis identified several key factors that might be associated with shortened DoT for patients with MM. The first factor is the patient's disease status—patients with poorer health status (CCI ≥4, ≥1 previous line of treatment, and comorbidities including pneumonia, depression, cancer, GERD, or bowel dysfunction), along with elevated laboratory parameters, showed longer DoT. Second was the management-related factors, including use of IMiDs and management of thromboprophylaxis: patients who were treated with IMiDs for MM showed extended DoT.

Strengths and Limitations

The strengths of the study include (1) this study is a precedent of data-driven research to understand the characteristics of patients with MM whose DoT was shortened; (2) by using AI, based on ML techniques, we were able to build a prediction model using many factors that are difficult to build with a conventional statistical model; (3) in this study, subgrouping was automatically done by the ML-based DoT predictive models as subgrouping analysis is required when analyzing real-world data due to heterogeneity of the patient population; (4) this study used the latest data from nationwide MDV claims database which was based on Japanese general population, and represented real-world clinical practice.

This study also has certain limitations. First, the database is restricted to the Japanese population; hence, caution should be taken while generalizing the results. This database only included the data of the insured population, and it was impossible to follow patients outside of the hospital or those who switched hospitals. Not all laboratory tests were conducted in the entire patient population, which again restricts the generalizability of the findings. Furthermore, since the cause of death was unavailable, it could not be determined whether the shorter DoT was due to treatment failure, toxicity, patient preference, or unrelated mortality. Additionally, death was captured only for hospitalized patients, and important prognostic factors for MM, such as cytogenetic profile, R-International Staging System stage, and presence of all plasmacytomas, were not collected. This may affect the clinical granularity and predictive robustness of the models. Their inclusion may further enhance both robustness and interpretability of future predictive tools. We partially addressed this by incorporating surrogate measures, such as CCI and ADL scores. Furthermore, it is difficult to evaluate the causal relationship between predictive factors and DoT, even if the relationship between the factors and shortening of DoT is assumed, and the presence of unknown confounding that affects the continuation of treatment cannot be ruled out.

Conclusion

This study is the first to apply ML to assess predictive factors for shortened DoT in Japanese patients with MM. Using the PWL model, an explainable deep learning method, we identified key patient and management factors influencing DoT. These insights can contribute to improving patient management in hospitals, optimizing care, and enhancing understanding of the burden of disease and health services for Japanese patients with MM in real-world settings. Further statistical analysis is needed

XSL•FO
RenderX

to refine patient segmentation and ensure the effective application of these insights in clinical environments.

Acknowledgments

The authors thank Rabi Panigrahy (PhD), Aafrin Khan, and Jyotsana Dixit (PhD) for writing assistance and Himabindu Gutha, CMPP, for additional editorial support (all from SIRO Clinpharm Pvt Ltd) for the development of the manuscript.

Data Availability

The data that support the findings of this study were derived from the MDV database. Restrictions apply to the availability of these data, which were used under license for this study.

Funding

This study was funded by Janssen Pharmaceutical K.K., Japan, and Johnson & Johnson.

Authors' Contributions

Conceptualization: HH, TI, SO, SI, KW, CS, KK, DBCW, DY
Data curation: SN, YY, TS
Formal analysis: KW, SN, YY, TS
Funding acquisition: KW
Methodology: HH, TI, SO, SI, KW, CS, KK, DBCW, DY, SN, YY, TS
Software: SN, YY, TS
Validation: HH, TI, SO, SI, KW, CS, KK, DBCW, DY, SN, YY, TS
Writing – original draft: KW
Writing – review and editing: HH, TI, SO, SI, KW, CS, KK, DBCW, DY, SN, YY, TS

Conflicts of Interest

HH has received consulting fees and honoraria from Johnson & Johnson, Ono, and Takeda; grants from Takeda, Kyowa Kirin, and BMS. TI has received consulting fees and honoraria from Ono, Takeda, BMS, Johnson & Johnson, and Sanofi; research fundings support from Takeda, Janssen, Pfizer, and BMS. SO has nothing to declare. SI has received consulting fees and honoraria from BMS, Johnson & Johnson, Takeda, Ono, Sanofi, AstraZeneca, and Pfizer; research funding support from BMS, Johnson & Johnson, Daiichi Sankyo, Amgen, Ono, AbbVie GK, GSK, Eli Lilly, Caelum Biosciences, Inc., Pfizer, Takeda, and Sanofi; grants from Chugai, Takeda, Ono, and Sanofi. KW and CS are employees of Johnson & Johnson. KK was an employee of Johnson & Johnson when the study was planned and is currently employed by Gilead Sciences K.K. since September 2022 and have share in Gilead Science. DBCW and DY are employees of Johnson & Johnson. SN is an employee of Hitachi Ltd. YY and TS were employed by Hitachi Ltd. during the time this research was conducted.

Multimedia Appendix 1
Table S1: The precision, recall and F1 score of all models.
[DOCX File , 15 KB - cancer_v12i1e75586_app1.docx]

Multimedia Appendix 2
Table S2: Details of characteristics of patients with MM treatment stratified by clusters at DoT 3, 6 and 12 months. Missing values from total samples: a4428, b4217, c4221, d4420, e4254, f3624. CCI, Charlson Comorbidity Index; DoC, degree of care; DoT, duration of treatment; eGFR, estimated glomerular filtration rate; GERD, gastro esophageal reflux disease; HTN, hypertension; LoT, line of treatment; MDV, Medical Data Vision; MM, multiple myeloma; T2DM, type 2 diabetes mellitus.
[DOCX File , 29 KB - cancer_v12i1e75586_app2.docx]

Multimedia Appendix 3
Two-dimensional cluster analysis using PCA at DoT of 3, 6 and 12 months Dimensional compression was performed by PCA to confirm the relationship between clusters in two-dimensional plots. Each result was drawn to correspond to a previously defined cluster. DoT, duration of treatment; PCA, principal component analysis.
[PNG File , 286 KB - cancer_v12i1e75586_app3.png]

References

1. Fazio M, Del Fabro V, Parrinello N, Allegra A, Markovic U, Botta C, et al. Multiple myeloma in 2023 ways: from trials to real life. Curr Oncol 2023;30(11):9710-9733 [FREE Full text] [doi: 10.3390/curroncol30110705] [Medline: 37999125]

XSL•FO
RenderX

2. World HO(GCOCT(. Global Cancer Observatory: Cancer Today (Japan). World Health Organization. URL: https://gco.iarc.who.int/media/globocan/factsheets/populations/392-japan-fact-sheet.pdf [accessed 2024-03-05]
3. Suzuki N, Okuyama T, Akechi T, Kusumoto S, Ri M, Inagaki A, et al. Symptoms and health-related quality of life in patients with newly diagnosed multiple myeloma: a multicenter prospective cohort study. Jpn J Clin Oncol 2022;52(2):163-169. [doi: 10.1093/jjco/hyab178] [Medline: 34865077]
4. Cancer R. Cancer Statistics in Japan—2023. Foundation for Promotion of Cancer Research. URL: https://ganjoho.jp/public/qa_links/report/statistics/pdf/cancer_statistics_2023.pdf [accessed 2024-03-05]
5. Cancer R. Cancer Statistics in Japan—2018. Foundation for Promotion of Cancer Research. URL: https://ganjoho.jp/public/qa_links/report/statistics/pdf/cancer_statistics_2018.pdf [accessed 2024-03-05]
6. Costello C. Newly diagnosed multiple myeloma: making sense of the menu. Hematology Am Soc Hematol Educ Program 2022;2022(1):539-550 [FREE Full text] [doi: 10.1182/hematology.2022000404] [Medline: 36485145]
7. Ogawa Y, Tobinai K, Ogura M, Ando K, Tsuchiya T, Kobayashi Y, et al. Phase I and II pharmacokinetic and pharmacodynamic study of the proteasome inhibitor bortezomib in Japanese patients with relapsed or refractory multiple myeloma. Cancer Sci 2008;99(1):140-144 [FREE Full text] [doi: 10.1111/j.1349-7006.2007.00638.x] [Medline: 17970782]
8. Uno S, Motegi Y, Minehata K, Aoki Y. Lenalidomide treatment of Japanese patients with myelodysplastic syndromes with 5q deletion: a post-marketing surveillance study. Int J Hematol 2023;118(4):432-442 [FREE Full text] [doi: 10.1007/s12185-023-03634-7] [Medline: 37493867]
9. Matsue K, Iwasaki H, Chou T, Tobinai K, Sunami K, Ogawa Y, et al. Pomalidomide alone or in combination with dexamethasone in Japanese patients with refractory or relapsed and refractory multiple myeloma. Cancer Sci 2015;106(11):1561-1567 [FREE Full text] [doi: 10.1111/cas.12772] [Medline: 26292221]
10. Murakami H, Shimizu K, Sawamura M, Suzuki K, Sugiura I, Kosugi H, et al. Phase II and pharmacokinetic study of thalidomide in Japanese patients with relapsed/refractory multiple myeloma. Int J Hematol 2009;89(5):636-641. [doi: 10.1007/s12185-009-0314-5] [Medline: 19399582]
11. Uno S, Midorikawa S, Inoue K, Ichikawa D, Ito T, Kuroda J, et al. Survival outcomes among patients with multiple myeloma in the era of novel agents: exploratory assessment using an electronic medical record database in Japan. PLoS One 2023;18(5):e0285947 [FREE Full text] [doi: 10.1371/journal.pone.0285947] [Medline: 37256880]
12. Handa H, Ishida T, Ozaki S, Mori A, Kato K, Iida S. Treatment pattern and clinical outcomes in multiple myeloma patients in Japan using the Medical Data Vision claims database. PLoS One 2023;18(4):e0283931 [FREE Full text] [doi: 10.1371/journal.pone.0283931] [Medline: 37023056]
13. Lee HC, Ailawadhi S, Gasparetto CJ, Jagannath S, Rifkin RM, Durie BGM, Connect® MM Registry. Treatment patterns and outcomes in elderly patients with newly diagnosed multiple myeloma: results from the Connect MM Registry. Blood Cancer J 2021;11(7):134 [FREE Full text] [doi: 10.1038/s41408-021-00524-1] [Medline: 34301916]
14. Shin J, Lee J, Kim K, Min C, Lee J, Suh C, Korean Multiple Myeloma 151 Investigators. Venous thromboembolism in relapsed or refractory multiple myeloma patients treated with lenalidomide plus dexamethasone. Int J Hematol 2019;109(1):79-90. [doi: 10.1007/s12185-018-2540-1] [Medline: 30291558]
15. Palumbo A, Bringhen S, Mateos M, Larocca A, Facon T, Kumar S. Erratum: Palumbo A, Bringhen S, Mateos M-V, et al. Geriatric assessment predicts survival and toxicities in elderly myeloma patients: an International Myeloma Working Group report. . 2015;125(13):2068-2074. Blood 2016;128(7):1020 [FREE Full text] [doi: 10.1182/blood-2016-07-727081] [Medline: 31265505]
16. Allegra A, Tonacci A, Sciaccotta R, Genovese S, Musolino C, Pioggia G, et al. Machine learning and deep learning applications in multiple myeloma diagnosis, prognosis, and treatment selection. Cancers (Basel) 2022;14(3):606 [FREE Full text] [doi: 10.3390/cancers14030606] [Medline: 35158874]
17. Mosquera Orgueira A, González Pérez MS, Diaz Arias J, Rosiñol L, Oriol A, Teruel AI, PETHEMA/GEM Cooperative Group. Unsupervised machine learning improves risk stratification in newly diagnosed multiple myeloma: an analysis of the Spanish Myeloma Group. Blood Cancer J 2022;12(4):76 [FREE Full text] [doi: 10.1038/s41408-022-00647-z] [Medline: 35468898]
18. Williams L, Khosravi B, Velimirovic M, Khouri J, Raza S, Mazzoni S. An ensemble machine learning model using gradient boosting identifies patients with disease progression in newly diagnosed multiple myeloma. Blood. 2023 2023;142(1):1990. [doi: 10.1182/blood-2023-188762]
19. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res 2001;125(1-2):279-284. [doi: 10.1016/s0166-4328(01)00297-2] [Medline: 11682119]
20. Kumagai S, Togashi Y, Kamada T, Sugiyama E, Nishinakamura H, Takeuchi Y, et al. The PD-1 expression balance between effector and regulatory T cells predicts the clinical efficacy of PD-1 blockade therapies. Nat Immunol 2020;21(11):1346-1358. [doi: 10.1038/s41590-020-0769-3] [Medline: 32868929]
21. Samuels J. Samuels J. One-Hot Encoding and Two-Hot Encoding. Lake Rumare: Imperial; 2024:2024.
22. Hirakawa Y, Yoshioka K, Kojima K, Yamashita Y, Shibahara T, Wada T, et al. Potential progression biomarkers of diabetic kidney disease determined using comprehensive machine learning analysis of non-targeted metabolomics. Sci Rep 2022 Sep 29;12(1). [doi: 10.1038/s41598-022-20638-1] [Medline: 36175470]

XSL•FO
RenderX

23. Shibahara T, Wada C, Yamashita Y, Fujita K, Sato M, Kuwata J, et al. Deep learning generates custom-made logistic regression models for explaining how breast cancer subtypes are classified. PLoS One 2023;18(5):e0286072 [FREE Full text] [doi: 10.1371/journal.pone.0286072] [Medline: 37216350]
24. Boldini D, Grisoni F, Kuhn D, Friedrich L, Sieber SA. Practical guidelines for the use of gradient boosting for molecular property prediction. J Cheminform 2023;15(1):73 [FREE Full text] [doi: 10.1186/s13321-023-00743-7] [Medline: 37641120]
25. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. : Association for Computing Machinery; 2016 Presented at: 22nd ACM SIGKDD International Conference; 2026 Aug 16; San Francisco, California, USA. [doi: 10.1145/2939672.2939785]
26. Filzmoser P, Liebmann B, Varmuza K. Repeated double cross validation. Journal of Chemometrics 2009;23(4):160-171. [doi: 10.1002/cem.1225]
27. Genolini C, Pingault J, Driss T, Côté S, Tremblay R, Vitaro F, et al. KmL3D: a non-parametric algorithm for clustering joint trajectories. Comput Methods Programs Biomed 2013;109(1):104-111. [doi: 10.1016/j.cmpb.2012.08.016] [Medline: 23127283]
28. Ko J, Furby H, Ma X, Long JD, Lu X, Slowiejko D, et al. Clustering and prediction of disease progression trajectories in Huntington's disease: An analysis of Enroll-HD data using a machine learning approach. Front Neurol 2022;13:1034269 [FREE Full text] [doi: 10.3389/fneur.2022.1034269] [Medline: 36793800]
29. Madhulatha T. An overview on clustering methods. arXiv 2012 [FREE Full text] [doi: 10.48550/arXiv.1205.1117]
30. Ministry of Education, Culture, Sports, Science and Technology, Ministry of Health, Labour and Welfare, Ministry of Economy, Trade and Industry. Ethical Guidelines for Medical and Biological Research Involving Human Subjects. Tokyo (JP). Ministry of Education, Culture, Sports, Science and Technology. URL: https://www.mext.go.jp/content/20250325-mxt_life-000035486-01.pdf [accessed 2026-02-12]
31. Hari P, Romanus D, Palumbo A, Luptakova K, Rifkin RM, Tran LM, et al. Prolonged duration of therapy is associated with improved survival in patients treated for relapsed/refractory multiple myeloma in routine clinical care in the United States. Clin Lymphoma Myeloma Leuk 2018;18(2):152-160 [FREE Full text] [doi: 10.1016/j.clml.2017.12.012] [Medline: 29395837]
32. Ailawadhi S, Ogbonnaya A, Murty S, Cherepanov D, Schroader BK, Romanus D, et al. Duration of frontline therapy and impact on clinical outcomes in newly diagnosed multiple myeloma patients not receiving frontline stem cell transplant. Cancer Med 2023;12(3):3145-3159 [FREE Full text] [doi: 10.1002/cam4.5239] [Medline: 36151787]
33. Boxley C, Fujimoto M, Ratwani RM, Fong A. A text mining approach to categorize patient safety event reports by medication error type. Sci Rep 2023 Oct 26;13(1):18354 [FREE Full text] [doi: 10.1038/s41598-023-45152-w] [Medline: 37884577]
34. Han J, Kim SY, Lee J, Lee WH. Brain age prediction: a comparison between machine learning models using brain morphometric data. Sensors (Basel) 2022 Oct 21;22(20):8077 [FREE Full text] [doi: 10.3390/s22208077] [Medline: 36298428]
35. Divyanshu D, Aggarwal D, Agarwal G, Agarwal G, Shahid M. Heart attack predication using machince learning. Journal of Propulsion Technology 2024;45(2):908-918 [FREE Full text]
36. Yamana H, Okada A, Ono S, Michihata N, Jo T, Yasunaga H. Machine learning-based models for outpatient prescription of Kampo formulations: an analysis of a health insurance claims database. J Epidemiol 2024;34(1):8-15 [FREE Full text] [doi: 10.2188/jea.JE20220089] [Medline: 36642513]
37. Golas SB, Shibahara T, Agboola S, Otaki H, Sato J, Nakae T, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. BMC Med Inform Decis Mak 2018;18(1):44 [FREE Full text] [doi: 10.1186/s12911-018-0620-z] [Medline: 29929496]
38. Fitri A, Sjahrir H, Bachtiar A, Ichwan M, Fitri FI, Rambe AS. Predictive model of diabetic polyneuropathy severity based on vitamin D level. Open Access Maced J Med Sci 2019;7(16):2626-2629 [FREE Full text] [doi: 10.3889/oamjms.2019.454] [Medline: 31777620]
39. Hurria A, Mohile S, Gajra A, Klepin H, Muss H, Chapman A, et al. Validation of a prediction tool for chemotherapy toxicity in older adults with cancer. J Clin Oncol 2016;34(20):2366-2371 [FREE Full text] [doi: 10.1200/JCO.2015.65.4327] [Medline: 27185838]
40. Nicholson C, Beattie L, Beattie M, Razzaghi T, Chen S. A machine learning and clustering-based approach for county-level COVID-19 analysis. PLoS One 2022;17(4):e0267558 [FREE Full text] [doi: 10.1371/journal.pone.0267558] [Medline: 35476849]
41. Marutho D, Handaka SH, Wijaya E, Muljono. The determination of cluster number at k-mean using elbow method and purity evaluation on headline news. 2018 Presented at: 2018 International Seminar on Application for Technology of Information and Communication; 2018 Sep 21-22; Semarang, Indonesia URL: https://ieeexplore.ieee.org/document/8549751 [doi: 10.1109/isemantic.2018.8549751]
42. Humaira H, Rasyidah R. Determining the appropiate cluster number using elbow method for k-means algorithm. 2024 Jan 24 Presented at: Determining the appropiate cluster number using elbow method for k-means algorithm. Proceedings of the 2nd Workshop on Multidisciplinary and Applications (WMA) , January , Padang, Indonesia; 2020; 2018; Padang, Indonesia. [doi: 10.4108/eai.24-1-2018.2292388]

XSL•FO
RenderX

43. Hassan BA, Rashid TA, Hamarashid HK. A novel cluster detection of COVID-19 patients and medical disease conditions using improved evolutionary clustering algorithm star. Comput Biol Med 2021;138:104866 [FREE Full text] [doi: 10.1016/j.compbiomed.2021.104866] [Medline: 34598065]
44. Terpos E, Mikhael J, Hajek R, Chari A, Zweegman S, Lee HC, et al. Management of patients with multiple myeloma beyond the clinical-trial setting: understanding the balance between efficacy, safety and tolerability, and quality of life. Blood Cancer J 2021;11(2):40 [FREE Full text] [doi: 10.1038/s41408-021-00432-4] [Medline: 33602913]
45. Handa H, Ishida T, Ozaki S, Iida S, Mori A, Sakai C. Applying artificial intelligence machine learning to discover factors that shorten duration of treatment in patients with multiple myeloma using real-world patient-level longitudinal data from Medical Data Vision database. Blood. 2023 2023:142. [doi: 10.1182/blood-2023-172849]
46. Kleber M, Ihorst G, Terhorst M, Koch B, Deschler B, Wäsch R, et al. Comorbidity as a prognostic variable in multiple myeloma: comparative evaluation of common comorbidity scores and use of a novel MM-comorbidity score. Blood Cancer J 2011;1(9):e35 [FREE Full text] [doi: 10.1038/bcj.2011.34] [Medline: 22829196]
47. Cordas Dos Santos DM, Saliba RM, Patel R, Bashir Q, Saini N, Hosing C, et al. Age is a prognostic factor for the overall survival of patients with multiple myeloma undergoing upfront autologous hematopoietic stem cell transplantation. Biol Blood Marrow Transplant 2020;26(6):1077-1083 [FREE Full text] [doi: 10.1016/j.bbmt.2019.11.028] [Medline: 31786242]
48. Palumbo A, Gay F, Cavallo F, Di Raimondo F, Larocca A, Hardan I, et al. Continuous therapy versus fixed duration of therapy in patients with newly diagnosed multiple myeloma. J Clin Oncol 2015;33(30):3459-3466 [FREE Full text] [doi: 10.1200/JCO.2014.60.2466] [Medline: 26282661]

Abbreviations

ADL: activity of daily living
AI: artificial intelligence
AUC: area under the curve
AUROC: area under the receiver operating characteristic curve
CCI: Charlson Comorbidity Index
DoT: duration of therapy
ICD-10: International Statistical Classification of Diseases, Tenth Revision
GERD: gastroesophageal reflux disease
IMiD: immunomodulatory drugs
LR: logistic regression
MDV: Medical Data Vision
ML: machine learning
MM: multiple myeloma
NDMM: newly diagnosed multiple myeloma
PWL: point-wise linear
RRMM: relapsed or refractory multiple myeloma
T2DM: type 2 diabetes mellitus
XGBoost: extreme gradient boosting

Edited by N Cahill; submitted 25.Apr.2025; peer-reviewed by C Madu, RT Potla; comments to author 22.May.2025; accepted 19.Oct.2025; published 19.Feb.2026.

Please cite as:
Handa H, Ishida T, Ozaki S, Iida S, Wattanakamolkul K, Sakai C, Kato K, Bin-Chia Wu D, Yu D, Nemoto S, Yamashita Y, Shibahara T
Assessment of Predictive Factors That Shorten Duration of Treatment in Patients With Multiple Myeloma Using AI: Real-World Longitudinal Study Using Data From Medical Data Vision Claims Database
JMIR Cancer 2026;12:e75586
URL: https://cancer.jmir.org/2026/1/e75586
doi:10.2196/75586
PMID:

©Hiroshi Handa, Tadao Ishida, Shuji Ozaki, Shinsuke Iida, Kittima Wattanakamolkul, Chika Sakai, Kenichi Kato, David Bin-Chia Wu, DaeYoung Yu, Shota Nemoto, Yasuho Yamashita, Takuma Shibahara. Originally published in JMIR Cancer (https://cancer.jmir.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction

XSL•FO
RenderX

in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Practical Recommendations for Implementing Remote Patient Monitoring for Patients Receiving Bispecific Antibodies for Multiple Myeloma

Cindy Varga[1], MD; Eric Schaefer[2], MD

[1]Department of Hematologic Oncology and Blood Disorders, Atrium Health Levine Cancer Institute, 1021 Morehead Medical Drive, Charlotte, NC, United States

[2]Department of Medical Oncology, Highlands Oncology Group, Fayetteville, AR, United States

Corresponding Author:
Cindy Varga, MD
Department of Hematologic Oncology and Blood Disorders, Atrium Health Levine Cancer Institute, 1021 Morehead Medical Drive, Charlotte, NC, United States

Abstract

Bispecific antibodies are novel immunotherapies approved for the treatment of patients with relapsed or refractory multiple myeloma (MM). While these therapies induce deep and durable responses, they are associated with adverse events, such as cytokine release syndrome and immune effector cell–associated neurotoxicity syndrome, that require specific protocols for monitoring, prevention, and treatment. Due to the potential risk of cytokine release syndrome and immune effector cell–associated neurotoxicity syndrome, current US labeling recommends inpatient administration during step-up dosing of bispecific antibodies, which typically occurs in academic centers and can often restrict the number of patients that can be treated. The administration of bispecific antibodies in the outpatient and community setting can enhance the patient experience and reduce resource use, which is a barrier to the widespread uptake of these treatments. With bispecific antibodies being evaluated in earlier lines of therapy for MM, the broader use of these agents will underscore the need for more accessible and less resource-intensive options that ensure safe outpatient administration. In this Viewpoint, we aim to provide practical recommendations built from our own real-world perspectives and case examples for institutions seeking to implement remote patient monitoring (RPM) strategies to support the outpatient dosing of bispecific antibodies in patients with MM. We discuss how the use of RPM has the potential to improve adverse event detection and management, improve patient outcomes, and decrease resource use and costs. While specific approaches will vary based on the needs of each treatment center, our experiences suggest that RPM implementation can be safe and feasible for both academic and community practices.

(JMIR Cancer 2026;12:e95949) doi:10.2196/95949

KEYWORDS

remote patient monitoring; bispecific antibodies; multiple myeloma; immunotherapy; outpatient monitoring

Introduction

Bispecific antibodies (teclistamab, talquetamab, linvoseltamab, and elranatamab) are novel immunotherapy agents approved for the treatment of relapsed or refractory multiple myeloma (MM) [1-8]. In heavily pretreated patients, bispecific antibodies induce deep and durable responses with predictable safety profiles [9]. In addition to common hematologic and nonhematologic adverse events (AEs) typically associated with MM therapies, patients treated with bispecific antibodies are at increased risk of cytokine release syndrome (CRS), immune effector cell–associated neurotoxicity syndrome (ICANS), and infections, all of which require patients to be closely monitored and managed [10-13]. AEs such as neutropenia and hypogammaglobulinemia contribute to the increased infection risk in patients receiving bispecific antibodies [14]. Recommended strategies to mitigate infection risk include vaccinations; the use of antibacterial, antifungal, and antiviral prophylaxis; and intravenous immunoglobulin supplementation regardless of immunoglobulin G blood levels [14,15]. In addition, reducing the frequency of bispecific antibody dosing in responding patients may help reduce the risk of infection [16]. CRS and ICANS events tend to occur during step-up dosing (SUD) of bispecific antibodies or following the first full treatment dose [9]. Therefore, bispecific antibodies are generally administered in hospitals and academic medical centers, with US labeling suggesting hospitalization for 24 to 48 hours following SUD to monitor for signs and symptoms of CRS and ICANS [4-6,8]. Prophylactic tocilizumab administered before the first dose [17,18] and pretreatment with dexamethasone, acetaminophen, and diphenhydramine before each SUD and the first full treatment dose [4-6,8] can help reduce CRS risk.

Inpatient treatment administration and subsequent monitoring for AEs can be resource intensive for hospital teams, limiting

XSL•FO
RenderX

the number of patients who can receive treatment with bispecific antibodies at any one time. Academic medical centers may be able to better handle the resource burden and costs of administering and monitoring treatment than community institutions, yet not all patients have access to academic centers [19]. Implementing bispecific antibody treatment in other practice settings will help expand patient access, reduce inpatient hospitalizations, and improve the patient experience [20,21]. These novel immunotherapies are being evaluated in earlier lines of therapy [9], with the National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology updating their recommendations in 2026 to include the use of teclistamab in combination with daratumumab for treatment of MM after one prior line of therapy including lenalidomide and a proteasome inhibitor [22]. Strategies that support the safe and effective administration of bispecific antibodies in outpatient clinics and community institutions are needed to facilitate the widespread use of these treatments in a broader population of patients with MM.

Accumulating evidence shows that bispecific antibodies can be safely administered in the outpatient setting. Outpatient vs inpatient monitoring for tarlatamab, a bispecific antibody approved for the treatment of small cell lung cancer, was directly compared in a phase 1 trial [23]. The safety profile of patients who received tarlatamab in the outpatient setting was similar to that observed with inpatient dosing, with no increase in the risk or severity of CRS or ICANS, and most CRS events were managed without hospitalization or an emergency department visit. In patients with MM, several real-world studies have shown outpatient administration of bispecific antibodies to be feasible, safe, and potentially cost-effective with appropriate prophylaxis and patient management [24-29].

Administration of bispecific antibodies in outpatient and community settings is feasible when supported by an appropriate infrastructure with careful planning, clear protocols, and a well-trained multidisciplinary team [21,30-32]. It is important that patients and their caregivers receive comprehensive support and education, including information on how to recognize the signs and symptoms of CRS and ICANS [21]. A practice-based road map has been developed for the outpatient administration of bispecific antibodies in patients with MM, with recommendations for infrastructure setup and logistics, including bed capacity and resource allocation; multidisciplinary team composition and training; patient selection; patient and caregiver education; SUD procedures; and AE prophylaxis, monitoring, and treatment [21]. The guidelines can be adapted to various health care environments considering available resources, infrastructure, and patient populations for each center.

Remote patient monitoring (RPM) is an example of one strategy being used to implement outpatient dosing programs [30]. RPM uses noninvasive devices (eg, wearables), electronic patient-reported outcomes (ePROs) manually entered into a digital device (eg, tablet or smartphone), and/or invasive devices (eg, cardiac monitors) to obtain patient physiological data and other information about their condition outside the traditional health care environment [33]. The information captured allows the care team to monitor conditions and escalate care as needed if a patient experiences an AE [34]. The use of RPM for patients receiving bispecific antibodies in the outpatient setting has the potential to expand patient access to these treatments, improve AE detection and management, decrease inpatient hospitalization time, streamline clinician workload (eg, by reducing the need for phone check-ins), lower overall health care costs, and improve patient outcomes [35].

Emerging evidence is demonstrating the benefit of RPM in patients with MM. For example, a small academic study evaluated the effectiveness of RPM wearable devices compared with standard-of-care monitoring to detect early CRS in patients with MM receiving bispecific antibodies and found that RPM was more sensitive and led to faster CRS detection than standard-of-care monitoring [36]. A retrospective single-center analysis of patients with MM treated with bispecific antibodies found that patients who received outpatient SUD with prophylactic dexamethasone combined with structured RPM had lower rates of CRS (13% vs 54%), recurrent CRS (0% vs 11%), and ICANS (7% vs 21%) compared with those who received standard inpatient SUD [37]. A larger study of patients with MM treated with teclistamab across 3 Mayo Clinic locations demonstrated the safety and feasibility of outpatient SUD using a technology package for RPM that comprised a tablet and Bluetooth-capable devices (thermometer, blood pressure monitor, and pulse oximeter) [31]. Among the 57 patients who completed outpatient SUD, CRS and ICANS occurred in 31.6% and 3.5%, respectively; all cases resolved, and no patients discontinued teclistamab treatment. In another Mayo Clinic study of 34 patients with MM treated with teclistamab and talquetamab SUD in an outpatient setting, including outpatient tocilizumab to treat CRS, 47% of patients were able to remain as outpatients even with an occurrence of CRS. In an evaluation of resource use, it was estimated that if all doses of teclistamab, talquetamab, and tocilizumab were given on an outpatient basis, it would have resulted in a total net revenue improvement of US $115,004 over the 7-month period analyzed [38].

In this Viewpoint, we aim to provide practical recommendations based on our own experiences for hematology and oncology physicians seeking to implement RPM for patients receiving bispecific antibodies for MM and provide support for the safe and feasible use and adoption of bispecific antibodies in the outpatient and community setting. We provide 2 case studies from our institutions that demonstrate successful RPM adoption at clinical practices using different systems and resourcing models. While these examples are specific to the administration of bispecific antibodies, the RPM approaches described herein can be applied to other therapies that require close monitoring for AEs. Of note, we discuss RPM implementation from a primarily US-based perspective, and therefore, the recommendations may not be broadly applicable to other countries due to differences in health care and payer systems.

Case Study Examples

The Atrium Health Levine Cancer Institute implemented an outsourcing program with Current Health and their care model Hospital at Home to monitor and provide care to patients in their homes [39]. Patients planned for bispecific antibody SUD

XSL•FO
RenderX

received a thermometer, blood pressure cuff, pulse oximeter, and wearable monitor to measure heart and respiratory rate, as well as a tablet to input their vital signs every 4 hours during waking hours. On the days between SUD, a mobile integrated health worker examined patients at home, evaluated the immune effector cell–associated encephalopathy score, and coordinated a virtual visit with a Hospital at Home internist. Patients also had round-the-clock virtual access to a team of Hospital at Home dedicated nurses, who are trained to manage phone calls from patients via a call service. Eight milligrams of prophylactic dexamethasone were administered the day after each SUD, and patients experiencing grade 1 CRS events were managed with acetaminophen and additional doses of dexamethasone as needed. Patients who developed persistent fever despite the above interventions, higher-grade CRS, or any-grade ICANS required hospital evaluation.

In addition to their established RPM program, Atrium Health is also exploring a temperature patch that measures core body temperature continuously for up to 90 days. The patch can be worn 24 hours a day, 7 days a week and is designed to alert the patient via mobile app if a fever develops, with the intent of identifying symptoms faster than self-reporting by the patient. This approach has shown to be an effective method for fever detection in patients with hematologic malignancies. Preliminary data on patients at high risk of febrile neutropenia who had recently undergone chemotherapy and stem cell transplantation show that continuous temperature monitoring with a wearable patch can be used for the accurate and early detection of fevers in an outpatient setting [40]. At Atrium Health, the temperature patch is currently being studied in patients treated with chimeric antigen receptor (CAR) T-cell therapy and may eventually be used to monitor all patients with cancer at risk of fever, including those receiving bispecific antibody treatment.

The Highlands Oncology Group is an independent practice that uses an ePRO approach for RPM in partnership with Canopy Care. Canopy's RPM system implements a smart device (eg, smartphone or tablet) for eligible patients, or phone calls for patients who do not have a smart device, to report symptoms to their care team. Symptom notifications are monitored by a dedicated triage nursing staff [41]. Patients with symptoms that exceed a predetermined severity threshold undergo further evaluation, including telephone triage, urgent outpatient office evaluation, or referral to the emergency department or hospital. Patients with symptoms of CRS or ICANS are referred to a partner hospital recommended by the practice.

An overview of the overall process for implementing RPM in the outpatient and community setting is illustrated in Figure 1, and specifics are discussed in the sections that follow.

XSL•FO
RenderX

Figure 1. Overview of key steps and considerations for implementing remote patient monitoring (RPM) in outpatient and community settings. *A dedicated team may consist of physicians, nurses (eg, bispecific registered nurses, and advanced practice registered nurses), and pharmacists. AE: adverse event.

XSL•FO
RenderX

Building the Business Case

Determining the benefits of RPM adoption may vary depending on the type of institution and the specific unmet needs. First, RPM may allow for broadening of patient access to bispecific antibody therapy. In the case of academic institutions, improving hospital bed capacity and expanding treatment throughput with the use of RPM may allow physicians to focus on patients who need more acute care. For community practices, RPM may provide a mechanism to safely retain patients rather than referring them to academic institutions or centers of excellence, enabling these institutions to grow and extend their reach without being limited by traditional resource constraints. Of note, RPM may allow for earlier detection of AEs [36] and better management of care in the outpatient and home settings that would positively impact patients' quality of life. Moreover, the implementation of RPM may improve institutional efficiency. As novel immunotherapies become more widely used in clinical practice, particularly as approvals are achieved in earlier lines of therapy, RPM will become increasingly more important for managing institutional workload and capacity considerations. With proper implementation and training, RPM has the potential to reduce the overall workload (eg, call volume) for institutions over time.

The successful implementation of RPM depends on the selection of the right technology to meet institutional and patient needs, as well as comprehensive training, institutional buy-in, and stakeholder engagement [34]. RPM may be seen as a new way of care delivery, yet when supported by adequate training and protocols, including patient and caregiver education on the signs and symptoms of potential complications (eg, CRS and ICANS) [21,42], RPM can advance the safe and effective administration of bispecific antibodies in the outpatient setting [31,36-38]. Ensuring the buy-in of all stakeholder groups (including executive leadership and management teams, health care providers [physicians and advanced practice providers, nurses, emergency room physicians, and pharmacists], and administrators) will help ensure successful implementation. Each practice should identify an RPM "champion" who can convey the benefit of RPM to stakeholders and guide the practice through the implementation process. In addition, the caregiver burden must also be considered as they may need to take time off work to care for the patient. For patients receiving bispecific antibodies, it is important for practices to establish standard operating procedures with specific criteria for outpatient SUD dictated by distance from the practice, caregiver availability, comorbidities, and disease characteristics [43] (Table 1).

Table . Recommended patient selection criteria for remote patient monitoring in outpatient and community settings [21,44].

Domain	Criteria
Disease burden	• ECOG[a] performance status of 0-2 • Low tumor burden
Comorbidities	• Well-controlled comorbidities • Good mental status
Prior CRS[b] or ICANS[c]	• No history of severe CRS or ICANS with prior dosing
Logistical considerations	• Close proximity to hospital (eg, within 60 min) • Dedicated caregiver support up to 48 h after the first target dose

[a]ECOG: Eastern Cooperative Oncology Group.

[b]CRS: cytokine release syndrome.

[c]ICANS: immune effector cell–associated neurotoxicity syndrome.

While implementing an RPM program can be time-consuming depending on the specific type of RPM and the infrastructure already in place, in the long term, the use of an RPM program will ultimately lead to a more streamlined process and better patient care delivery. Given that RPM models will differ between clinical practices, having a clear understanding of the institution's individual needs and working with an RPM vendor to tailor an approach specific to the practice can streamline the process and ensure that an effective solution is implemented. Securing institutional alignment is central as organizational support drives smoother implementation and long-term success of RPM.

Stakeholders also need to identify key performance indicators that will measure the success of the RPM program upon implementation, including measures of clinical, operational, and financial success. Clinical success can be measured via patient throughput and improvements in patient outcomes, such as lower-grade CRS events or fewer hospital readmissions. For example, Atrium Health Levine Cancer Institute conducted a study that implemented the use of RPM and prophylactic dexamethasone after each outpatient SUD, which limited the severity of CRS events compared with standard-of-care inpatient monitoring (41% grade 1 events, 19% grade 2 events, and no grade 3 or 4 events with outpatient SUD vs 33% grade 1 events, 12.5% grade 2 events, and 8% grade 3 or 4 events with inpatient SUD) and reduced the incidence of ICANS (6% vs 17%, respectively) [39]. The rate of recurrent CRS with subsequent SUD was also lower with outpatient SUD vs inpatient monitoring (32% vs 46%, respectively). In addition, a reduction in the total days of readmission (42 vs 105 days, respectively) and total inpatient days per patient (1.1 vs 7.8 days, respectively) was observed with outpatient SUD vs standard inpatient SUD.

XSL•FO
RenderX

Operational success can be defined when the RPM system is seamlessly integrated into the daily operations of the practice and those involved can shift their focus to other priorities, indicating that the system is functioning well without constant attention. Improvements in operational workflow result from the nurses or care teams spending less time answering unnecessary calls (allowing them to focus on other responsibilities) as well as from the reduced need for inpatient visits. At Atrium Health, the fielding and triaging of calls by the Hospital at Home team has helped alleviate the burden on providers in the clinic or hospital wards.

Finally, financial success is linked to the clinical and operational metrics (eg, reducing hospital stays and readmissions can lead to cost savings, which are quantifiable financial benefits). Compelling data from Atrium Health Levine Cancer Institute showed that fewer patients who received outpatient SUD with RPM and prophylactic dexamethasone required tocilizumab than those who received inpatient SUD based on the standard-of-care protocol (13% vs 42% of patients, respectively), resulting in reduced estimated total tocilizumab-related per-patient costs (US $868 vs US $2909, respectively) [39]. Overall, the mean estimated cost per patient was US $6444 with outpatient SUD vs US $29,911 with standard inpatient SUD. The use of ePROs at the Highlands Oncology Group to monitor infections in patients with hematologic malignancies has also demonstrated substantial cost savings, with reductions in infection-related hospitalizations and emergency department visits yielding an estimated US $977,695 annual savings per 1000 patients [41]. It is important to note that these results are based on our own single-center experiences; more cost-effectiveness data will emerge as more institutions adopt RPM to manage outpatient administration of bispecific antibodies.

Infrastructure Assessment and Device Considerations

The existing infrastructure and available resources within a clinical practice should be evaluated prior to RPM decision-making as these elements will inform the appropriate RPM solution. Existing infrastructure may include a dedicated phone nurse, an established electronic medical record (EMR) system (eg, Encore, Flatiron, or Epic), and overall IT maturity, as well as the ability to admit patients in the event of an AE. For example, Highlands Oncology uses an RPM system that functions as a hybrid solution approach, where the Highlands Oncology advanced practice registered nurse system standard operating procedures assist with patient monitoring. This model has decreased the risk of infection-related hospitalization rates by 52% and the risk of infection-related emergency department visits by 33% over 2 years for patients with hematologic cancers [41], and the first exclusive RPM patient was enrolled in July 2026. Moreover, the RPM system needs to be flexible so that it can be adjusted based on the institution's individual needs. Resources must also be available to provide answers to specific questions and implement algorithms for care (eg, "if patient says x, you do y"). Typically, RPM vendors will offer this type of support (tailored to meet the particular needs of the institution) to help reduce burden on the champion and overall care team [34].

The infrastructure at Atrium Health Levine Cancer Institute exemplifies collaboration between the physicians and the Hospital at Home nurses. This allows for an open line of communication and establishment of a proactive escalation plan that initiates physician involvement when necessary, thereby alleviating the burden of unnecessary calls, and allows for timely intervention to prevent hospitalization. In a case example, the on-call physician received a call from a Hospital at Home nurse about a patient who received SUD at Atrium Health Levine Cancer Institute followed by prophylactic dexamethasone and later experienced a fever. The patient took acetaminophen for the fever, which did not resolve within 1 hour, so the patient was administered an additional dose of dexamethasone. Even though this lowered the fever, the nurse and on-call physician proactively discussed next steps (tocilizumab administration at the infusion center) in case the patient's fever spiked again. This case study demonstrates how a model of open communication between the physician and Hospital at Home nurse was able to provide timely treatment guidance; develop a proactive escalation plan; and, ultimately, keep this patient with grade 1 CRS out of the hospital.

The evaluation of potential vendors for RPM should be based on the level of existing infrastructure within the institution. An initial assessment of available RPM solutions that considers the practice setup of the institution should be conducted. An assessment of other institutions with established RPM programs should also be conducted. One important consideration is whether the institution has the capacity or resources for dedicated phone nurses or if that component needs to be outsourced to an agency such as Hospital at Home or even the RPM vendor if such services are offered.

There are several factors that can facilitate implementation regardless of the RPM system or device. For example, enthusiasm and an openness to change, logistical maturity to handle patient escalation, and financial maturity to sustain the program will support the initial implementation and contribute to long-term success. Most importantly, all stakeholders need to be appropriately trained on the use of new therapies, including management of common AEs and appropriate triaging and escalation procedures, and education and training can be a time-consuming aspect of integration. Implementation of RPM can be straightforward with adequate training on how to recognize AEs (eg, CRS or ICANS) over the phone, as well as how to administer guidance, advice, and care for patients. On the basis of clinical experience from the Atrium Health Levine Cancer Institute and the Highlands Oncology Group, we recommend that institutions have the following infrastructure in place for the successful adoption of RPM for bispecific antibodies: (1) established hematology and oncology floors with clinicians who are familiar with and comfortable with caring for patients with cancer; (2) a pharmacist who is aware of tocilizumab, a pharmacy that stocks it, and a champion who understands the protocols and how to use tocilizumab; and (3) adequate training for all team members (eg, emergency room and critical care staff, infusion nurses, pharmacists, and hospice nocturnists). In the case of community setting triage, it is

XSL•FO
RenderX

recommended to have a "buddy system" with another clinician at a more experienced site. It should be noted that all these needs can be adapted to RPM of patients receiving other types of drugs with different safety profiles and monitoring requirements (eg, CAR T-cell therapy).

Technological systems should be adjustable to various settings and patient demographics. While vendors often provide technological support (eg, implementation managers), some level of IT maturity is required for implementation (eg, an established EMR system). However, it is a misconception that implementation of RPM can only be successful with full EMR integration. In the Current Health model at Atrium Health Levine Cancer Institute, Hospital at Home nurses have a separate dashboard for monitoring vital signs. If there is a change in vital signs, Hospital at Home nurses will document the vital sign change (eg, fever) in the Epic EMR system at Atrium Health, where it is then visible to the physicians. Thus, this monitoring approach can be successful either with or without full integration into the hospital's EMR system. It is also important to remember that limited resource barriers can be overcome. Having well-trained nursing staff available to answer patient calls and triage patients with serious AEs is critical. However, if capacity does not allow for internal phone nurses, outsourcing provides an option to eliminate this barrier, such as the Current Health Clinical Command Center. In addition, while pilot studies are recommended, they are not required. Pilot studies can provide learnings that inform a scale-up program; however, to be successful, a pilot study needs to be approached as a priority project that is well planned and executed with intention of full deployment in the future.

Reimbursement and Financial Challenges and Considerations

In the case of bispecific antibodies, RPM is currently not considered part of chronic care management in the United States, making billing and reimbursement something that needs to be established. While virtual visits with a physician can be billed to insurance, clinicians may not have established billing requirements for RPM, which is the case for the Atrium Health Levine Cancer Institute. In some programs, however, the cost savings of monitoring patients in an outpatient setting help offset the costs associated with RPM. Reimbursement considerations differ for Highlands Oncology as insurance payments do not cover the cost of an overnight nurse call and advanced practice registered nurse call. Additional billing options, including telehealth, are also being considered. In the United States, there are 9 reimbursement codes for RPM of physiological parameters in 2026, including a new code that covers RPM device supply and data transmission for 2 to 15 days of a 30-day period (Current Procedural Terminology code 99445) [45]. This code is particularly relevant for patients receiving bispecific antibodies who need to be monitored in the first few days after outpatient SUD.

Reimbursement strategies will depend on the type of treatment as well as the needs and capabilities of each institution. For some institutions, reimbursement for RPM may occur through billing codes as part of service. In the event that RPM is not included as part of the billing code, practices will need to explore other reimbursement mechanisms. In some cases, device provisioning can be considered, allowing the institution to deliver and manage the device. At Atrium Health Levine Cancer Institute, RPM (including monitoring or paramedic visits) does not incur a separate fee; the patient or insurance company is charged only for the virtual visits with the internist from Hospital at Home.

Contracting and Legal Considerations and Liabilities

Institutional buy-in and commitment can drive urgency and expedite the process of RPM adoption and implementation. Academic medical centers may face more bureaucratic hurdles than community practices. In academic medical centers and larger institutions, contracting can be a significant source of delays due to the extensive processes and guardrails involved, such as clinical implementation, data sharing, and IT requirements. These centers typically have more complex procedures to navigate. For example, integrating RPM devices into a large already established EMR system can introduce additional complexities, requiring collaboration between the institution and the vendor. Despite this, the benefits of the RPM systems and care models outweigh this potential hurdle. Conversely, community and independent practices generally face fewer contractual and logistical hurdles, allowing for a nimbler approach. While contracting can be a time-consuming part of the RPM process, implementation can be rapid once all paperwork has been signed, a process that is generally made smoother with institutional buy-in and input.

The legal liability of RPM is similar to that of traditional care, but the main challenge is adapting to the new paradigm of remote care. Physicians must accept some risk in remote monitoring, balanced by the benefits of outpatient care. Patients are no longer under constant direct supervision as they would be in a hospital, and this shift in thinking can be challenging for physicians who are accustomed to being able to check on patients in person in the event of a concern. Accepting this new model requires trust and confidence in the benefits of outpatient care for both the patients and the team of health care providers. Ultimately, it is up to the physician to determine the right balance between direct involvement and RPM, and this requires a personalized approach for each patient.

Potential Limitations of RPM

While RPM may offer several advantages for patients and providers, it is not without limitations. Programs will need to consider how to manage device malfunction or failure and ensure that patients have sufficient internet connectivity [34,46]. Some patients may have slower connections and data caps that could limit transmission of data; however, this can be addressed using RPM gateways that connect directly to providers and bypass the patient's home network [46]. Robust security measures are needed to prevent unauthorized access to patient data and protect patient privacy [46]. Digital literacy gaps may impact the success of RPM, particularly among older adult

XSL•FO
RenderX

patients who may find new technologies disorienting [46-48]. Socioeconomic barriers may also limit access to and adoption of RPM [46,47]. Proactive engagement and education are needed to ensure inclusive implementation of RPM for all eligible patients [48].

Conclusions

The practical recommendations discussed herein give health care providers the tools needed to successfully implement RPM for patients receiving bispecific antibodies in outpatient clinics and nonacademic centers, giving patients with MM broader access to bispecific antibodies and leading to improved patient outcomes. The use of RPM for patients with MM receiving bispecific antibodies has the potential to profoundly improve patient care and the patient experience. Potential benefits include increased access to care, earlier AE detection and better AE management, and lower inpatient hospitalization time. By shifting the way in which health care providers think about patient care, RPM can also improve efficiencies in the health care system by reducing resource burdens (eg, hospital staffing and capacity) and lowering overall health care costs. Community practices often refer their patients to academic centers for the SUD phase [42] and may lose them as a consequence. Decentralizing SUD for bispecific antibodies to community physicians empowers them to remain central in the care of their patients rather than transferring management exclusively to tertiary care centers. This allows for continuity of care, potentially leading to better outcomes, and facilitates trust and comfort between the patient and health care team.

Taken together, RPM implementation is a feasible strategy to optimize outpatient dosing of bispecific antibodies for both academic and community practices, and specific approaches will vary based on the needs and goals of each treatment center. Importantly, RPM protocols and management pathways are broadly applicable to other drugs (eg, CAR T-cell therapy [49]), different types of AEs, and diseases outside the oncology setting.

Acknowledgments

The authors thank Marlee Bogema, FNP-C, for helpful discussions during the preparation of this manuscript.

Funding

Editorial and medical writing support was provided by Lisa Shannon, PharmD, and Kimberly Fuller, PhD, of Lumanity and was funded by Johnson & Johnson. Support for article processing fees was provided by Johnson & Johnson.

Authors' Contributions

Writing—review and editing: CV, ES
Both authors approved the final version of the manuscript before submission.

Conflicts of Interest

CV reports a consultancy for Johnson & Johnson Innovative Medicine and research funding from Janssen and Lava Therapeutics. ES reports honoraria from Amgen, Johnson & Johnson, and Pfizer and a consultancy for Daiichi Sankyo.

References

1. TECVAYLI: summary of product characteristics. : Janssen (Belgium); 2024 URL: https://appliedsciencecommons.net/artifacts/18465371/souhrn-udaju-o-pripravku-1/19365840/ [accessed 2026-09-24]
2. TALVEY (talquetamab): summary of product characteristics. European Medicines Agency. URL: https://www.ema.europa.eu/en/documents/product-information/talvey-epar-product-information_en.pdf [accessed 2026-09-27]
3. Elrexfio: summary of product characteristics. European Commission. URL: https://ec.europa.eu/health/documents/community-register/2025/20250908167245/anx_167245_en.pdf [accessed 2026-09-24]
4. TECVAYLI (teclistamab-cqyv) injection, for subcutaneous use. U.S. Food and Drug Administration. URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2026/761291s018lbl.pdf [accessed 2026-09-24]
5. TALVEY (talquetamab-tgvs) injection, for subcutaneous use. U.S. Food and Drug Administration. 2023. URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2023/761342s000lbl.pdf [accessed 2026-09-24]
6. ELREXFIO (elranatamab-bcmm) injection, for subcutaneous use. U.S. Food and Drug Administration. URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2026/761345s007lbl.pdf [accessed 2026-09-24]
7. LYNOZYFIC (linvoseltamab): summary of product characteristics. European Medicines Agency. URL: https://www.ema.europa.eu/en/documents/product-information/lynozyfic-epar-product-information_en.pdf [accessed 2026-09-27]
8. LYNOZYFIC (linvoseltamab-gcpt) injection, for intravenous use. : Regeneron Pharmaceuticals, Inc; 2025 URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2025/761400s000lbl.pdf [accessed 2026-10-02]
9. Devasia AJ, Chari A, Lancman G. Bispecific antibodies in the treatment of multiple myeloma. Blood Cancer J 2024 Sep 12;14(1):158. [doi: 10.1038/s41408-024-01139-y] [Medline: 39266530]
10. Holstein SA, Grant SJ, Wildes TM. Chimeric antigen receptor T-cell and bispecific antibody therapy in multiple myeloma: moving into the future. J Clin Oncol 2023 Sep 20;41(27):4416-4429. [doi: 10.1200/JCO.23.00512] [Medline: 37471687]

XSL•FO
RenderX

11. Moreau P, Garfall AL, van de Donk NW, et al. Teclistamab in relapsed or refractory multiple myeloma. N Engl J Med 2022 Aug 11;387(6):495-505. [doi: 10.1056/NEJMoa2203478] [Medline: 35661166]
12. Chari A, Minnema MC, Berdeja JG, et al. Talquetamab, a T-cell-redirecting GPRC5D bispecific antibody for multiple myeloma. N Engl J Med 2022 Dec 15;387(24):2232-2244. [doi: 10.1056/NEJMoa2204591] [Medline: 36507686]
13. Lesokhin AM, Tomasson MH, Arnulf B, et al. Elranatamab in relapsed or refractory multiple myeloma: phase 2 MagnetisMM-3 trial results. Nat Med 2023 Sep;29(9):2259-2267. [doi: 10.1038/s41591-023-02528-9] [Medline: 37582952]
14. Raje N, Anderson K, Einsele H, et al. Monitoring, prophylaxis, and treatment of infections in patients with MM receiving bispecific antibody therapy: consensus recommendations from an expert panel. Blood Cancer J 2023 Aug 1;13(1):116. [doi: 10.1038/s41408-023-00879-7] [Medline: 37528088]
15. Banerjee R, Mohan M, Rejeski K, et al. Immunoglobulin prophylaxis should be initiated after bispecific antibody therapy in multiple myeloma, regardless of IgG levels. Blood Adv 2025 Sep 23;9(18):4720-4726. [doi: 10.1182/bloodadvances.2025016490] [Medline: 40690761]
16. Smits F, Groen K, Korst CL, et al. Immunoglobulin supplementation and longer dosing intervals reduce risk of infections in patients with RRMM treated with teclistamab. Blood Cancer J 2026 Feb 21;16(1):26. [doi: 10.1038/s41408-026-01451-9] [Medline: 41720763]
17. Nierengarten MB. Recommendations highlight important updates specific to systemic therapy options for those with newly diagnosed or previously treated multiple myeloma. Cancer 2025 Sep 1;131(17):e70014. [doi: 10.1002/cncr.70014] [Medline: 40887732]
18. Scott SA, Marin EM, Maples KT, et al. Prophylactic tocilizumab to prevent cytokine release syndrome (CRS) with teclistamab: a single-center experience. Blood Cancer J 2023 Dec 20;13(1):191. [doi: 10.1038/s41408-023-00963-y] [Medline: 38114481]
19. Vardell VA, Ermann DA, Tantravahi SK, et al. Impact of academic medical center access on outcomes in multiple myeloma. Am J Hematol 2023 Jan;98(1):41-48. [doi: 10.1002/ajh.26759] [Medline: 36266759]
20. Pan D, Richter J. Management of toxicities associated with BCMA, GPRC5D, and FcRH5-targeting bispecific antibodies in multiple myeloma. Curr Hematol Malig Rep 2024 Dec;19(6):237-245. [doi: 10.1007/s11899-024-00740-z] [Medline: 39145912]
21. Garfall AL, Banerjee R, Frenzel L, et al. A roadmap to implementing outpatient administration of bispecific antibodies in multiple myeloma. Front Oncol 2025;15:1630146. [doi: 10.3389/fonc.2025.1630146] [Medline: 40809030]
22. Kumar SK, Callander NS, Adekola K, et al. Multiple myeloma, version 5.2026. J Natl Compr Canc Netw 2026;24(1):e260001. [doi: 10.6004/jnccn.2026.0001]
23. Chiang AC, Olmedo Garcia ME, Carlisle JW, et al. Safety of tarlatamab with 6-8-h outpatient versus 48-h inpatient monitoring during cycle 1: DeLLphi-300 phase 1 substudy. ESMO Open 2025 Apr;10(4):104538. [doi: 10.1016/j.esmoop.2025.104538] [Medline: 40187110]
24. Ali A, Kaur A, Yang F, et al. Economic value of tocilizumab prophylaxis to prevent cytokine release syndrome (CRS) during outpatient teclistamab (Tec) or talquetamab (Tal) initiation for relapsed/refractory multiple myeloma (RRMM) [Poster]. Clin Lymphoma Myeloma Leuk 2025 Sep;25(Supplement 2):S48. [doi: 10.1016/S2152-2650(25)03475-5]
25. Chen C, Weatherby H, Masih-Khan E, et al. Outpatient step-up dosing of teclistamab (TEC): an implementation process shifting from the academic inpatient (IP) to community outpatient (OP) setting [Poster]. Clin Lymphoma Myeloma Leuk 2025 Sep;25(Supplement 2):S52. [doi: 10.1016/S2152-2650(25)03481-0]
26. Dhakal B, Ko G, He J, et al. Outcomes of teclistamab (Tec) step-up dosing in outpatient/hybrid settings among a large sample of relapsed refractory multiple myeloma (RRMM) patients treated with Tec [Poster]. Clin Lymphoma Myeloma Leuk 2025 Sep;25(Supplement 2):S57-S58. [doi: 10.1016/S2152-2650(25)03490-1]
27. Dao P, Nguyen T, San M, et al. Real world accelerated step-up dosing of teclistamab and talquetamab in the outpatient setting is feasible and associated with low incidence of cytokine release syndrome [Poster]. Clin Lymphoma Myeloma Leuk 2025 Sep;25:S66. [doi: 10.1016/S2152-2650(25)03504-9]
28. Raff L, Abrams J, Barisonek C, et al. Outcomes of outpatient step-up dosing (SUD) of teclistamab and talquetamab in patients with relapsed/refractory multiple myeloma (RRMM): findings from a large network of community practices in the USA [Poster]. Clin Lymphoma Myeloma Leuk 2025 Sep;25(Supplement 2):S77. [doi: 10.1016/S2152-2650(25)03523-2]
29. Varga C, Khouri J, Oveisi D, et al. Real-world safety outcomes and healthcare resource utilization (HCRU) during outpatient, inpatient, and hybrid step-up dosing (SUD) of teclistamab (Tec) and talquetamab (Tal): a chart review study [Poster]. Clin Lymphoma Myeloma Leuk 2025 Sep;25(Supplement 2):S77-S78. [doi: 10.1016/S2152-2650(25)03524-4]
30. Olivieri DJ, Banerjee R. Impatient for outpatient: operationalizing bispecific antibodies for multiple myeloma in the ambulatory setting. JCO Oncol Pract 2025 May;21(5):596-598. [doi: 10.1200/OP-24-00921] [Medline: 39705656]
31. Sandahl TB, Soefje SA, Fonseca R, et al. Real-world safety and health care resource utilization of teclistamab under an outpatient model for step-up dosing administration. JCO Oncol Pract 2025 May;21(5):702-709. [doi: 10.1200/OP-24-00489] [Medline: 39705632]
32. Mahmoudjafari Z, Ali A, Davis J, Sandahl T, Nachar V, Mancini R. Seamless navigation of bispecific therapies: optimizing management and outpatient access with a focus on coordination. J Adv Pract Oncol 2024 Sep 11;11:1-16. [doi: 10.6004/jadpro.2024.15.8.15] [Medline: 39802527]

XSL•FO
RenderX

33. Rockey-Bartlett C, Morelli J, Coffel M, Geracitano J, Lafata JE, Khairat S. Effect of remote patient monitoring on healthcare use among patients with cancer: a systematic review. Digit Health 2025;11:20552076251384220. [doi: 10.1177/20552076251384220] [Medline: 41084602]
34. Remote patient monitoring playbook. American Medical Association. 2022. URL: https://www.ama-assn.org/system/files/ama-remote-patient-monitoring-playbook.pdf [accessed 2026-09-24]
35. Essell J, Derman BA, Brooks B, et al. Remote therapeutic monitoring to support adoption of bispecific antibody therapy in community practices. J Clin Pathways 2025;11(3):19-26. [doi: 10.25270/jcp.2025.05.02]
36. Vanhellemont A, Coolbrandt A, Vandeneede E, et al. Remote patient monitoring for early detection of cytokine release syndrome in myeloma patients: a comparative study between standard care and remote monitoring. Blood 2024 Nov 5;144(Supplement 1):3664. [doi: 10.1182/blood-2024-203013]
37. Cirstea D, Kim EB, O'Neill S, et al. Bispecific outpatient step-up dosing with prophylactic dexamethasone: the BOSS program experience in myeloma [Poster]. Clin Lymphoma Myeloma Leuk 2025;25(Supplement 2):S53-S54. [doi: 10.1016/S2152-2650(25)03484-6]
38. Puttkammer JR, Barreto JN, Jensen CJ, et al. Outpatient management of bispecific related toxicities: an observational study of safety outcomes and resource utilization. JCO Oncol Pract 2026 Jan;22(1):83-90. [doi: 10.1200/OP-24-00930] [Medline: 40233295]
39. Varga C, Knight M, Davis D, et al. Prophylactic dexamethasone (dex) with outpatient step-up dosing (SUD) of bispecific antibodies (BsAb) in multiple myeloma (MM) v. standard of care (SOC) inpatient observation [Poster]. Clin Lymphoma Myeloma Leuk 2025;25(Supplement 2):S102-S103. [doi: 10.1016/S2152-2650(25)03562-1]
40. Verma N, Haji-Abolhassani I, Ganesh S, et al. A novel wearable device for continuous temperature monitoring & fever detection. IEEE J Transl Eng Health Med 2021;9:2700407. [doi: 10.1109/JTEHM.2021.3098127] [Medline: 34765323]
41. Essell J, Ascha M, Schaefer E, et al. Remote therapeutic monitoring reduces hospitalization due to infection in patients being treated for hematological malignancy. Blood 2025 Nov 3;146(Supplement 1):131. [doi: 10.1182/blood-2025-131]
42. Bellerive C, McNeill A, Faiman B, Richards TA. Optimizing transitions of care in multiple myeloma immunotherapy: nurse roles. Clin J Oncol Nurs 2025 Jan 17;29(1):E7-E16. [doi: 10.1188/25.CJON.E7-E16] [Medline: 42599895]
43. Derman BA, Roach M, Lin D, et al. Panel interview of oncology practices with emergent experience of teclistamab in the real world: the TecPIONEER study. Curr Med Res Opin 2024 May 16;16:1-6. [doi: 10.1080/03007995.2024.2352856] [Medline: 38745451]
44. Essell J, Ascha M, Shaefer E, et al. Remote therapeutic monitoring reduces hospitalization due to infection in patients being treated for hematological malignancy [Abstract]. Presented at: 67th American Society of Hematology (ASH) Annual Meeting and Exposition; Dec 6-9, 2025. [doi: 10.1182/blood-2025-131]
45. Medicare and Medicaid programs; CY 2026 payment policies under the physician fee schedule and other changes to part B payment and coverage policies; Medicare shared savings program requirements; and Medicare prescription drug inflation rebate program. Federal Register. URL: https://www.federalregister.gov/documents/2025/11/05/2025-19787/medicare-and-medicaid-programs-cy-2026-payment-policies-under-the-physician-fee-schedule-and-other [accessed 2026-09-24]
46. Claggett J, Petter S, Joshi A, Ponzio T, Kirkendall E. An infrastructure framework for remote patient monitoring interventions and research. J Med Internet Res 2024 May 30;26:e51234. [doi: 10.2196/51234] [Medline: 38815263]
47. Serrano LP, Maita KC, Avila FR, et al. Benefits and challenges of remote patient monitoring as perceived by health care practitioners: a systematic review. Perm J 2023 Dec 15;27(4):100-111. [doi: 10.7812/TPP/23.022] [Medline: 37735970]
48. Prospero E, Yazigi A, Rouzier R, Franzoi MA, L'Orphelin JM. Digital illiteracy and access to remote patient monitoring in oncology: a multicentre study of healthcare providers' perceptions and patient-reported digital literacy. EClinicalMedicine 2026;94:103866. [doi: 10.1016/j.eclinm.2026.103866] [Medline: 42005927]
49. Majhail NS, Cox T, Larson S, et al. Outpatient administration of chimeric antigen receptor T-cell therapy using remote patient monitoring. JCO Oncol Pract 2025 Nov;21(11):1601-1608. [doi: 10.1200/OP-25-00062] [Medline: 40249898]

Abbreviations

AE: adverse event
CAR: chimeric antigen receptor
CRS: cytokine release syndrome
EMR: electronic medical record
ePRO: electronic patient-reported outcome
ICANS: immune effector cell–associated neurotoxicity syndrome
MM: multiple myeloma
RPM: remote patient monitoring
SUD: step-up dosing

XSL•FO
RenderX

Edited by M Balcarras; submitted 24.Mar.2026; peer-reviewed by L Sharp, S Jabbar; revised version received 09.Sep.2026; accepted 10.Sep.2026; published 05.Nov.2026.

Please cite as:
Varga C, Schaefer E
Practical Recommendations for Implementing Remote Patient Monitoring for Patients Receiving Bispecific Antibodies for Multiple Myeloma
JMIR Cancer 2026;12:e95949
URL: https://cancer.jmir.org/2026/1/e95949
doi:10.2196/95949

© Cindy Varga, Eric Schaefer. Originally published in JMIR Cancer (https://cancer.jmir.org), 5.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of a Retrieval-Augmented Generation–Based Large Language Model for Evidence-Based Herb and Supplement Information in Cancer Care

Yen-Nien Hou[1], PharmD; Jyothirmai Gubili[1], MS; Pulkit Jain[2], MS; Chun Sing Lam[1], PhD; Avijit Chatterjee[2], PhD; Jun J Mao[1], MD, MSCE

[1]Integrative Medicine and Wellness Service, Memorial Sloan Kettering Cancer Center, 321 E 61st Street, 4th Fl., New York, NY, United States

[2]Digital, Informatics and Technology Solutions (DigITs), Artificial Intelligence & Machine Learning, Memorial Sloan Kettering Cancer Center, New York, NY, United States

Corresponding Author:
Jun J Mao, MD, MSCE
Integrative Medicine and Wellness Service, Memorial Sloan Kettering Cancer Center, 321 E 61st Street, 4th Fl., New York, NY, United States

Abstract

Our study describes the development and evaluation of a retrieval-augmented generation–based large language model to improve the quality of responses to provider questions about herbs and dietary supplements.

(JMIR Cancer 2026;12:e86073) doi:10.2196/86073

KEYWORDS

herbs; dietary supplements; cancer; health information; artificial intelligence

Introduction

Patients with cancer increasingly seek herbs and dietary supplements for symptom relief [1]. However, concurrent use of supplements with cancer treatment can lead to harmful adverse events and herb-drug interactions [2]. Additionally, many oncology providers have limited knowledge/training to counsel patients [3]. There is also a growing amount of misinformation about cancer treatments and exaggerated benefits of natural products for improving outcomes including curing cancer [4].

The *About Herbs* app—developed by the Integrative Medicine Service at Memorial Sloan Kettering Cancer Center (MSK)—delivers content from the *About Herbs* database that contains the latest research on 317 entries including herbs, minerals, and vitamins [5]. Currently, the app uses a keyword-based search system, which is inadequate to answer complex queries typically asked by providers. This can hinder their communication with patients on this important topic.

Since 2022, large language models (LLMs) are being increasingly used to obtain health-related information. While LLMs outperform online searches, they are prone to producing inaccurate responses or hallucinations [6], which can be especially harmful in oncology settings. Retrieval-augmented generation (RAG) is a technique that can improve the accuracy of responses by integrating the LLM with a searchable knowledge base [7]. To enhance the search functionality of the *About Herbs* app, we developed and evaluated a RAG-based LLM grounded in the *About Herbs* database.

Methods

Overview

We selected 26 monographs with extensive clinical data, which included commonly queried herbs/supplements in oncology practice at MSK, as the knowledge source for developing the RAG-based LLM (*AboutHerbsAI*): active hexose correlated compound, aloe vera, ashwagandha, astragalus, biotin, Boswellia, broccoli sprouts, calcium, cannabis, carnitine, coenzyme Q10, dong quai, flaxseed, folate, glucosamine, glutamine, green tea, licorice, melatonin, omega-3, probiotics, psilocybin, soy, turmeric, vitamin C, and vitamin D.

Development of *AboutHerbsAI* is described in Multimedia Appendix 1 [7,8].

Evaluation of *AboutHerbsAI* versus Current State-of-the Art LLMs

We developed 5 commonly asked questions (see Multimedia Appendix 1) surrounding herb/supplement use in cancer care, focusing on safety and dietary precautions, symptom management, supportive benefits during treatment, expected effects, and potential herb–drug interactions. These domains reflect common patient counseling considerations.

To compare with *AboutHerbsAI*, we chose 9 general-purpose LLMs available in April 2025 from 3 major vendors—OpenAI

XSL•FO
RenderX

(gpt-35-turbo-0613, gpt-4-turbo-2024-04-09, and gpt-4o-20240513); Anthropic (anthropic.claude-3-opus-20240229-v1:0, anthropic.claude-3-haiku-20240307-v1:0, and anthropic.claude-3‐5-sonnet-20240620-v1:0); and Meta (meta.llama3-1-8b-instruct-v1:0, meta.llama3-1-70b-instruct-v1:0, and meta.llama3-1-405b-instruct-v1:0)—to reflect the systems most likely to be accessed by clinicians (LLM settings are in Multimedia Appendix 1). Based on the QUEST (quality of information, understanding and reasoning, expression style and persona, safety and harm, and trust and confidence) framework for human evaluation of LLMs in health care as well as other RAG evaluation frameworks [9] and expert input, we arrived at 3 metrics to capture the quality of answers: (1) completeness ("Does the answer include all the pertinent information?"), (2) correctness ("Does the answer include false or true information?"), and (3) conciseness ("Does the answer include any non-pertinent information?"). The answers were rated on a binary scale of *true* or *false*.

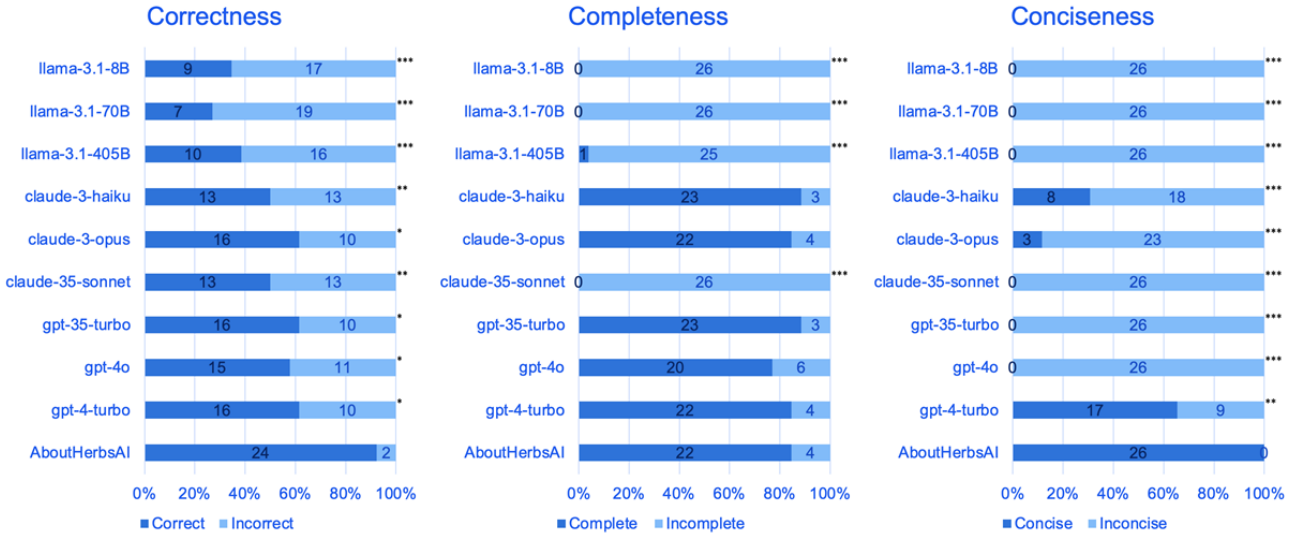
We randomly distributed the 26 monographs among 3 reviewers. They independently rated the answers to the 5 questions for all LLMs for the set of monographs assigned to them. The reviewers were not aware of which LLM each answer was coming from. We compiled the reviewer ratings and aggregated them to derive our final score for a particular LLM on a given metric. Interrater reliability assessment details are in Multimedia Appendix 1.

To compare performance between *AboutHerbsAI* and other LLMs, a chi-square test was used to assess whether the proportion of positive ratings differed significantly between models for each metric.

Results

In total, we evaluated 1300 question-answer pairs across 5 expert-curated questions, 26 monographs, and 9 LLMs and *AboutHerbsAI*. Interrater agreement was substantial for across rater pairs (κ=0.66‐0.74). Our model was superior to the 9 LLMs in both correctness and conciseness of responses with no significant differences in completeness (Figure 1).

Figure 1. Comparison of RAG-based LLM (*AboutHerbsAI*) with 9 general-purpose LLMs available in April 2025 from 3 major vendors (OpenAI, Anthropic, and Meta). LLM: large language model; RAG: retrieval-augmented generation. *$P \leq .05$; **$P \leq .01$; ***$P \leq .001$.



Correctness

AboutHerbsAI provided correct responses in 92.3% (24/26) cases compared to the leading LLMs (Claude-3-opus, GPT-35-turbo, and GPT-4-turbo) at 61.5% (16/26; *P*=.02). It also generated significantly more correct responses than other LLMs tested (range: 26.9% [7/26] to 57.7% [15/26]; all *P*<.02).

Completeness

AboutHerbsAI provided complete responses in 84.6% (22/26) cases compared to the leading LLMs (Claude-3-haiku and GPT-35-turbo) at 88.5% (23/26) with no statistically significant differences (*P*>.99). It also generated significantly more complete responses than Llama-3.1-8B, Llama-3.1-70B, Llama-3.1-405B and Claude-35-sonnet (range: 0% [0/26] to 3.8% [1/26]; all *P*<.001).

Conciseness

AboutHerbsAI provided concise responses in 100% (26/26) cases compared to the leading LLM (GPT-4-turbo) at 65.4% (17/26; *P*=.003). It also generated significantly more concise responses than all other LLMs (range: 0% [0/26] to 30.8% [8/26]; all *P*<.001).

Discussion

This study found that a RAG-based LLM substantially improved correctness and conciseness of responses to common questions about herbs and dietary supplements, when compared with 9 general-purpose LLMs, but slightly underperformed in completeness. We consider this acceptable given the benefit of grounding responses in an evidence-based database rather than open internet sources, which are prone to misinformation.

XSL•FO
RenderX

These results align with prior concerns about LLM hallucinations [6] and complementary and alternative medicine misinformation [4] and support our approach as a proof-of-concept study for integrating evidence-grounded LLM assistance [7] to improve the quality and accuracy of responses. This study is timely as oncology care is becoming increasingly digitized. An important limitation is the confinement of our model to a curated database that limits response breadth, especially when newer evidence is not yet included. Second, its performance was assessed in a controlled setting with limited content and predefined questions. Therefore, generalizability, and real-world queries—particularly those using vague or nonmedical/nontechnical language (eg, "hangover")—may pose challenges [10].

Future work will include system enhancements and evaluation of the clinical/patient-facing impacts of *About Herbs*. Database curation will involve regular ingestion of the latest literature to ensure timeliness. To support patients with varying health literacy needs, the app will offer a configurable layperson-oriented version to deliver information. Potential clinical implementations include embedding the RAG-based LLM within the *About Herbs* app to help physicians deliver evidence-based recommendations to minimize risks associated with the concomitant use of supplements during cancer treatment. Evaluation will involve assessing clinical significance, including effects on herb-related counseling, documentation, and safety, alongside integration into electronic medical records and clinic workflows.

Acknowledgments

We attest that no artificial intelligence–assisted technologies were used in the writing, editing, or preparation of this manuscript.

Funding

This work is supported in part by a National Institutes of Health/National Cancer Institute Cancer Center grant (P30 CA008748); the Laurance S. Rockefeller Fund that funds the MSK Herbal Education and Research in Oncology (HERO) program; and the Memorial Sloan Kettering's Translational and Integrative Medicine Research Fund.

Data Availability

All data generated in this study and evaluation code applied in this study is available at the online GitHub repository [11]. The code for the retrieval-augmented generation pipeline is available from the corresponding author, upon reasonable request, due to the potential commercial nature of the research.

Authors' Contributions

Conceptualization: JJM, YNH, PJ
Development of retrieval-augmented generation–based large language model: PJ
Data curation: YNH, PJ
Formal analysis: PJ, YNH, CSL
Methodology: JJM, YNH
Resources: JJM
Writing: YNH (lead), JG (equal), PJ (supporting), CSL (supporting), AC (supporting), JJM (supporting)

Conflicts of Interest

JJM has received grants from Tibet CheeZheng Tibetan Medicine Co Ltd and Zhongke Health International LLC to the Memorial Sloan Kettering Cancer Center outside the submitted work, serves in an unpaid role as co-president of the Society for Acupuncture Research, and is co-founder of Greatly Health. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Methods, including development and evaluation of *AboutHerbsAI*, LLM settings, and the interrater reliability assessment.
[DOCX File, 16 KB - cancer_v12i1e86073_app1.docx]

References

1. Sanford NN, Sher DJ, Ahn C, Aizer AA, Mahal BA. Prevalence and nondisclosure of complementary and alternative medicine use in patients with cancer and cancer survivors in the United States. JAMA Oncol 2019 May 1;5(5):735-737. [doi: 10.1001/jamaoncol.2019.0349] [Medline: 30973579]
2. Lam CS, Koon HK, Ma CT, et al. Real-world data on herb-drug interactions in oncology: a scoping review of pharmacoepidemiological studies. Phytomedicine 2022 Aug;103:154247. [doi: 10.1016/j.phymed.2022.154247] [Medline: 35716539]

XSL•FO
RenderX

3. Samuels N, Ben-Arye E. Exploring herbal medicine use during palliative cancer care: the integrative physician as a facilitator of pharmacist-patient-oncologist communication. Pharmaceuticals (Basel) 2020 Dec 11;13(12):455. [doi: 10.3390/ph13120455] [Medline: 33322582]
4. Loeb S, Langford AT, Bragg MA, Sherman R, Chan JM. Cancer misinformation on social media. CA Cancer J Clin 2024;74(5):453-464. [doi: 10.3322/caac.21857] [Medline: 38896503]
5. Hou YN, Deng G, Mao JJ. Practical application of "About Herbs" website: herbs and dietary supplement use in oncology settings. Cancer J 2019;25(5):357-366. [doi: 10.1097/PPO.0000000000000403] [Medline: 31567464]
6. De Angelis L, Baglivo F, Arzilli G, et al. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. Front Public Health 2023;11:1166120. [doi: 10.3389/fpubh.2023.1166120] [Medline: 37181697]
7. Lewis P, Perez E, Piktus A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. Presented at: 34th Conference on Neural Information Processing Systems (NeurIPS 2020); Dec 6-12, 2020.
8. Meskó B. Prompt engineering as an important emerging skill for medical professionals: tutorial. J Med Internet Res 2023 Oct 4;25:e50638. [doi: 10.2196/50638] [Medline: 37792434]
9. Tam TYC, Sivarajkumar S, Kapoor S, et al. A framework for human evaluation of large language models in healthcare derived from literature review. NPJ Digit Med 2024 Sep 28;7(1):258. [doi: 10.1038/s41746-024-01258-7] [Medline: 39333376]
10. Mei L, Liu S, Wang Y, Bi B, Cheng X. SLANG: new concept comprehension of large language models. Presented at: Conference on Empirical Methods in Natural Language Processing; Nov 12-16, 2024.
11. Pulkitcollier/LLM-RAG-evaluation-for-herbs-and-supplements-in-cancer-care. GitHub. URL: https://github.com/pulkitcollier/LLM-RAG-evaluation-for-herbs-and-supplements-in-cancer-care [accessed 2026-06-26]

Abbreviations

LLM: large language model
MSK: Memorial Sloan Kettering Cancer Center
QUEST: quality of information, understanding and reasoning, expression style and persona, safety and harm, and trust and confidence
RAG: retrieval-augmented generation

Edited by M Balcarras; submitted 17.Oct.2025; peer-reviewed by G Gabriels, HY Hsu, O Ajayi, Y Hou; revised version received 04.Mar.2026; accepted 05.Mar.2026; published 09.Jul.2026.

Please cite as:
Hou YN, Gubili J, Jain P, Lam CS, Chatterjee A, Mao JJ
Evaluation of a Retrieval-Augmented Generation–Based Large Language Model for Evidence-Based Herb and Supplement Information in Cancer Care
JMIR Cancer 2026;12:e86073
URL: https://cancer.jmir.org/2026/1/e86073
doi:10.2196/86073

© Yen-Nien Hou, Jyothirmai Gubili, Pulkit Jain, Chun Sing Lam, Avijit Chatterjee, Jun J Mao. Originally published in JMIR Cancer (https://cancer.jmir.org), 9.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Participation in Cervical and Breast Cancer Screening Among Working Women in Japan: Panel Survey Study

Miho Satoh[1], PHN, RN, PhD; Naoko Sato[2], RN, PhD

[1]Department of Nursing, Yokohama City University, 3-9 Fukuura, Kanazawa-ku, Yokohama, Japan

[2]Department of Nursing, Fukushima Medical University, Fukushima, Japan

Corresponding Author:
Miho Satoh, PHN, RN, PhD
Department of Nursing, Yokohama City University, 3-9 Fukuura, Kanazawa-ku, Yokohama, Japan

Abstract

With growing concern over high cervical and breast cancer mortality rates, this study was designed to examine the factors influencing attendance in breast and cervical cancer screening among working women in Japan by conducting a secondary analysis using nationally representative panel data from the Japanese Household Panel Survey (JHPS) and the Keio Household Panel Survey (KHPS) between 2019 and 2022, which showed that reduced screening attendance was associated with higher conscientiousness and neuroticism (possibly because of increased anxiety or fear of adverse results), suggesting that health-promoting behaviors and personality traits do not always align with preventive health actions, such as cancer screening.

(JMIR Cancer 2026;12:e86885) doi:10.2196/86885

KEYWORDS

cervical cancer; breast cancer; cancer screening; personality trait; working women

Introduction

Breast and cervical cancers have high incidence and mortality in Japan [1]. Mammography screening for breast cancer (ages 40‐69 years) aims to detect cancer early, while cervical cytology (ages >20 years) identifies human papillomavirus infection and precancerous changes to prevent cancer [2]. Both are recommended biennially and organized through municipal services under the Cancer Control Act. Japan's breast and cervical cancer screening rates (47.4% and 43.6%) remain lower than those of other OECD (Organisation for Economic Co-operation and Development) countries [1]. With increasing workforce participation, promoting screening among working women is crucial. Working women face time constraints [3], while psychosocial factors influence health behaviors [4,5]. This study explored factors affecting screening attendance among working women, examining personality traits and risk perception.

Methods

Study Design and Data

This study analyzed data from the Japanese Household Panel Survey (JHPS) and the Keio Household Panel Survey (KHPS), administered by the Panel Data Research Center (PDRC) [6]. The KHPS has surveyed 4005 individuals since 2004 who were born between 1935 and 1984, while the JHPS surveyed 4000 individuals since 2009 who were born after 1927. We analyzed data from Waves 2019‐2022, including baseline demographics. The sample included women aged up to 70 years, following Japan's breast and cervical cancer screening guidelines, which recommend screening for women older than 20 years for cervical cancer and those older than 40 years for breast cancer [2]. "Working women" were defined as female respondents employed full-time or part-time at each survey wave, with employment status as a time-varying covariate. Unemployed women were excluded from that wave's sample. PDRC sampling weights maintained national representativeness.

Variables

Dependent Variables

Breast and cervical cancer screening attendance between 2019 and 2022 was assessed. At each survey wave, participants reported whether they had undergone breast cancer screening (mammography) and cervical cancer screening (cervical cytology) within the past year, yielding a binary outcome (attended vs did not attend) for each cancer type. The dependent variable was a repeated binary outcome measured at four time points (2019, 2020, 2021, and 2022).

Independent Variables

Personality traits were assessed using the Japanese version of the Ten-Item Personality Inventory (TIPI-J) [7], which measures five dimensions: extraversion, agreeableness, conscientiousness, neuroticism, and openness. Each trait is evaluated on a 7-point Likert scale from 1 ("Strongly Agree") to 7 ("Strongly Disagree").

XSL•FO
RenderX

Risk aversion was assessed by asking respondents whether they would carry an umbrella based on rain probability in weather forecasts. A higher risk aversion ratio indicates stronger risk avoidance. This measure captures domain-general risk aversion, not cancer-specific risk perception. The JHPS/KHPS umbrella item is a validated behavioral proxy for general risk preference in Japanese panel studies [6]. While domain-general risk aversion may not perfectly reflect cancer-related health risk appraisals, research suggests that general risk preferences predict health-protective behaviors including cancer screening.

Self-rated health was measured using a 5-point scale from "Excellent" to "Poor." Mental health was assessed using the 6-item Kessler Psychological Distress Scale [8], which evaluates depressive and anxiety symptoms over the past month.

Sociodemographic variables included age, marital status, having children or not, caregiving responsibilities, academic background, employment status, and residential area.

Analysis

Participants completing at least two survey waves were included. A generalized linear mixed model with a binomial distribution and logit link function was used, with random intercepts for within-person correlation. For cervical cancer screening, women aged 20‐70 years were included; for breast cancer screening, women aged 40‐70 years were included, following Japanese guidelines. Wave-level missing cases were excluded while retaining completed waves, and participants with missing time-invariant covariates were removed.

Time-invariant variables were modeled as between-person predictors; time-varying covariates, at wave level. TIPI-J subscale scores (range 2‐14) were entered without standardization, with odds ratios reflecting 1-unit increments. The model incorporated demographics, personality traits, and time-varying factors using restricted pseudo-likelihood via GENLINMIXED in SPSS (version 29.0; IBM Corp). Poststratification weights from the PDRC, calibrated by age, sex, and region against census statistics, were applied owing to WEIGHT BY. GENLINMIXED's inability to specify stratification and clustering variables for complex survey variance estimation (a noted limitation).

Analyses were performed using SPSS Statistics 29.0, with significance at $P<.05$ (two-tailed).

Ethical Considerations

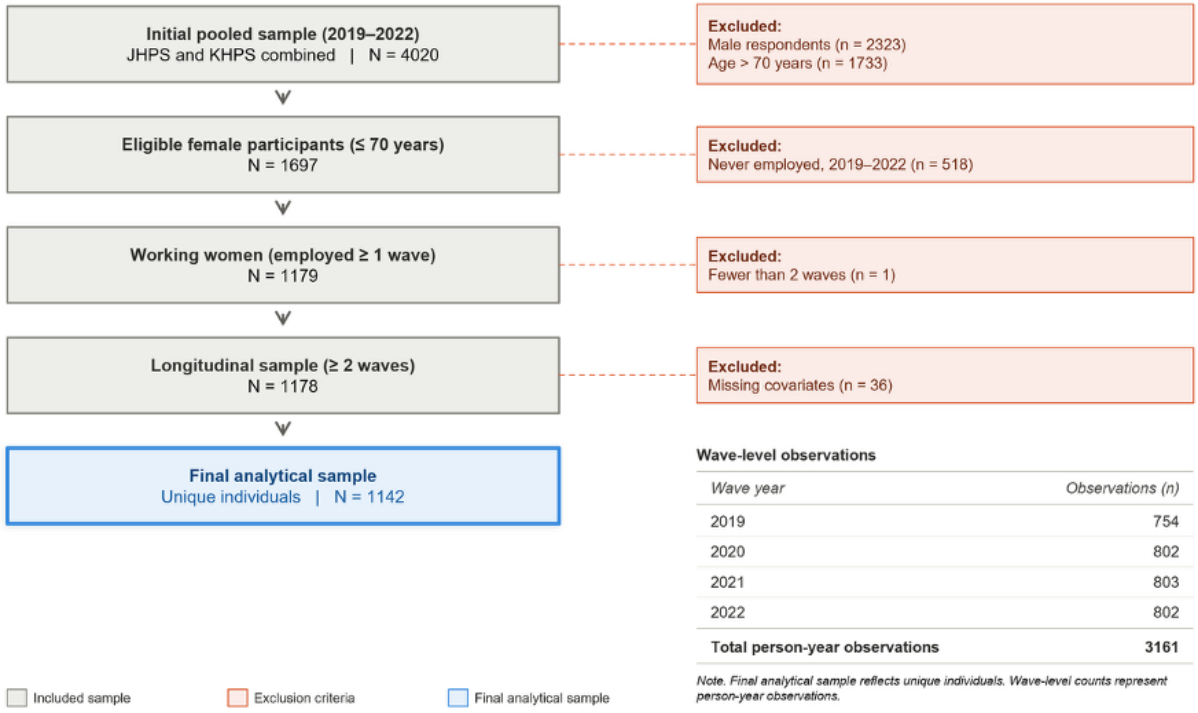
The JHPS/KHPS data used in this study were provided by the PDRC, Keio University (data provision number: 5841). This study was conducted as a secondary analysis of existing, fully deidentified panel survey data collected by the PDRC under a standing ethics framework. No new data were collected from human participants in the course of this study. In accordance with the Ethical Guidelines for Medical Research Involving Human Subjects (Ministry of Education, Culture, Sports, Science and Technology and Ministry of Health, Labour and Welfare, Japan; revised 2021) [9], secondary analyses of preexisting, anonymized datasets that were originally collected under informed consent do not require separate institutional review board approval. The original JHPS/KHPS surveys were conducted in compliance with the PDRC's own ethical framework, and all participants provided written informed consent regarding the purpose of the study, the use of their data, assurance of anonymity, strict protection of personal information, and secondary use by the PDRC. On this basis, a separate ethics review application was not submitted for this study. The corresponding author confirms that the study was conducted in accordance with the Declaration of Helsinki.

Results

Figure 1 provides an overview of the participant selection process.

XSL•FO
RenderX

Figure 1. Flowchart of participant selection. JHPS: Japanese Household Panel Survey; KHPS: Keio Household Panel Survey.



Wave-level observations

Wave year	*Observations (n)*
2019	754
2020	802
2021	803
2022	802
Total person-year observations	**3161**

Note. Final analytical sample reflects unique individuals. Wave-level counts represent person-year observations.

Table 1 summarizes the factors associated with participation in cervical and breast cancer screenings.

XSL•FO
RenderX

Table . Factors associated with participation in cervical cancer screenings and breast cancer screenings.

	Wave 2019 (n=754)	Wave 2020 (n=802)	Wave 2021 (n=803)	Wave 2022 (n=802)	Cervical cancer screening				Breast cancer screening			
					Odds ratio	Lower limit of 95% CI	Upper limit of 95% CI	*P* value	Odds ratio	Lower limit of 95% CI	Upper limit of 95% CI	*P* value
Cervical cancer screening participation (20 - 70 years), n (%)[a]	119 (15.8)	124 (15.5)	117 (14.6)	142 (17.7)	__[b]	—	—	—	—	—	—	—
Breast cancer screening participation (40 - 70 years), n (%)	119 (17.4)	134 (18.8)	118 (15.7)	129 (17.7)	—	—	—	—	—	—	—	—
Year (reference: 2022）												
2019	—	—	—	—	1.28	0.78	2.13	.33	1.03	0.62	1.71	.92
2020	—	—	—	—	1.01	0.64	1.60	.95	0.91	0.57	1.45	.69
2021	—	—	—	—	1.37	0.87	2.21	.19	1.15	0.72	1.84	.57
Age group[c], mean (SD)[d]												
29 - 39 years	89 (11.8)	84 (10.5)	48 (6.0)	73 (9.1)	—	—	—	—	—	—	—	—
40 - 49 years	269 (35.7)	253 (31.5)	222 (27.6)	242 (30.2)	0.74	0.34	1.60	.45	—	—	—	—
50 - 59 years	253 (33.6)	285 (35.5)	307 (38.2)	288 (35.9)	1.08	0.48	2.40	.86	1.48	0.91	2.40	.11
60 - 70 years	143 (19)	180 (22.4)	226 (28.1)	199 (24.8)	1.10	0.46	2.62	.84	1.38	0.76	2.53	.29
Extraversion, mean (SD)	8.60 (2.55)	—	—	—	0.92	0.84	1.01	.09	0.92	0.84	1.01	.09
Openness, mean (SD	7.59 (2.17)	—	—	—	1.04	0.93	1.16	.51	1.02	0.91	1.14	.72
Conscientiousness, mean (SD)	7.79 (2.16)	—	—	—	0.88	0.79	0.98	.02	0.86	0.77	0.97	.01
Neuroticism, mean (SD)	8.08 (2.16)	—	—	—	0.88	0.78	0.99	.04	0.87	0.77	0.98	.02
Agreeableness, mean (SD)	9.89 (1.82)	—	—	—	1.01	0.90	1.15	.84	0.96	0.85	1.09	.57
Risk aversion, mean (SD)	47.78 (16.73)	46.94 (16.55)	48.21 (16.96)	47.68 (16.32)	1.01	1.00	1.02	.21	1.01	1.00	1.02	.15
Mental health, mean (SD)	5.65 (5.15)	5.49 (5.00)	4.91 (4.52)	5.04 (4.73)	0.98	0.93	1.03	.40	0.99	0.95	1.04	.74
Self-rated health, mean (SD)	3.35 (0.89)	3.41 (0.91)	3.47 (0.89)	3.41 (0.87)	1.03	0.80	1.31	.83	1.01	0.80	1.29	.91

XSL•FO
RenderX

		Wave 2019 (n=754)	Wave 2020 (n=802)	Wave 2021 (n=803)	Wave 2022 (n=802)	Cervical cancer screening				Breast cancer screening			
						Odds ratio	Lower limit of 95% CI	Upper limit of 95% CI	*P* value	Odds ratio	Lower limit of 95% CI	Upper limit of 95% CI	*P* value
Marital status, n (%)													
	Reference: unmarried	201 (26.7)	220 (27.4)	218 (27.1)	219 (27.3)	—	—	—	—	—	—	—	—
	Married	553 (73.3)	582 (72.6)	585 (72.9)	583 (72.7)	0.76	0.43	1.34	.34	0.72	0.40	1.28	.27
Child status, n (%)													
	Reference: ≥15 years old	538 (71.4)	565 (70.4)	565 (70.4)	565 (70.4)	—	—	—	—	—	—	—	—
	None	120 (15.9)	136 (17)	140 (17.4)	137 (17.1)	0.90	0.44	1.83	.77	0.78	0.38	1.60	.50
	<15 years old	96 (12.7)	101 (12.6)	98 (12.2)	100 (12.5)	0.75	0.39	1.44	.39	0.64	0.33	1.25	.19
Caregiver role status, n (%)													
	Reference: Yes	138 (18.3)	155 (19.3)	157 (19.6)	151 (18.8)	—	—	—	—	—	—	—	—
	No	616 (81.7)	647 (80.7)	646 (80.4)	651 (81.2)	1.02	0.58	1.78	.95	1.18	0.68	2.03	.56
Education background, n (%)													
	Reference: university/graduate school	313 (46.8)	331 (46.7)	328 (45.9)	334 (46.8)	—	—	—	—	—	—	—	—
	Junior/senior high school	130 (19.4)	239 (33.7)	242 (33.9)	235 (33)	0.71	0.39	1.31	.28	0.60	0.33	1.10	.10
	Junior/technical college	226 (33.8)	139 (19.6)	144 (20.2)	144 (20.2)	0.85	0.52	1.38	.50	0.74	0.45	1.21	.23
Employment status, n (%)													
	Reference: temporary	531 (70.4)	552 (68.8)	551 (68.6)	554 (69.1)	—	—	—	—	—	—	—	—
	Permanent	223 (29.6)	250 (31.2)	252 (31.4)	248 (30.9)	0.96	0.59	1.55	.86	1.03	0.64	1.66	.91

XSL•FO
RenderX

	Wave 2019 (n=754)	Wave 2020 (n=802)	Wave 2021 (n=803)	Wave 2022 (n=802)	Cervical cancer screening				Breast cancer screening			
					Odds ratio	Lower limit of 95% CI	Upper limit of 95% CI	*P* value	Odds ratio	Lower limit of 95% CI	Upper limit of 95% CI	*P* value
Region of residence, n (%)												
Reference: towns or villages[e]	62 (8.2)	65 (8.1)	66 (8.2)	67 (8.4)	—	—	—	—	—	—	—	—
Government-designated city[f]	453 (60.1)	482 (60.1)	483 (60.1)	487 (60.7)	1.44	0.67	3.06	.35	1.59	0.76	3.33	.22
City[g]	239 (31.7)	255 (31.8)	254 (31.6)	248 (30.9)	1.47	0.72	2.98	.29	2.24	1.11	4.53	.02

[a]For categorical variables.

[b]Not applicable.

[c]Reference category for age differed by outcome: 29 - 39 years for cervical cancer screening and 40 - 49 years for breast cancer screening.

[d]For continuous variables.

[e]Population is <50,000.

[f]Population is ≥500,000.

[g]Population is ≥50,000 to <500,000.

For cervical cancer screening, women with higher conscientiousness (odds ratio [OR] 0.88, 95% CI 0.79 - 0.98; *P*=.02) and neuroticism (OR 0.88, 95% CI, 0.78 - 0.99; *P*=.04) showed lower attendance likelihood.

For breast cancer screening, women with higher conscientiousness (OR 0.86, 95% CI 0.77 - 0.97; *P*=.01) and neuroticism (OR 0.87, 95% CI 0.77 - 0.98; *P*=.02) showed reduced attendance.

Discussion

This study showed that women with high conscientiousness and neuroticism were less likely to participate in cancer screening. While conscientious individuals typically show strong self-discipline linked to preventive health behaviors [4,10], our findings revealed increased conscientiousness correlated with reduced screening uptake. Japanese women in demanding roles may prioritize work and family over preventive health care. Higher neuroticism may increase anxiety about diagnoses, discouraging screening [11], though both require investigation.

Study limitations include potential recall and social desirability biases from self-administered questionnaires, selection bias from listwise exclusion of missing data, general rather than cancer-specific risk aversion measures, and lack of health care trust and access barrier data. Using women in their 30s as a reference group may be misleading, since mammography is not recommended for this age in Japan. Future research should examine structural and psychosocial factors affecting screening.

Acknowledgments

We are grateful to the Keio University Panel Data Research Center for providing

Keio Household Panel Survey (KHPS)/Japanese Household Panel Survey (JHPS) data. ChatGPT was used for English language editing and proofreading. All artificial intelligence–generated content was reviewed and verified by the authors, who take full responsibility for the accuracy and integrity of the work.

Funding

This study did not receive any specific grants from funding agencies in the public, commercial, or nonprofit sectors. The authors declared that no financial support was received for this work.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. OECD Health Statistics. OECD. 2024. URL: https://www.oecd.org/en/data/datasets/oecd-health-statistics.html [accessed 2025-04-30]
2. Ministry of Health, Labour and Welfare, Japan. Cancer screening guidelines. 2016 URL: https://www.mhlw.go.jp/bunya/kenkou/gan_kenshin.html [accessed 2025-04-30]
3. Mon HM, Robb KA, Demou E. Effectiveness of workplace cancer screening interventions: A systematic review. BMC Cancer 2024 Aug 12;24(1):999. [doi: 10.1186/s12885-024-12649-0] [Medline: 39134945]
4. Emanuel AS, Kiviniemi MT, Howell JL, et al. Avoiding cancer risk information. Soc Sci Med 2015 Dec;147:113-120. [doi: 10.1016/j.socscimed.2015.10.058] [Medline: 26560410]
5. Hajek A, Kretzler B, König HH. Personality and the use of cancer screenings. A systematic review. PLoS ONE 2020;15(12):e0244655. [doi: 10.1371/journal.pone.0244655] [Medline: 33370379]
6. Overview of the research at PDRC. Keio University. URL: https://www.pdrc.keio.ac.jp/en/about/overview/ [accessed 2025-04-30]
7. Oshio A, Abe S, Cutrone P. Development, reliability, and validity of the Japanese Version of Ten Item Personality Inventory (TIPI-J). Jpn J Personal 2012;21(1):40-52. [doi: 10.2132/personality.21.40]
8. Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [doi: 10.1017/s0033291702006074] [Medline: 12214795]
9. Ethical Guidelines for Medical and Biological Research Involving Human Subjects. Ministry of Education, Culture, Sports, Science and Technology (MEXT). URL: https://www.mext.go.jp/content/20250325-mxt_life-000035486-01.pdf [accessed 2025-04-30]
10. Hajek A, Becher H, Brenner H, et al. Personality and the use of cancer screenings - results of the German National Cohort. Prev Med Rep 2024 May;41:102677. [doi: 10.1016/j.pmedr.2024.102677] [Medline: 38533391]
11. Niedzwiedz CL, Robb KA, Katikireddi SV, Pell JP, Smith DJ. Depressive symptoms, neuroticism, and participation in breast and cervical cancer screening: Cross-sectional and prospective evidence from UK Biobank. Psychooncology 2020 Feb;29(2):381-388. [doi: 10.1002/pon.5272] [Medline: 31663185]

Abbreviations

JHPS: Japanese Household Panel Survey
KHPS: Keio Household Panel Survey
OECD: Organisation for Economic Co-operation and Development
OR: odds ratio
PDRC: Panel Data Research Center
TIPI-J: Japanese version of the Ten-Item Personality Inventory

Edited by M Balcarras; submitted 31.Oct.2025; peer-reviewed by F Liu, M Solbjor; revised version received 30.Mar.2026; accepted 31.Mar.2026; published 28.Apr.2026.

Please cite as:
Satoh M, Sato N
Participation in Cervical and Breast Cancer Screening Among Working Women in Japan: Panel Survey Study
JMIR Cancer 2026;12:e86885
URL: https://cancer.jmir.org/2026/1/e86885
doi:10.2196/86885

© Miho Satoh, Naoko Sato. Originally published in JMIR Cancer (https://cancer.jmir.org), 28.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Research Letter

Remote Patient Monitoring Use Among Commercially Insured Adults With Cancer

Joseph H Joo[1,2,3], MD, MS; Changchuan Jiang[1,2], MD, MPH; Jessica I Billig[2,4], MD, MSc; Nadia Lieu[1,2], MS; Shining Yang[1,2], MA; Yujia Jin[1], MS; Arthur S Hong[1,3], MD, MPH; Joshua M Liao[1,2,3], MD, MSc

[1]Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, United States

[2]Program on Policy Evaluation and Learning, Dallas, TX, United States

[3]Department of Health Economics, Systems, and Policy, University of Texas Southwestern Medical Center, Dallas, TX, United States

[4]Department of Plastic Surgery, University of Texas Southwestern Medical Center, Dallas, TX, United States

Corresponding Author:
Joseph H Joo, MD, MS
Department of Internal Medicine
University of Texas Southwestern Medical Center
5323 Harry Hines Blvd
Dallas, TX, 75390
United States
Phone: 1 (214) 648 3111
Email: Joseph.Joo@UTSouthwestern.edu

Abstract

Our study describes the characteristics of remote patient monitoring use among commercially insured patients with cancer from 2019 to 2023.

(JMIR Cancer 2026;12:e84788) doi:10.2196/84788

KEYWORDS

remote patient monitoring; oncology; cancer; commercially insured; utilization

Introduction

Remote patient monitoring (RPM) enables continuous monitoring of vital signs, including temperature, heart rate, and blood pressure. By capturing this information while patients are at home or in their communities, RPM can detect early disease status changes and treatment-related adverse events and potentially improve outcomes through early intervention [1]. As such, RPM holds substantial promise for cancer treatment. In particular, RPM can support timely interventions to facilitate earlier recognition of decompensation, including fever and hypoxia. In the United States, however, outside of several single-centered trials, data on RPM use among patients with cancer are very limited in oncology [2,3]. Therefore, the aim of our nationwide study was to describe the characteristics of RPM use among commercially insured patients with cancer.

Methods

We used 2019-2023 Merative MarketScan commercial claims, which contain longitudinal commercial insurance claims in the United States. The claims included nearly 40 million unique patients in the outpatient setting for the study period. We used Current Procedural Terminology codes (99091, 99453, 99454, 99457, 99458) to identify RPM services and International Classification of Diseases, Tenth Revision, Clinical Modification codes to identify patients with cancer. Cancer types were classified per American Cancer Society classifications and stratified by the cancer status (active vs history of cancer) [4]. Patients were included if they received RPM and were enrolled in insurance at least 11 out of 12 months in a calendar year.

We described the number and characteristics of patients with cancer with respect to age, sex, region, clinical complexity (ie, Charlson Comorbidity Index) and insurance plan type (health maintenance organization vs preferred provider organization vs other) [5]. We also assessed the total number of months of RPM use and the average number of months of RPM use per patient.

Because of deidentified claims data, the study was determined to be nonhuman research and therefore exempt from institutional review board at the University of Texas Southwestern Medical Center.

XSL•FO
RenderX

Results

Between 2019 and 2023, RPM was used on 236 patients (Table 1). Half of our sample were 55-64 years old (118/236, 50.0%) and located in the southern region of the United States (117/236, 49.6%), with considerable clinical complexity (106/236, 44.9% with a Charlson Comorbidity Index of 4+). Over half of patients receiving RPM were female (141/236, 59.7%). The plurality of patients was enrolled in a preferred provider organization (89/236, 37.7%), followed by high-deductible health (64/236, 27.1%) and health maintenance organization (39/236, 16.5%) plans.

Among the 236 patients, 86.9% (205/236) used RPM for one year, while 13.1% (31/236) used RPM for 2 or more years. These patients accrued a total of 611 total months of RPM, equating to approximately 2.6 months of RPM use per patient. Of all the RPM use in patients with cancer, 45.7% (448/981) of RPM use was from those with other cancers (ie, cancer of unknown primary, rare cancers, cancer subtypes, and pre-cancers), followed by patients with cancers of the blood and lymph system (122/981, 12.4%) and patients with breast cancer (121/981, 12.3%). RPM was also used among patients with digestive (79/981, 8.1%), reproductive (77/981, 7.9%), urinary (46/981, 4.7%), endocrine (36/981, 3.7%), lung and chest (30/981, 3.1%), and skin cancers (15/981, 1.5%). RPM was least frequently used among patients with bone and soft tissue cancer (1/981, 0.1%), brain and nervous system cancer (3/981, 0.3%), and head and neck cancer (3/981, 0.3%)

Table 1. Characteristics of the patients with cancer using remote patient monitoring (RPM).

	Patients using RPM in 2019-2023 (N=236), n (%)	Patients using RPM for 1 year (n=205), n (%)	Patients using RPM for 2 or more years (n=31), n (%)
Age (y)			
18-34	21 (8.9)	21 (10.2)	0 (0)
35-44	32 (13.6)	28 (13.7)	4 (12.9)
45-54	64 (27.1)	56 (27.3)	8 (25.8)
55-64	118 (50)	99 (48.3)	19 (61.3)
>64	1 (0.4)	1 (0.5)	0 (0)
Sex			
Female	141 (59.7)	115 (56.1)	26 (83.9)
Male	95 (40.3)	90 (43.9)	5 (16.1)
Region			
South	117 (49.6)	101 (49.3)	16 (51.6)
West	45 (19.1)	41 (20)	4 (12.9)
North Central	35 (14.8)	34 (16.6)	1 (3.2)
Northeast	39 (16.5)	29 (14.1)	10 (32.3)
Insurance			
High deductible	64 (27.1)	59 (28.8)	5 (16.1)
Health maintenance organization	39 (16.5)	33 (16.1)	6 (19.4)
Preferred provider organization	89 (37.7)	77 (37.6)	12 (38.7)
Other[a]	38 (16.1)	31 (15.1)	7 (22.6)
Missing	6 (2.5)	5 (2.4)	1 (3.2)
Charlson Comorbidity Index score			
0-1	60 (25.4)	50 (24.4)	10 (32.3)
2-3	70 (29.7)	62 (30.2)	8 (25.8)
4+	106 (44.9)	93 (45.4)	13 (41.9)

[a]Other insurance includes basic/major medical plan, comprehensive plan, exclusive provider organization plan, noncapitated (non-cap) point-of-service plan, and capitated (cap) or partially capitated (part cap) point-of-service plan.

XSL•FO
RenderX

Discussion

In this nationwide US study, RPM use was limited to less than 1% (236/6,437,829) of patients with cancer and focused on particular patient groups. These exploratory findings offer several implications for future work.

First, our finding that RPM in the cancer setting was used more in older female patients reflects findings from prior noncancer studies that also showed female patients between the ages of 45 and 64 years to engage most with their clinicians through RPM [6,7]. Future research should evaluate facilitators and barriers to RPM use among younger patients with cancer, particularly given increases in early-onset cancer in the United States and likely greater familiarity with digital interventions such as RPM. Further work is needed to better understand why nearly half of the patients with RPM use were in the south region.

Second, our findings reflect uneven RPM use across cancer types, with disproportionately higher use among other cancers group (cancers of unknown primary, rare cancers, cancer subtypes, and pre-cancers) over other types such as blood and lymph system cancer. Considering that urinary, breast, lung, and digestive cancers are the most common cancer diagnoses in the United States, more work is needed to evaluate the appropriateness and impact of such findings [8]. Because not only different types of cancer but also the stages of treatment (eg, in remission, undergoing chemotherapy/radiation, postsurgery) could largely impact the indications for RPM use, future research must inform clinicians and policymakers which patients with cancer could most benefit from RPM.

Third, our finding that a patient used RPM for an average of 2.6 months indicates that RPM use is still nascent in the cancer setting. In contrast to more RPM use in chronic diseases such as hypertension (with one studying showing patients using RPM about 9 months), there is a dearth of evidence for longitudinal RPM use in cancer [9,10]. Future research is needed to evaluate the outcomes associated with RPM use to better motivate implementation among health systems.

Study limitations included descriptive design and inability to evaluate the associations between RPM use with cancer outcomes. Although RPM use in cancer could have been more prevalent through clinical trials, which are not captured in our study, the goal of our analysis was to assess real-world adoption within commercial claims data. Furthermore, our study does not capture adoption in patients with public health insurances. Ultimately, our results provide early evidence of opportunity to increase RPM use and potentially fruitful areas for future research and implementation.

Acknowledgments

We attest that no artificial intelligence–assisted technologies were used in the writing, editing, or preparation of this manuscript.

Funding

This research protocol was funded by the University of Texas Southwestern National Institutes of Health Clinical and Translational Science Award grant 1U54TR002361.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to data use and licensing restrictions from Merative MarketScan Commercial Database but are available from the corresponding author on reasonable request. The dataset analyzed during this study is the Merative MarketScan Commercial Database, which is one of the largest repositories of commercial claims data available. MarketScan is a registered trademark of Merative Corporation in the United States.

Authors' Contributions

Conceptualization: JHJ, JML
Data curation: SY, YJ
Formal analysis: SY, YJ
Funding acquisition: JML
Investigation: JHJ, JML
Methodology: JHJ, CJ, JIB, ASH, JML
Project administration: JHJ, JML
Resources: JHJ, JML
Software: SY, YJ
Supervision: JML
Validation: JHJ, JML
Visualization: JHJ, CJ, JIB, NL, SY, YJ, ASH, JML
Writing - original draft: JHJ, CJ, JIB, NL, SY, YJ, ASH, JML
Writing - review & editing: JHJ, CJ, JIB, NL, SY, YJ, ASH, JML

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. DeRogatis MJ, Pellegrino AN, Wang N, Higgins M, Dubin J, Issack P, et al. Enhancing recovery and reducing readmissions: the impact of remote monitoring on acute postoperative care in outpatient total joint arthroplasty. J Orthop 2024 Dec;58:111-116. [doi: 10.1016/j.jor.2024.06.028] [Medline: 39114426]
2. Pritchett JC, Paludo JC, Borah BJ, Haddad TC. Remote patient monitoring to transform management of febrile neutropenia in cancer patients. NEJM Catalyst 2024 Mar 20;5(4):1. [doi: 10.1056/cat.23.0365]
3. Pritchett JC, Borah BJ, Desai AP, Xie Z, Saliba AN, Leventakos K, et al. Association of a remote patient monitoring (RPM) program with reduced hospitalizations in cancer patients with COVID-19. JCO Oncology Practice 2021 Sep;17(9):e1293-e1302. [doi: 10.1200/op.21.00307]
4. Cancer types. American Cancer Society. URL: https://www.cancer.org/cancer/types.html [accessed 2026-02-02]
5. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson Comorbidity Index: a critical review of clinimetric properties. Psychother Psychosom 2022;91(1):8-35 [FREE Full text] [doi: 10.1159/000521288] [Medline: 34991091]
6. Adepoju O, Dang P, Nguyen H, Mertz J. Equity in digital health: assessing access and utilization of remote patient monitoring, medical apps, and wearables in underserved communities. Inquiry 2024;61:469580241271137 [FREE Full text] [doi: 10.1177/00469580241271137] [Medline: 39323052]
7. Kirkland E, Schumann SO, Schreiner A, Heincelman M, Zhang J, Marsden J, et al. Patient demographics and clinic type are associated with patient engagement within a remote monitoring program. Telemed J E Health 2021 Aug 01;27(8):843-850 [FREE Full text] [doi: 10.1089/tmj.2020.0535] [Medline: 34115942]
8. Cancer stat facts: common cancer sites. National Cancer Institute: Surveillance, Epidemiology, and End Results Program (SEER Program). URL: https://seer.cancer.gov/statfacts/html/common.html [accessed 2026-02-02]
9. Smith W, Colbert BM, Namouz T, Caven D, Ewing JA, Albano AW. Remote patient monitoring is associated with improved outcomes in hypertension: a large, retrospective, cohort analysis. Healthcare (Basel) 2024 Aug 09;12(16):1583 [FREE Full text] [doi: 10.3390/healthcare12161583] [Medline: 39201142]
10. Mahuron K, Esslin P, Sun V. Randomized controlled trial-perioperative telemonitoring of patient-generated health data in gastrointestinal oncologic (GI) surgery: assessing outcomes. Ann Surg 2026 Jan 01;283(1):48-56. [doi: 10.1097/sla.0000000000006724]

Abbreviations

RPM: remote patient monitoring

Edited by N Cahill, M Balcarras; submitted 25.Sep.2025; peer-reviewed by I Kudel, T Yu, WW Tarazi; comments to author 17.Nov.2025; revised version received 26.Jan.2026; accepted 27.Jan.2026; published 12.Feb.2026.

Please cite as:
Joo JH, Jiang C, Billig JI, Lieu N, Yang S, Jin Y, Hong AS, Liao JM
Remote Patient Monitoring Use Among Commercially Insured Adults With Cancer
JMIR Cancer 2026;12:e84788
URL: https://cancer.jmir.org/2026/1/e84788
doi:10.2196/84788
PMID:

©Joseph H Joo, Changchuan Jiang, Jessica I Billig, Nadia Lieu, Shining Yang, Yujia Jin, Arthur S Hong, Joshua M Liao. Originally published in JMIR Cancer (https://cancer.jmir.org), 12.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of a Web-Based Experiential Learning Intervention for the Public to Reduce Cancer Stigma: Tutorial on the Application of Intervention Mapping

Miyako Tsuchiya[1,2*], CPsychol, PhD; Keiichiro Adachi[3*], CP, PhD; Akiko Kimata[2,4*], RN, PhD; Kaori Kumagai[5*], RN, PHN, PhD

[1]Research Institute of Nursing, Musashino University, 3-3-3, Ariake, Koto-ku, Tokyo, Japan

[2]Former Division of Health Care Delivery, Survivorship and Policy Research, Institute of Cancer Control, National Cancer Center, Tokyo, Japan

[3]Faculty of Health Sciences, Yamaguchi University Graduate School of Medicine, Ube, Japan

[4]School of Nursing, Mejiro University, Saitama, Japan

[5]School of Nursing and Social Services, Health Sciences University of Hokkaido, Ishikari, Japan

[*]all authors contributed equally

Corresponding Author:
Miyako Tsuchiya, CPsychol, PhD
Research Institute of Nursing, Musashino University, 3-3-3, Ariake, Koto-ku, Tokyo, Japan

Abstract

Background: Stigma may negatively impact individuals throughout the continuum of cancer care and survivorship. Multitheory and multilevel intervention programs are necessary to reduce stigma but remain globally limited.

Objective: This tutorial aims to illustrate the development of a web-based experiential learning intervention, "Friend Supporter," designed for the public, which simulates scenarios to foster empathy and helping intentions. We applied the intervention mapping (IM) approach, which is rooted in the socioecological model, using the first four steps.

Methods: In step 1, key issues faced by cancer survivors and influential factors were identified through empirical evidence and literature reviews on cancer stigmas and psychological theories. A multidisciplinary planning team assessed issue-related logic. In step 2, a logic model of change was created based on step 1 findings. In step 3, we designed program themes and a structure using systematic reviews and needs surveys among the public (n=1076) and cancer survivors (n=473), while applying theoretical change methods and practical strategies. Step 4 integrated prior findings. Inputs from an expert panel (n=5) and the public (n=13) using the think-aloud approach were used to refine the materials and functions, with educational resources for program providers also developed.

Results: Step 1 revealed that public misconceptions and attitudes worsened the quality of life of cancer survivors. Step 2 identified 3 long-term outcomes: reduced public bias, improved responses to disclosure via acquired skills, and support aligned with survivors' aspirations. The short-term primary outcome was helping intention. Personal factors (knowledge, skills, and self-efficacy) were also expected to improve helping intention as mediators. In step 3, the literature review showed that multicomponent online interventions effectively reduced stigma. The survey indicated the public most needed information on "how to interact with friends diagnosed with cancer" (317/1076, 29.5%), regardless of whether they had a friend diagnosed with cancer (χ^2_1=0.98; *P*=.32). Participants with no friends diagnosed with cancer were more likely to require information concerning "survival rates of all types of cancer" (χ^2_1=7.3; *P*=.007). Preferred delivery modes were booklets or leaflets (529/1076, 49.2%) and the internet (texts and figures: 460/1076, 42.8%). Cancer survivors wanted their friends to understand "the possibility of a cure as a result of early detection and treatment" (193/473, 40.8%). To produce program materials, we applied stigma and discrimination, protection motivation, social cognitive, and learning theories. The 5-module program included self-learning, role-plays, worksheets, and written feedback from clinical psychologists. Step 4 confirmed the feasibility of the program with minor refinement. We then developed a practical guide for program providers' future implementation.

Conclusions: IM is useful for systematically developing web-based multitheory and multilevel interventions. "Friend Supporter" offers a promising approach to enhance supportive behaviors and reduce cancer stigma. Quantitative evaluation is underway using the final 2 IM steps (implementation and evaluation) to determine real-world effectiveness.

(JMIR Cancer 2026;12:e71166) doi:10.2196/71166

KEYWORDS

cancer; digital intervention development; intervention mapping; stigma; survivorship

XSL•FO
RenderX

Introduction

Background

Cancer stigma is a global public health concern, contributing to poor health and social inequality among survivors [1]. Stigma, which can be categorized as involving both public stigma and internalized stigma, is produced through complex societal processes [2]. The term "cancer" may automatically evoke fears of death (negative stereotypes or ignorance) or emotional discomfort (prejudice), which then may lead to discrimination toward cancer survivors. Survivors who are aware of this public stigma, in turn, may experience diminished self-esteem, intense self-blame, and shame [3], reflecting internalized stigma.

Public stigma undermines the cancer care continuum by reducing screening uptake, delaying diagnosis, and lowering treatment adherence, which can contribute to poor survival outcomes [1,4,5]. The literature further demonstrates that stigma negatively affects survivorship through impaired mental health [6-8], weakened social relationships [9], and reduced social participation [8,10]. Survivors of breast, cervical, prostate, and lung cancer are particularly vulnerable; the associations reported between these survivors and sexuality or risk-related behaviors (eg, smoking) may intensify both public and internalized stigma [3,7,9,11,12], further hinder timely access to care [4,8,13], and worsen quality of life (QoL) [3,14,15].

In Japan, however, policy interventions are not primarily aimed at reducing mortality through stigma mitigation. Instead, as articulated in the Fourth Basic Plan for Cancer Control [16], the national focus is on cancer survivorship through promoting coexistence with cancer by reducing stigma, enhancing public understanding, and fostering a supportive environment for survivors. This approach emphasizes citizen engagement and social inclusion to improve survivors' QoL and to remove stigma-related barriers to social participation.

To reduce stigma, multitheory and multilevel interventions are recommended [13,17-20]. These approaches are grounded in the socioecological model, which emphasizes the dynamic interactions between intrapersonal factors and broader environmental contexts, such as interpersonal relationships, organizational structures, community settings, and policy frameworks [21]. Most illness-related stigma reduction programs focus on HIV/AIDS, leprosy, tuberculosis, mental illness or substance use disorders, and epilepsy [17-19,22], with limited numbers of cancer-specific intervention programs [1,23]. In Japan, 1 intervention dealing with romantic relationships among Japanese youngsters has been reported [24]. Most interventions rely on education and/or intergroup contact [17-19,21,24]. Sustaining intergroup contact poses challenges due to infectious risks (eg, COVID-19) and the burden on patients with cancer. Thus, interventions that minimize in-person contact while maintaining efficacy are needed.

Intervention content and delivery must also reflect relationships that survivors have (eg, with friends, colleagues, employers, teachers, caregivers, neighbors, and even those with no prior connection), survivors' cancer types, and available provider resources (eg, human and infrastructure). However, developing multiple intervention versions for each context is impractical [25]. Instead, adapting core components may be a more feasible approach. We focused on friendships, which are vital to survivors' well-being across the lifespan [26,27]. According to stigma research, engaging with hypothetical friends has been shown to help reduce psychological distance and foster empathetic understanding of an individual's experience of stigma [28].

Brief Description of the Intervention

We developed a 5-week web-based experiential learning program named "Friend Supporter" that was designed for the adult public with no history of cancer diagnosis (hereafter, the public). In this program, experiential learning is defined as a method that allows the program participants to learn the contents of the program and deepen their understanding of practical conduct through simulation. The program centers on a hypothetical scenario: "What if your friend were diagnosed with cancer?" This framing enables broad participation regardless of prior connection to cancer survivors as friends. The primary outcome of the program was helping intentions, assessed immediately after program completion. In addition, the program was designed to foster long-term outcomes, including supportive interactions with adult cancer survivors, increased support provision, and stigma reduction.

The intervention was conceptually designed for broad applicability across cancer types in adulthood, with a theoretical foundation that addressed common relational challenges and support needs identified in diverse contexts [29]. Through text-based role-play, program participants practice empathetic listening and supportive responses across diverse scenarios. These scenarios vary by cancer type, age, gender, occupation, disclosure context, and conversational content. These varied scenarios are intended to help participants adapt to different contexts. By applying core interpersonal skills learned through the program, they can respond more effectively across various situations. The program is grounded in the socioecological model and primarily focuses on intrapersonal and interpersonal levels, but with further potential effects at organizational and community levels.

Objective

This paper aimed to illustrate the developmental processes of the 5-week web-based experiential learning program "Friend Supporter," using the first 4 steps of the intervention mapping (IM) approach [30].

Methods

The IM Approach

Overview

IM, rooted in the socioecological model, is a systematic method for developing, intervening, and implementing health promotion programs by integrating theories, empirical evidence, and practical input [30]. IM emphasizes that interventions operate within dynamic systems, where their impact is influenced by interaction across environmental levels [30]. Thus, IM has been widely recommended for designing multilevel interventions

XSL•FO
RenderX

[31]. This framework has been applied in diverse digital interventions; for example, pain self-management apps [32] and mobile-based self-compassion programs for patients with cancer [33], and web-based exercise interventions for patients with diabetes [34].

IM comprises 6 steps. Step 1 involves the identification of issues and factors, step 2 involves the establishment of program outcomes, step 3 involves program design, step 4 involves the production of materials, step 5 involves implementation planning, and step 6 involves implementation plan evaluation. Each step includes the completion of several tasks that form the basis for the subsequent step; however, the procedure is not always performed in a linear manner [30].

Step 1: Identification of Issues and Factors

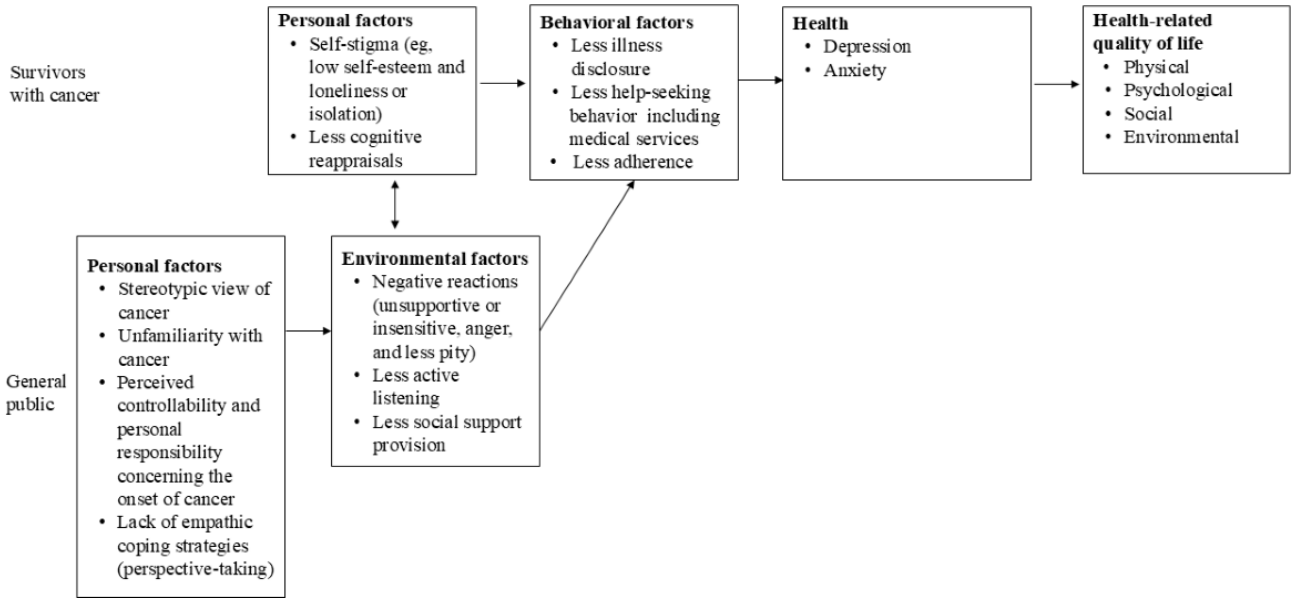
In step 1, we identified issues in cancer survivors' QoL and associated personal, behavioral, and environmental factors [30]. To do this, we conducted needs assessments focusing on breast cancer survivors [35,36]. This initial focus was chosen due to the availability of relevant Japan-based datasets and the stigma literature documenting relational challenges associated with breast cancer globally. We also investigated the public perceptions of cancer [37,38], existing literature reviews on cancer stigma [39], and relevant stigma and psychological theories [40-42]. This allowed us to provide a conceptual foundation for designing an intervention with broader applicability.

A multidisciplinary planning team was established that included a health psychologist, a clinical psychologist, a nurse, and an oncologist (another clinical psychologist and another nurse subsequently joined the planning team). Three members had the experience of information website development. All members discussed the logic model of the problem to reach a consensus.

Step 2: Establishment of the Program Outcomes

In step 2, we created a logic model of change that was intended to illustrate how the intervention program could modify or reduce the factors directly and indirectly associated with cancer survivors' QoL, as identified in step 1; this model served as a theoretical framework rather than a definitive causal pathway [30]. As the logic model of the problem (Figure 1) shows, there were 2 possible populations for which the intervention program could be developed.

Figure 1. A logic model of the problem.



The multidisciplinary planning team discussed which populations needed to have interventions and what needed to be changed to improve the QoL of cancer survivors, in accordance with the national policy in Japan. Based on these discussions, the planning team decided to develop an intervention program designed for the public. This decision was grounded in the reality that many cancer survivors resume their usual lives and reintegrate into their communities after discharge. If a new intervention program could reduce stigma and discrimination among program participants, they might become advocates for cancer survivors. This, in turn, could empower cancer survivors and promote a society where living with cancer is more widely accepted.

A logic model of change was drafted to guide the intervention design. Behavioral and environmental impact, performance objectives (long-term outcomes), and change objectives (short-term outcomes) were established through a stepwise logic modeling process. Change objectives were mapped to personal determinants and organized in a matrix. These were theorized to support individual behavioral and interpersonal changes.

Step 3: Program Design

Overview

Step 3 involved generating themes, components, and sequences of components for the intervention program [30]. Theory-, evidence-, and practice-based change methods should be determined to optimize the effectiveness and adaptability of an intervention [30]. We reviewed existing systematic reviews and conducted 2 surveys (one among the public and another among cancer survivors) in Japan. This approach was adopted because

XSL•FO
RenderX

customization based on user preferences facilitates smoother implementation [43]. To design the program, various theoretical and practical change methods were discussed among the planning team.

Existing Systematic Review of Literature

To identify existing effective interventions that could achieve the performance objectives established in step 2, the Cochrane Database of Systematic Reviews (CDSR) was searched. To gain insight into the intervention design and components, we reviewed interventions that were indicated in the findings of a meta-analysis as having significant effects.

Needs Survey Among the Public

To investigate the necessary content and preferable modes of intervention delivery, a cross-sectional online survey was conducted among the adult public (≥20 y) in Japan.

The detailed survey procedures have been previously reported in a study concerning a cancer-related scale development process among the public [44]. Potential adult Japanese participants were conveniently recruited from the panels of a research company. The company developed and tested the survey website for this study. Participants voluntarily accessed the closed survey and read the information sheets before participating. Each participant ticked the statement "I agree to participate in this study" before responding to the questionnaire. Question items were randomly presented to minimize order bias. The researchers created all the study materials except for the survey website and had no direct contact with the participants. The response rate was reported in a previously published article based on the same survey dataset [44]. As the recruitment and data collection were conducted by the research company, the investigators did not have access to individual-level invitation procedures and response tracking.

Data concerning demographics, willingness to disclose their cancer to surrounding people (families, friends, colleagues, and neighbors) if the participants were diagnosed with cancer, and whether they had friends with cancer were obtained. The participants were instructed to read a scenario describing a situation where a hypothetical friend talked about a cancer diagnosis, after which the participants were asked to respond to the following question: "Reflecting on your reactions to have better relationships with the friend with cancer, please tell us what you want to know?" There were 24 items with dichotomous responses. Additionally, a detail concerning preferred modes of delivery when the participants received the relevant information was requested.

The study variables were summarized using descriptive statistics (mean and SD or frequencies and percentages). Chi-square tests were conducted to examine differences in information needs and preferred modes of delivery between the participants whose friends had been diagnosed with cancer and those who had not in real life. Information needs were categorized into the relevant change objectives identified in step 2.

Needs Survey Among Cancer Survivors

To investigate the necessary content for the public, a cross-sectional online survey was conducted among adult cancer survivors in Japan. The detailed survey procedures have been previously reported in a study concerning cancer survivors' psychological distress with cancer disclosure [45]. Potential adult Japanese participants were recruited from the panels of a research company, which were independent of those involved in the public survey. The recruitment process, consent process, survey design, and data collection methods were identical to those used in the public survey.

Demographic and cancer-related information was collected. The participants, including those who had been diagnosed with cancer 5 years previously and those who had already disclosed their cancer diagnosis to friends, were asked to respond to the following question: "To have better relationships, what information would you want your friends to know when you tell them of your cancer diagnosis?" There were 20 items with dichotomous responses.

The study variables were summarized using descriptive statistics (mean and SD or frequencies and percentages). Information needs were categorized into the relevant change objectives identified in step 2.

Setting Up the Theoretical and Practical Change Methods

Considering the findings from the CDSR search and the 2 surveys, the planning team discussed the appropriate mode of delivery and themes and sequences of the components of the intervention program. In parallel, theoretical change methods for each change objective were selected by referring to a taxonomy of behavior change methods in IM [46] where applicable. The planning team discussed various practical change methods to reach a consensus.

Step 4: Program Materials

Overview

Step 4 involved drafting, pilot testing, refining, and finalizing the program material [30]. The planning team drafted and reviewed the initial content. While the evaluation framework was initially conceptualized during step 2 through the logic model of change, the relevant behavioral and environmental impact was further structured during step 4, as part of the intervention logic model. A multidisciplinary expert panel evaluated the conceptual and practical validity of the materials. The IT team deployed materials for pilot testing. The adult public participants engaged in think-aloud sessions and individual interviews. Based on these findings, we refined the materials, finalized the intervention logic model, and produced accompanying educational resources.

Internal Material Production

The planning team reviewed the theoretical and practical change methods identified in step 3 and proposed that the program consist of 5 modules corresponding to the themes and sequences established in step 3. Prior to the development of the module materials, the planning team established formatting and stylistic guidelines to ensure consistency and accessibility. These included: standardized font type and size across all modules; use of concise and short sentences written in polite Japanese; and active use of visual elements such as images, charts, and diagrams to support comprehension. These guidelines were

XSL•FO
RenderX

applied throughout the material development process to ensure a coherent learning experience across modules.

The planning team was divided into subgroups according to their expertise (eg, medicine and nursing, and psychology). Each subgroup created specific component materials, using credible sources and empirical findings, where applicable. To facilitate the design and review process, materials were initially created in slide format (Microsoft PowerPoint), allowing for visual structuring and annotation. The fourth component of the program, which features role-playing scenarios, required a distinctive approach in relation to the other components. The program participants could choose from multiple cases and progress through them interactively. This component was visually prototyped to ensure clarity, usability, and flexibility. First-person narratives were used in role-play case stories, based on evidence from a systematic review showing their effectiveness in reducing mental health prejudice [47]. In the role-play case stories, a character (a cancer survivor) shares a personal experience via text on a screen (eg, "I've been diagnosed with cancer—take care of yourself too"). The program participants, acting as the character's friend, are prompted to respond through text-based interactions, simulating a supportive conversation. Finally, the planning team independently reviewed all the initial materials, and their feedback was incorporated through iterative refinements.

Multidisciplinary Expert Inputs

A multidisciplinary expert panel (n=5), including a medical oncologist, a psychiatrist, an oncology nurse, and researchers in pedagogy and sociology, independently evaluated the intervention logic model and paper-based materials. Electronic versions with voice recordings in role-play components were also provided. Regarding the evaluation of the intervention logic model, 16 questions were developed and assessed using a 5-point Likert scale (completely agree/agree/neither agree nor disagree/disagree/completely disagree), referring to relevant logic model guides [48,49]. To evaluate the materials, 29 questions were similarly assessed using the same scale. Open-ended questions were also included to gather qualitative feedback. Responses to the Likert scale were summarized by frequency. To guide subsequent refinements, particular attention was given to comments associated with "disagree" and "completely disagree" responses.

Pilot Testing

After refinement reflecting the expert panel's concerns, an IT team deployed "Friend Supporter" on the Internet. The planning team provided the IT team with detailed instructions regarding slide navigation and hyperlinking (in both PowerPoint presentation and PDF formats) to ensure that the materials could be deployed effectively. Revisions were conducted collaboratively through embedded comments within the slides, enabling iterative refinement. The program participants would be expected to complete one module per week using the internet, and modules they had already completed could not be reviewed.

To investigate the feasibility, usability, and acceptability of this web-based program, we conducted a pilot study using a think-aloud approach among the public in Japan (this study has already been published [50]). Participants were recruited using snowball sampling. Participant characteristics were summarized using frequency, and the interview transcripts were analyzed using content analysis. Based on the findings, the program was further refined, the intervention logic model was finalized, and educational materials for implementation were developed.

Ethical Considerations

Ethical approval for the survey conducted among the public was obtained from the institutional research board of the National Cancer Center (2017 - 378). Ethical approval for the survey among cancer survivors was obtained separately from the institutional research board of the National Cancer Center (2017 - 134). For the 2 surveys, appropriate electronic informed consent was obtained from all participants. The surveys were anonymous. Participants in the public survey received a 2,000-yen gift voucher (approximately 17 USD at the time of study).Ethical approval for pilot testing was obtained from the institutional research board of the National Cancer Center (2021 - 020). Electronic informed consent was obtained from all participants. Participants in the interviews received a 2,000-yen QUO card (a prepaid purchasing card), which was approximately 15 USD. No personally identifiable information, including initials, was included in the transcripts and the published results. All procedures were conducted in accordance with the ethical standards of the institutional research board and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Results

Step 1: Identification of Issues and Factors

The relevant literature shows that some members of the public have negative reactions toward cancer survivors due to beliefs about cancer (eg, "cancer equals death" and "incurable disease" [37]), lack of familiarity with survivors, and perceptions of the risk factors [37-39]. These interactions can adversely affect survivors' self-concepts, cognitive reappraisals, social disclosure [35,36], help-seeking behaviors (eg, less medical use), treatment adherence [39], mental health status [39], and QoL [36]. Psychological theories such as stigma attribution theory [40,41] and empathetic coping theory [42] have been used to elaborate on the links between the general public's attitudes and the impact of such attitudes on cancer survivors. The findings underscore the need to address cancer stigma across the cancer experience, including coexistence, early detection, and clinical care. Figure 1 illustrates the relationship among health problems, QoL, and associated factors. The multidisciplinary planning team agreed to this logic model of the problem.

Step 2: Establishment of the Program Outcomes

The logic model of change was structured around three key components: performance objectives, change objectives, and personal determinants. This stepwise structure enabled a systematic evaluation of how targeted changes in personal determinants may lead to behavioral shifts, ultimately contributing to a more supportive environment for cancer survivors. While improved QoL of cancer survivors was the ultimate goal, this was not directly measured in this model.

XSL•FO
RenderX

To clarify the logic of change, the primary outcome was defined as strengthening the public's intention to support hypothetical friends with cancer, representing the short-term learning effect of the program. This proximal change was considered a key indicator of immediate program success. To achieve the primary outcome, the intervention targeted intrapersonal mediators of change, such as increased knowledge about cancer and cancer survivors and greater self-efficacy, as well as interpersonal skills such as empathetic coping skills. Change objectives for each performance objective were presented together with personal determinants (knowledge, skills, self-efficacy, and attitudes) in a matrix; for example, regarding change objectives aimed at reducing stereotypes and prejudice about cancer and cancer survivors, personal determinants associated with knowledge focused on "increasing accurate knowledge of cancer and cancer survivors" (Multimedia Appendix 1). These short-term shifts were theorized to contribute to the long-term goals outlined in the three performance objectives and were expected to emerge over a span of 2 to 3 years.

Step 3: Program Design

Existing Systematic Review of Literature

The CDSR search found that there were no systematic reviews about cancer stigma reduction interventions for the public. However, there was 1 article about mental illness stigma reduction interventions [47], which reported that mass media intervention, first-person narratives, and multiple-component interventions were effective in reducing prejudice about mental illness. Additionally, a meta-analysis revealed that 2 reviewed interventions [51,52] significantly decreased prejudice and increased knowledge of mental illness and of affected patients. These interventions included information about mental illnesses and relevant patients, addressing negative reactions toward such patients and providing practical guides to help these patients using patient-related stories. However, these interventions did not address communication issues shown in Multimedia Appendix 1 or the established outcome. The planning team decided not to adapt these existing programs in the cancer context.

Needs Survey Among the Public

A total of 1076 people participated in the survey. The descriptive statistics showed that the mean age was 47.2 (SD 11.4) years, 35.4% (381/1076) were women, and 17.5% (188/1076) had friends with cancer [44]. The participants reported that if they were diagnosed with cancer, they would "strongly agree/agree" to disclose their cancer diagnoses to their families (920/1076, 85.5%), friends (474/1076, 44.1%), colleagues (480/1076, 44.6%), and neighbors (104/1076, 9.7%).

After reading scenarios depicting hypothetical friends with cancer, the participants identified the following 5 most sought types of information: "how to interact with friends diagnosed with cancer" (317/1076, 29.5%), "types of cancer treatment" (288/1076, 26.8%), "side-effects of cancer treatment" (249/1076, 23.1%), "survivors' desire for relationships with their friends" (212/1076, 19.7%), and "what cancer survivors do not want friends to say" (202/1076, 18.8%). These information needs were ranked and mapped to the relevant change objectives identified in step 2 (Multimedia Appendix 2). Chi-square tests revealed that the participants whose friends had not been diagnosed with cancer were more likely to require information concerning "survival rates of all types of cancer" (χ^2_1 =7.3; *P*=.007) compared with those whose friends had been diagnosed (Multimedia Appendix 3).

Preferable modes of delivery were booklets or leaflets (529/1076, 49.2%), the Internet (texts and figures: 460/1076, 42.8%; text only: 297/1076, 27.6%; text and video: 155/1076, 14.4%), apps (75/1076, 7%), DVDs (46/1076, 4%), and other (24/1076, 2%). There were no differences in these preferences between those participants whose friends had not been diagnosed with cancer and those whose friends had been diagnosed (Multimedia Appendix 4).

Needs Survey Among the Cancer Survivors

A total of 473 cancer survivors participated in the survey. The descriptive statistics showed that the ages ranged from 25 to 85 years, and that 49.7% (235/473) were men. The participants had primarily developed breast cancer (118/473, 24.9%), colorectal cancer (68/473, 14%), and prostate cancer (51/473, 11%), with most at the early stage (grades 0/I/II) of cancer (311/473, 65.8%) [45].

The 5 most important types of information that the cancer survivors wanted their friends to know included: "the possibility of a cure as a result of early detection and treatment" (193/473, 40.8%), "types of cancer treatment" (154/473, 32.6%), "survivors continuing their social life during/after cancer treatment" (130/473, 27.5%), "outpatient cancer treatment" (116/473, 24.5%), and "risk factors of cancer" (114/473, 24.1%). These information needs were categorized into the relevant change objectives identified in step 2 (Multimedia Appendix 5).

Setting Up the Theoretical and Practical Change Methods

According to the CDSR findings and the 2 surveys, the planning team decided to develop a multiple-component and web-based (texts and figures) intervention program. The contents of the program were to include information about interpersonal skills, cancer and its treatments, and cancer survivors' social life, which would meet the needs of both the public and cancer survivors. Referring to the two reviewed interventions in the CDSR search [51,52], the structure and sequences would have 5 thematic components, in the following order: (1) cancer, treatment, and survivors; (2) emotional and cognitive reactions to hypothetical friends with cancer; (3) reasons for cancer survivors' illness disclosure to friends and expected responses from them; (4) guiding principles and role-plays on how to listen to make survivors feel safe; and (5) survivors' desire for relationships with and support from their friends.

Several behavioral change methods that matched the determinants and change objectives were selected. Regarding knowledge (increasing accurate knowledge of cancer and cancer survivors), details on stereotype-inconsistent information were derived from stigma and discrimination theory [53], and details on framing were derived from protection motivation theory

XSL•FO
RenderX

[54,55]. Regarding skills (acquiring empathetic coping strategies to use when told about hypothetical friends' cancer diagnosis), details on empathy training were derived from stigma and discrimination theory [56], details on modeling from social cognitive theory [57], and details on feedback from learning theories [58]. To facilitate understanding of how theoretical and practical change methods were selected and translated into the program components, a comprehensive table is presented in Multimedia Appendix 6. The planning team agreed to these practical change methods.

Step 4: Program Materials

Internal Material Production and Refinements

After individually reviewing all the draft materials, the planning team observed notable variation in both the quantity and presentation styles across the 5 modules. Through discussions, the planning team decided to standardize the duration of each module to within 30 minutes and to clearly state the learning objectives at the beginning of each module to help program participants understand the intended goals.

Multidisciplinary Experts' Inputs and Refinements

The multidisciplinary expert panel reported some disagreements concerning the validity of the intervention logic model and program materials. Regarding the logic model, disagreements were reported on "concrete descriptions of the activities" (n=1), "the impact does not deviate from the scope of the stated outcome of the activity" (n=1), "the approximate time is clearly stated in which program participants would complete each module of the activity" (n=3), and "the time until all the activities would be completed is clearly stated" (n=3). We added the estimated time to complete each module and the total duration required to complete the program.

Regarding the materials, disagreements were reported on the "appropriateness of content" (n=3), "appropriateness of sentence expression" (n=2), "cost effectiveness" (n=5), "usefulness for short-term outcomes" (n=2), "appropriateness of audio recording with case stories" (n=1), "appropriateness of the use of worksheets" (n=2), and "appropriateness of feedback for worksheets" (n=2).

Most disagreements of the expert panel concerned participant burden in relation to the program. The planning team prioritized the expert panel's suggestions concerning addressing matters that might adversely affect the program participants' learning motivation and understanding. Consequently, the volume of materials was reduced by deleting unnecessary quizzes and worksheets, and voice recordings of the case stories. Terminology (eg, the 5-y survival rate) was clarified by adding explanations, and layouts were refined by separating the contents into 2 slides. The planning team discussed the refined materials and reached a consensus.

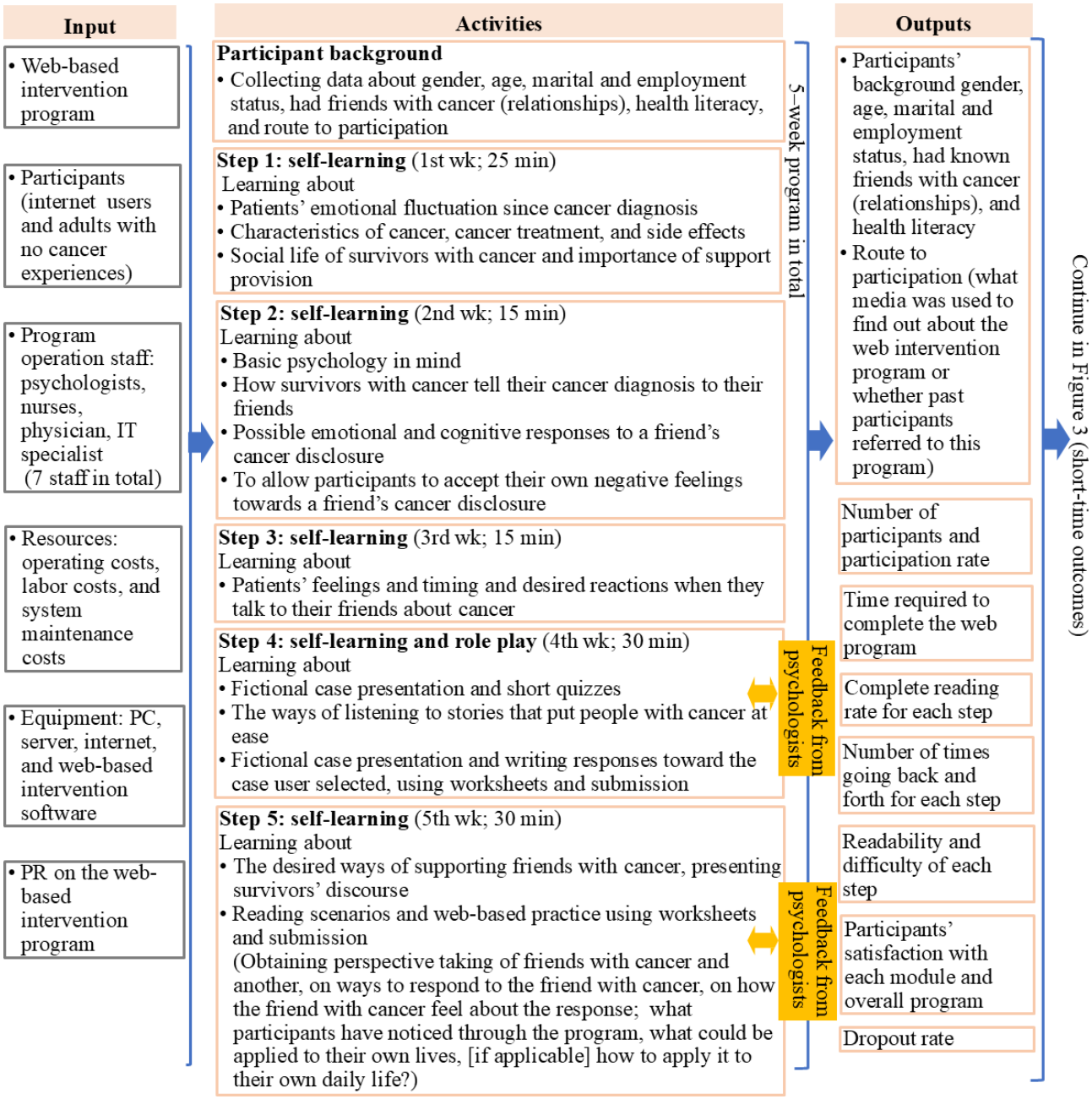
Pilot Testing and Refinements

A total of 13 adults participated in the think-aloud session. The descriptive statistics showed that the ages ranged from 20 to 63 years. The program's visualization, content, quantity, and written feedback on worksheets were well received among the program participants. However, some older participants expressed a preference for printed or PDF materials to review in advance [50].

The interview data (n=8) [50] showed that some participants preferred to revisit previously studied modules freely and requested access to external links to the case stories to explore a wider range of survivors' experiences and situational contexts. Based on these findings, both printed and PDF materials were developed. The IT team added functions to allow program participants to freely browse complete modules. The planning team finalized the intervention logic model (the final version is shown in Figures 2 and 3) and developed a practical guide for program providers to support future implementation.

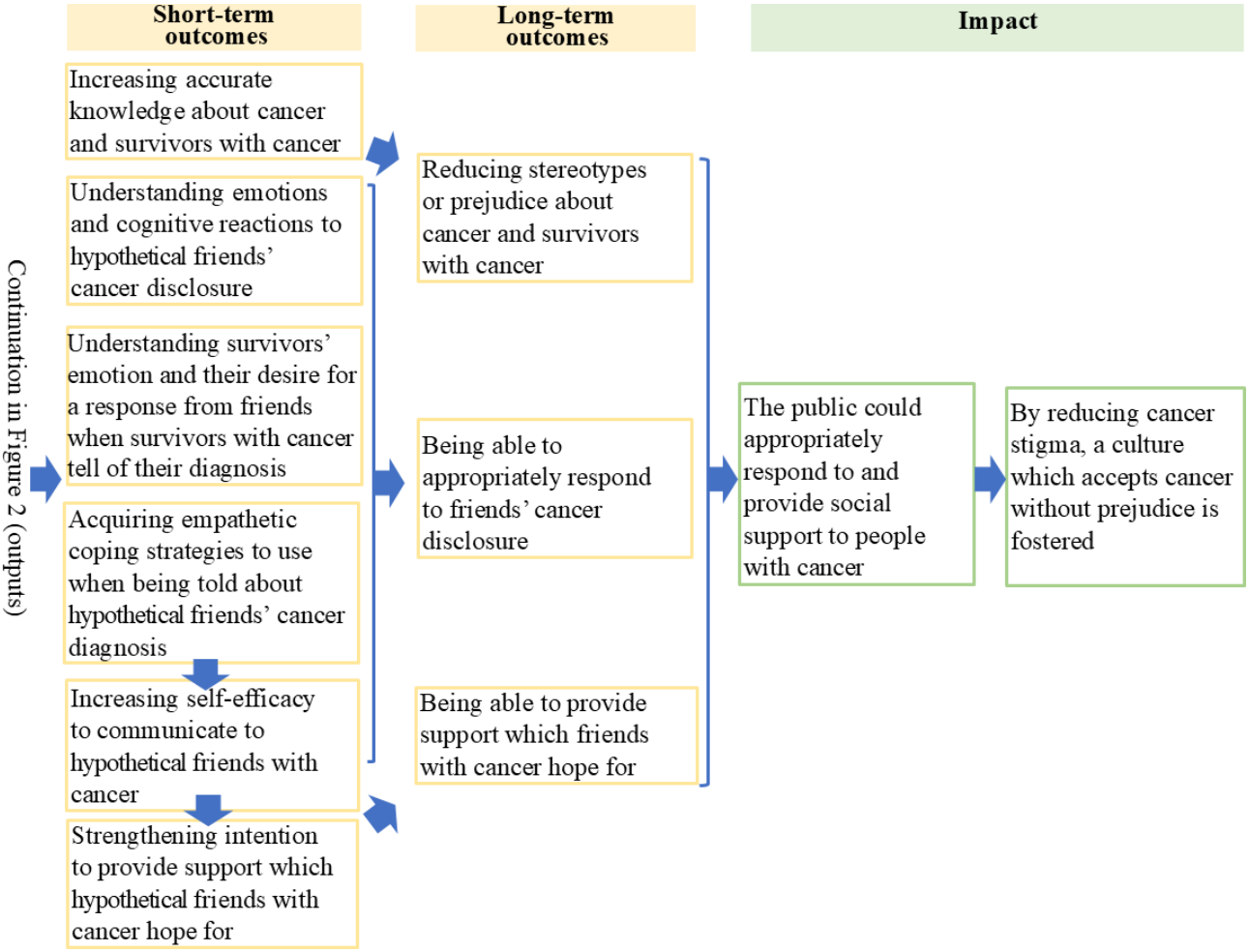
XSL•FO
RenderX

Figure 2. Intervention logic model: from input to output. PR: public relations.



XSL•FO
RenderX

Figure 3. Intervention logic model: from outcomes to impact.



Furthermore, the interview data showed that the intervention program was found to reduce participants' distancing attitudes and increase their intention to support hypothetical friends [50]. These promising results suggest the program's potential to enhance helping behaviors toward cancer survivors (the snapshot of the "Friend Support" is shown in Multimedia Appendix 7).

Discussion

Principal Findings

This study is the first to apply the IM approach to develop a web-based experiential learning program aimed at reducing cancer stigma among the public.

We set out and described the first four steps of the IM approach [30]. While adapting existing evidence-based intervention programs can be more efficient than creating new ones, there remains a notable lack of programs aiming to reduce cancer stigma [1,23]. Therefore, it was considered worthwhile to undertake and report on a step-by-step developmental process that integrates empirical findings, behavioral change theories, and practical input. The incorporation of expert panel input and user feedback through a think-aloud approach significantly enhanced content refinement, as well as the relevance and feasibility of the final product.

The intervention logic model allowed for the development of the hypothesis that the primary goal of this proposed intervention program would be achieved, namely, to increase helping intentions toward hypothetical friends with cancer. We set up multiple mediators: increasing knowledge, empathetic skills, and self-efficacy. Individuals with heightened helping intentions are more likely to respond supportively when someone discloses a cancer diagnosis. Over time, such attitudinal changes may contribute to the three specified long-term outcomes: improved interpersonal responsiveness, increased support provision, and reduced stigma. If supportive behaviors become widespread across interpersonal, organizational, and community levels, the subsequent effects are more likely to foster a culture of inclusion, one that accepts cancer without prejudice and reduces discrimination against cancer survivors. While QoL was not explicitly defined as a direct outcome in the logic model, it serves as a foundational construct that underpins the broader objectives of the program. By fostering supportive environments and reducing stigma, the program is expected to indirectly contribute to enhancing the QoL of cancer survivors.

Comparison With Prior Work

Prior studies have shown that public stigma contributes to internalized stigma among cancer survivors [6-14,59]. This dynamic aligns with the socio-ecological framework, which recognizes the interplay between societal attitudes and individual experiences. Whereas most digital interventions using IM have focused on intrapersonal factors to support individuals with cancer [32,33], our program uniquely targets the public, situated

XSL•FO
RenderX

in the outer layer of the socio-ecological model, to promote broader cultural and relational change. Similar logic has been applied in intervention development targeting parental behavior, with the intention of indirectly influencing children's physical activity levels by modifying relational dynamics [60].

Previous stigma reduction efforts have largely emphasized knowledge acquisition, including exposure to patient narratives [18-22,24]. However, social cognitive theory [61] suggests that knowledge alone may be insufficient to drive behavioral change; attitudes toward behavioral intentions are also essential components in predicting supportive actions toward individuals with cancer. Our experiential learning integrates empathy training together with cancer-related knowledge to foster emotional and cognitive empathy, and to enhance self-efficacy and helping intentions toward cancer survivors.

Unlike prior stigma studies that focused on specific cancer types [4,6,8,9,11-13,15,59], our program addresses stigma toward cancer survivors more broadly and is applicable to any kind of relationship. Centering on relational challenges and support needs common to diverse cancer contexts [29], the materials were intentionally designed for wide applicability. In addition to fostering transferable interpersonal skills, the program also equips participants with foundational knowledge about cancer, treatment processes, and the social lives of patients, namely survivorship. The program's modular design allows for the inclusion of supplementary materials tailored to specific cancer types, enhancing its adaptability across diverse contexts.

Strength and Limitations

This project demonstrates several key strengths in relation to the development of a web-based experiential learning intervention for cancer stigma reduction. First, the use of the IM framework enabled a systematic integration of theory, empirical evidence, and practical input. A multidisciplinary planning team guided the process through structured consensus-building, ensuring that diverse perspectives were reflected and that certain professional opinions were not overly influential. Second, the logic model helped identify and target essential mediators—knowledge, skills, and self-efficacy—within the program materials. The public survey data further informed the content and delivery preferences, aligning the intervention with the needs of both the public and cancer survivors. Appropriate reflection on the public's needs, together with cancer survivors' needs, may help motivate participation, implementation, and dissemination [43]. Third, the inclusion of communication-focused components and empathy training was undertaken in response to stigma-related challenges identified in prior research [56]. Various techniques, such as modeling, feedback, and first-person narratives [47], were incorporated to enhance self-efficacy and promote prosocial attitudes. The clinical psychologists' written feedback on worksheets may provide further assurance and facilitate prosocial attitudes. Finally, expert reviews and end user feedback contributed to the credibility, feasibility, and relevance of the program, increasing the likelihood of achieving its performance objectives and supporting future implementation.

Regarding limitations, the investigation of the research on cancer stigma was limited to creating a logic model of the problem. However, recent findings suggest that public stigma may intersect with structural factors such as education and employment [10,62-64]. Future studies should address structural dimensions of stigma, which were not explicitly targeted in the current program. Further, we used the end user approach. However, while we did not invite cancer survivors into the expert panel, we compensated for this limitation by using the findings of the needs survey among cancer survivors. Finally, we did not establish a specific performance objective or change objectives for the environmental impact. In the future, appropriate change methods and practical applications should be identified and combined.

Practical Implications

We plan to assess the effectiveness of this web-based intervention program, "Friend Support," quantitatively, using randomized clinical trials among the public. If its effectiveness is confirmed, implementation studies could be performed in the future using steps 5 and 6 of the IM approach. To facilitate dissemination, intervention packages, including materials for users, a practical guide for program providers, and feedback training workshops for researchers and clinicians, have been prepared. A developed practical guide would be helpful for staff training, not only in terms of providing feedback on worksheets but also for understanding how to deliver the intervention program effectively.

For future implementation, collaboration with schools, companies, and community-based organizations, such as nongovernmental organizations or civil groups, could support localization and sustainable delivery. While the current program is framed around the role of a friend, extending this application to other social roles, such as workplace colleagues or teachers, may enhance its relevance. Core components of the program, particularly those related to interpersonal empathy and supportive communication, can be retained while adapting delivery methods to suit different contexts. For instance, incorporating reflective exercises that prompt program participants to consider how they might respond if a coworker or student were diagnosed with cancer could deepen empathy and foster supportive engagement in professional or institutional settings. These adaptations may also help address stigma in environments where social dynamics are shaped by hierarchical or role-based interactions.

The intervention has the potential to serve as a long-term resource for the public, promoting supportive attitudes and behaviors toward cancer survivors. Although the program was developed in response to Japan's legal emphasis on stigma reduction in cancer survivorship, future research may explore whether stigma reduction can also encourage preventive behaviors, such as cancer screening and health care-seeking attitudes and behaviors [4,5,8,13].

Conclusions

IM is a useful framework for integrating theory, evidence, and practical inputs to develop a web-based experiential learning intervention program designed for the public to reduce cancer stigma. A qualitative pilot test demonstrated the program's feasibility, but further evaluation involving the public is required

XSL•FO
RenderX

before implementation can be considered. We hope this tutorial paper will help researchers and educators systematically develop similar web-based experiential learning interventions, particularly those using simulation to address stigma and promote an inclusive and supportive community.

Acknowledgments

The authors give their thanks to Ms Risako Fujita, who, as a multidisciplinary planning team member, undertook the data preparation and analysis of the 2 surveys.

Funding

This study was supported by JSPS KAKENHI (Grant Number JP16H05216) and Mitsubishi Foundation Grants for Social Welfare Activities. The funding bodies were not involved in the study design, data collection, data analysis, interpretation, writing of the report, or submission of the manuscript for publication.

Data Availability

The data are not available as they are not held in a public archive owing to participant privacy reasons.

Authors' Contributions

Conceptualization: MT (lead), KA (supporting), KK (supporting), AK (supporting)
Data curation: MT (lead), AK (supporting)
Formal analysis: MT (lead), KA (supporting), AK (supporting), KK (supporting)
Investigation: MT (lead), AK (supporting), KA (supporting), KK (supporting)
Methodology: MT (lead), KA (supporting), AK (supporting), KK (supporting)
Visualization: MT (lead), AK (supporting), KK (supporting), KA (supporting)
Writing – original draft: MT (lead), KA (supporting), AK (supporting), KK (supporting)
Writing – review & editing: MT (lead), all authors (supporting)
Supervision: MT (lead), KA (lead)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Matrix of the performance objectives and change objectives by personal determinants.
[PDF File, 117 KB - cancer_v12i1e71166_app1.pdf]

Multimedia Appendix 2
Information priorities in the public and corresponding change objectives (n=1076).
[PDF File, 77 KB - cancer_v12i1e71166_app2.pdf]

Multimedia Appendix 3
Information needs of the public in relation to having friends with cancer (n=1076).
[PDF File, 92 KB - cancer_v12i1e71166_app3.pdf]

Multimedia Appendix 4
Preferred mode of information delivery among the public in relation to having friends with cancer (n=1076).
[PDF File, 109 KB - cancer_v12i1e71166_app4.pdf]

Multimedia Appendix 5
Survivors' knowledge priorities for friends and corresponding change objectives (n=473).
[PDF File, 107 KB - cancer_v12i1e71166_app5.pdf]

Multimedia Appendix 6
Theoretical methods and practical applications for change objectives and program components.
[PDF File, 120 KB - cancer_v12i1e71166_app6.pdf]

XSL•FO
RenderX

Multimedia Appendix 7
Snapshot of “Friends Supporter.”
[PDF File, 961 KB - cancer_v12i1e71166_app7.pdf]

References

1. Heley K, Vedham V, Vanderpool RC. The global landscape of cancer stigma research and practice: perspectives, measures, interventions, and actions for change. J Natl Cancer Inst Monogr 2024 Jun 5;2024(63):1-3. [doi: 10.1093/jncimonographs/lgae009] [Medline: 38836526]
2. Link BG, Phelan JC. Conceptualizing stigma. Annu Rev Sociol 2001 Aug;27(1):363-385. [doi: 10.1146/annurev.soc.27.1.363]
3. Larkin D, Birtle AJ, Bradley L, et al. A systematic review of disease related stigmatization in patients living with prostate cancer. PLoS One 2022;17(2):e0261557. [doi: 10.1371/journal.pone.0261557] [Medline: 35148315]
4. Chambers SK, Dunn J, Occhipinti S, et al. A systematic review of the impact of stigma and nihilism on lung cancer outcomes. BMC Cancer 2012 May 20;12(1):184. [doi: 10.1186/1471-2407-12-184] [Medline: 22607085]
5. Vrinten C, Gallagher A, Waller J, Marlow LAV. Cancer stigma and cancer screening attendance: a population based survey in England. BMC Cancer 2019 Jun 11;19(1):566. [doi: 10.1186/s12885-019-5787-x] [Medline: 31185949]
6. Kang NE, Kim HY, Kim JY, Kim SR. Relationship between cancer stigma, social support, coping strategies and psychosocial adjustment among breast cancer survivors. J Clin Nurs 2020 Nov;29(21-22):4368-4378. [doi: 10.1111/jocn.15475] [Medline: 32860289]
7. Warner ET, Park ER, Luberto CM, Rabin J, Perez GK, Ostroff JS. Internalized stigma among cancer patients enrolled in a smoking cessation trial: The role of cancer type and associations with psychological distress. Psychooncology 2022 May;31(5):753-760. [doi: 10.1002/pon.5859] [Medline: 34797953]
8. Tang WZ, Yusuf A, Jia K, et al. Correlates of stigma for patients with breast cancer: a systematic review and meta-analysis. Support Care Cancer 2023 Jan;31(1):55. [doi: 10.1007/s00520-022-07506-4] [Medline: 36526859]
9. Otsuka T, Majima T. The stigma process experienced by women diagnosed with cervical precancerous lesions [Article in Japanese]. J Chiba Acad Nurs Sci 2024;29:1-10. [doi: 10.1016/S0021-5198(19)66838-1]
10. Shim S, Kang D, Bae KR, et al. Association between cancer stigma and job loss among cancer survivors. Psychooncology 2021 Aug;30(8):1347-1355. [doi: 10.1002/pon.5690] [Medline: 33817907]
11. Liu XH, Zhong JD, Zhang JE, Cheng Y, Bu XQ. Stigma and its correlates in people living with lung cancer: a cross-sectional study from China. Psychooncology 2020 Feb;29(2):287-293. [doi: 10.1002/pon.5245] [Medline: 31654596]
12. Melhem SJ, Nabhani-Gebara S, Kayyali R. Latency of breast cancer stigma during survivorship and its influencing factors: a qualitative study. Front Oncol 2023;13:1075298. [doi: 10.3389/fonc.2023.1075298] [Medline: 36998442]
13. Hamann HA, Ver Hoeve ES, Carter-Harris L, Studts JL, Ostroff JS. Multilevel opportunities to address lung cancer stigma across the cancer control continuum. J Thorac Oncol 2018 Aug;13(8):1062-1075. [doi: 10.1016/j.jtho.2018.05.014] [Medline: 29800746]
14. Fujisawa D, Umezawa S, Fujimori M, Miyashita M. Prevalence and associated factors of perceived cancer-related stigma in Japanese cancer survivors. Jpn J Clin Oncol 2020 Oct 22;50(11):1325-1329. [doi: 10.1093/jjco/hyaa135] [Medline: 32776092]
15. Pan S, Wang L, Zheng L, et al. Effects of stigma, anxiety and depression, and uncertainty in illness on quality of life in patients with prostate cancer: a cross-sectional analysis. BMC Psychol 2023 Apr 25;11(1):129. [doi: 10.1186/s40359-023-01159-6] [Medline: 37098648]
16. The Fourth Basic Plan for Cancer Control [Report in Japanese]. : Ministry of Health Labour and Welfare; 2023 URL: https://www.mhlw.go.jp/content/10900000/001522768.pdf [accessed 2025-08-20]
17. Stangl AL, Earnshaw VA, Logie CH, et al. The health stigma and discrimination framework: a global, crosscutting framework to inform research, intervention development, and policy on health-related stigmas. BMC Med 2019 Feb 15;17(1):31. [doi: 10.1186/s12916-019-1271-3] [Medline: 30764826]
18. Gronholm PC, Henderson C, Deb T, Thornicroft G. Interventions to reduce discrimination and stigma: the state of the art. Soc Psychiatry Psychiatr Epidemiol 2017 Mar;52(3):249-258. [doi: 10.1007/s00127-017-1341-9] [Medline: 28144713]
19. Kemp CG, Jarrett BA, Kwon CS, et al. Implementation science and stigma reduction interventions in low- and middle-income countries: a systematic review. BMC Med 2019 Feb 15;17(1):6. [doi: 10.1186/s12916-018-1237-x] [Medline: 30764820]
20. Rao D, Elshafei A, Nguyen M, Hatzenbuehler ML, Frey S, Go VF. A systematic review of multi-level stigma interventions: state of the science and future directions. BMC Med 2019 Feb 15;17(1):41. [doi: 10.1186/s12916-018-1244-y] [Medline: 30770756]
21. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15(4):351-377. [doi: 10.1177/109019818801500401] [Medline: 3068205]
22. Livingston JD, Milne T, Fang ML, Amari E. The effectiveness of interventions for reducing stigma related to substance use disorders: a systematic review. Addiction 2012 Jan;107(1):39-50. [doi: 10.1111/j.1360-0443.2011.03601.x] [Medline: 21815959]

XSL•FO
RenderX

23. Yamazaki-Tan J, Harrison NJ, Marshall H, Gartner C, Runge CE, Morphett K. Interventions to reduce lung cancer and COPD-related stigma: a systematic review. Ann Behav Med 2024 Oct 18;58(11):729-740. [doi: 10.1093/abm/kaae048] [Medline: 39197098]
24. Yoshida K, Matsui Y. A pilot study for the development and evaluation of an educational program to reduce stigma toward cancer and cancer survivors: focusing on dating and marriage after cancer diagnosis. J Cancer Educ 2022 Jun;37(3):806-811. [doi: 10.1007/s13187-020-01885-0] [Medline: 32980978]
25. Bradbury K, Steele M, Corbett T, et al. Developing a digital intervention for cancer survivors: an evidence-, theory- and person-based approach. NPJ Digit Med 2019;2:85. [doi: 10.1038/s41746-019-0163-4] [Medline: 31508496]
26. Amati V, Meggiolaro S, Rivellini G, Zaccarin S. Social relations and life satisfaction: the role of friends. Genus 2018;74(1):7. [doi: 10.1186/s41118-018-0032-z] [Medline: 29755134]
27. Lewis DWG, Al-Shawaf L, Russell EM, Buss DM. Friends and happiness: an evolutionary perspective on friendship. In: Demir M, editor. Friendship and Happiness: Across the Life-Span and Cultures: Springer; 2015:37-57. [doi: 10.1007/978-94-017-9603-3_3]
28. Paluck EL, Porat R, Clark CS, Green DP. Prejudice reduction: progress and challenges. Annu Rev Psychol 2021 Jan 4;72:533-560. [doi: 10.1146/annurev-psych-071620-030619] [Medline: 32928061]
29. Tsuchiya M, Tazaki M, Fujita R, Kodama S, Takata Y. A mixed-method systematic review of unmet care and support needs among Japanese cancer survivors. J Cancer Surviv 2024 Dec;18(6):2008-2021. [doi: 10.1007/s11764-023-01439-5] [Medline: 37531042]
30. Eldredge LKB, Markham CM, Ruiter RAC, Fernández ME, Kok G, Parcel GS. Planning Health Promotion Programs: An Intervention Mapping Approach, th edition: Jossey-Bass; 2016.
31. Hall KL, Oh A, Perez LG, et al. The ecology of multilevel intervention research. Transl Behav Med 2018 Nov 21;8(6):968-978. [doi: 10.1093/tbm/iby102] [Medline: 30428073]
32. Abahussin AA, West RM, Wong DC, Ziegler LE, Allsop MJ. Supporting pain self-management in patients with cancer: app development based on a theoretical and evidence-driven approach. JMIR Cancer 2023 Oct 9;9:e49471. [doi: 10.2196/49471] [Medline: 37812491]
33. Austin J, Drossaert CHC, van Dijk J, et al. Integrating top-down and bottom-up requirements in eHealth development: the case of a mobile self-compassion intervention for people with newly diagnosed cancer. JMIR Cancer 2022 Aug 1;8(3):e37502. [doi: 10.2196/37502] [Medline: 35916691]
34. Hadjiconstantinou M, Schreder S, Brough C, et al. Using intervention mapping to develop a digital self-management program for people with type 2 diabetes: tutorial on MyDESMOND. J Med Internet Res 2020 May 11;22(5):e17316. [doi: 10.2196/17316] [Medline: 32391797]
35. Tsuchiya M, Horn S, Ingham R. Social disclosure about lymphoedema symptoms: a qualitative study among Japanese breast cancer survivors. Psychol Health Med 2015;20(6):680-684. [doi: 10.1080/13548506.2014.986140] [Medline: 25482185]
36. Tsuchiya M, Horn S, Ingham R. Development of the psycho-social discomfort scale (PsDS): investigation of psychometric properties among Japanese breast cancer survivors. Psychooncology 2012 Feb;21(2):161-167. [doi: 10.1002/pon.1880] [Medline: 22271536]
37. Tsuchiya M. Lay beliefs, knowledge, and attitudes towards cancer: a pilot study in Japan. Asian Pac J Cancer Prev 2015;16(8):3247-3251. [doi: 10.7314/apjcp.2015.16.8.3247] [Medline: 25921127]
38. Tsuchiya M. Lay people's psychological reactions and helping intention after friends' cancer disclosure: an exploratory analysis using vignettes. Eur J Cancer Care (Engl) 2019 Nov;28(6):e13150. [doi: 10.1111/ecc.13150] [Medline: 31448849]
39. Fujisawa D, Hagiwara N. Cancer stigma and its health consequences. Curr Breast Cancer Rep 2015 Sep;7(3):143-150. [doi: 10.1007/s12609-015-0185-0]
40. Weiner B, Perry RP, Magnusson J. An attributional analysis of reactions to stigmas. J Pers Soc Psychol 1988 Nov;55(5):738-748. [doi: 10.1037//0022-3514.55.5.738] [Medline: 2974883]
41. Weiner B. Inferences of responsibility and social motivation. Adv Exp Soc Psychol 1995;27:1-47. [doi: 10.1016/S0065-2601(08)60402-5]
42. Delongis A, O'Brién T. An interpersonal framework for stress and coping: an application to the families of Alzheimer's patients. In: Stephens MA, Crowther JH, Hobfoll SE, Tennenbaum DL, editors. Stress and Coping in Later-Life Families: Hemisphere Publishing Corp; 1990:221-239. [doi: 10.4324/9781315803074-13]
43. Jorm AF, Kitchener BA, Reavley NJ. Mental Health First Aid training: lessons learned from the global spread of a community education program. World Psychiatry 2019 Jun;18(2):142-143. [doi: 10.1002/wps.20621] [Medline: 31059620]
44. Kimata A, Kumagai K, Kondo N, Adachi K, Fujita R, Tsuchiya M. Development and validation of the Cancer Knowledge Scale for the general population: An item response theory approach. Patient Educ Couns 2022 Jan;105(1):151-157. [doi: 10.1016/j.pec.2021.05.007] [Medline: 34020840]
45. Tsuchiya M, Adachi K, Kumagai K, Kondo N, Kimata A. Cancer disclosure to friends: Survey on psychological distress and perceived social support provision. Eur J Cancer Care 2022 Sep;31(5):e13332. [doi: 10.1111/ecc.13332] [Medline: 33006241]

XSL•FO
RenderX

46. Kok G, Gottlieb NH, Peters GJY, et al. A taxonomy of behaviour change methods: an Intervention Mapping approach. Health Psychol Rev 2016 Sep;10(3):297-312. [doi: 10.1080/17437199.2015.1077155] [Medline: 26262912]
47. Clement S, Lassman F, Barley E, et al. Mass media interventions for reducing mental health-related stigma. Cochrane Database Syst Rev 2013 Jul 23;2013(7):CD009453. [doi: 10.1002/14651858.CD009453.pub2] [Medline: 23881731]
48. Using logic models to bring together planning, evaluation, and action: logic model development guide. : Kellogg WK Foundation; 2004 URL: https://www.betterevaluation.org/sites/default/files/2021-11/Kellogg_Foundation_Logic_Model_Guide.pdf [accessed 2024-06-20]
49. Knowlton LW, Phillips CC. The Logic Model Guidebook: Better Strategies for Great Results: Sage Publications, Inc; 2013. URL: https://uk.sagepub.com/en-gb/eur/the-logic-model-guidebook/book237460 [accessed 2025-12-13]
50. Tsuchiya M, Fujita R, Adachi K, Kumagai K, Kimata A, Kondo N. Web-based experiential learning intervention Friend Supporter to reduce cancer stigma: Feasibility, usability, and acceptability trial. J Health Psychol Res 2023;35(2):133-145. [doi: 10.11560/jhpr.220714179]
51. Finkelstein J, Lapshin O, Wasserman E. Randomized study of different anti-stigma media. Patient Educ Couns 2008 May;71(2):204-214. [doi: 10.1016/j.pec.2008.01.002] [Medline: 18289823]
52. Jorm AF, Kitchener BA, Fischer JA, Cvetkovski S. Mental health first aid training by e-learning: a randomized controlled trial. Aust N Z J Psychiatry 2010 Dec;44(12):1072-1081. [doi: 10.3109/00048674.2010.516426] [Medline: 21070103]
53. Bos AER, Schaalma HP, Pryor JB. Reducing AIDS-related stigma in developing countries: the importance of theory- and evidence-based interventions. Psychol Health Med 2008 Aug;13(4):450-460. [doi: 10.1080/13548500701687171] [Medline: 18825583]
54. Van 't Riet J, Cox AD, Cox D, et al. Does perceived risk influence the effects of message framing? a new investigation of a widely held notion. Psychol Health 2014;29(8):933-949. [doi: 10.1080/08870446.2014.896916] [Medline: 24579986]
55. Werrij MQ, Ruiter RAC, Van't Riet J, De Vries H. Message framing. In: Writing Health Communication: An Evidence-Based Guide: Sage Publications Ltd; 2011:134-143. [doi: 10.4135/9781446288054.n9]
56. Batson CD, Chang J, Orr R, Rowland J. Empathy, attitudes, and action: can feeling for a member of a stigmatized group motivate one to help the group? Pers Soc Psychol Bull 2002 Dec;28(12):1656-1666. [doi: 10.1177/014616702237647]
57. Kelder SH, Hoelscher D, Perry CL. How individuals, environments, and health behavior interact: social cognitive theory. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior: Theory, Research, Practice, 5th edition: Jossey-Bass; 2015:159-182.
58. Kazdin AE. Behavior Modification in Applied Setting: Waveland Press Inc; 2013.
59. Coleman D, Hurtado-de-Mendoza A, Montero A, et al. Stigma, social support, and spirituality: associations with symptoms among Black, Latina, and Chinese American cervical cancer survivors. J Cancer Surviv 2024 Jun;18(3):710-726. [doi: 10.1007/s11764-022-01283-z] [Medline: 36417116]
60. Wilhite BC, Hennessy E, Economos C, et al. Content development for a multilevel intervention to operationalize behavior change theory and improve parenting behaviors for physical activity: descriptive study. JMIR Pediatr Parent 2025 Aug 26;8:e73510. [doi: 10.2196/73510] [Medline: 40856625]
61. Ajzen I, Driver BL. Prediction of leisure participation from behavioral, normative, and control beliefs: an application of the theory of planned behavior. Leis Sci 1991 Jan;13(3):185-204. [doi: 10.1080/01490409109513137]
62. Takata Y, Tsuchiya M, Tazaki M, Takahashi M. An explanatory study of childhood and adolescent/young adult cancer survivors' difficulties and information and support needs during job-seeking: focusing on Illness-related difficulties [Report in Japanese]. J Japan Acad Health Behav Sci 2018;33:29-38 [FREE Full text]
63. Tsuchiya M, Tazaki M, Takata Y, Takahashi M. Factors affecting the decision to disclose illness and responses from human resource professionals to survivors of childhood, adolescent, and young-adult cancer. Jpn J Pediatr Hematol Oncol 2019;56:189-197. [doi: 10.11412/jspho.56.189]
64. Tsuchiya M. Health labour sciences research grant (research for promotion of cancer control program) co-investigators' research report [Report in Japanese]. : Ministry of Health, Labour and Welfare; 2019 URL: https://mhlw-grants.niph.go.jp/system/files/report_pdf/202108009A-buntan4.pdf [accessed 2025-06-20]

Abbreviations

CDSR: The Cochrane Database of Systematic Reviews
IM: intervention mapping
QoL: quality of life

XSL•FO
RenderX

Edited by J Lau; submitted 11.Jan.2025; peer-reviewed by J Studts, K Yeary, L Wreyford; accepted 01.Dec.2025; published 27.Jan.2026.

Please cite as:
Tsuchiya M, Adachi K, Kimata A, Kumagai K
Development of a Web-Based Experiential Learning Intervention for the Public to Reduce Cancer Stigma: Tutorial on the Application of Intervention Mapping
JMIR Cancer 2026;12:e71166
URL: https://cancer.jmir.org/2026/1/e71166
doi:10.2196/71166

© Miyako Tsuchiya, Keiichiro Adachi, Akiko Kimata, Kaori Kumagai. Originally published in JMIR Cancer (https://cancer.jmir.org), 27.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on https://cancer.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX